This document summarizes the material on fabrication and falsification in my much longer paper, "Ethical and Data-Integrity Problems in the Second Lancet Survey of Mortality in Iraq." (http://personal.rhul.ac.uk/uhte/014/Ethics and Data Integrity_8_09_08.pdf)

1. Comparison of L2 with other sources (Section 3.4)

a. Aggregate violent-death figures.

Iraq Family Health Survey (IFHS) – **L2 exceeds the IFHS by a factor of 6.6** for almost exactly the same coverage period – about **500,000 additional violent deaths** in L2 when we compare like with like.

Iraq Living Conditions Survey (ILCS) – **L2 exceeds ILCS by more than factor of 3** through May, 2004.

Iraq Body Count Project (IBC) – **L2 exceeds IBC by a factor of 12.**

b. Geographical distribution of violent deaths – L2 compared to all three other sources

L2 has an extraordinary **concentration of deaths in central governorates surrounding Baghdad but proportionately much lower in Baghdad** compared to the other sources.

c. Timing

Divergence with other sources grows sharply over time. For the last 13 months, **L2 exceeds the (unadjusted) IFHS by a factor of 11 and IBC by a factor of 16.7**.

Geography and timing patters are consistent with a fabrication hypothesis:

  – L2 acknowledged that IBC has reasonable coverage in Baghdad - deaths are packed into other areas (besides Baghdad) known to be dangerous and plausibly less covered by the media.
  – L2 is less out of line with other sources during the coverage periods of the ILCS and L1 – deaths mushroom in L2 after the discipline of these prior surveys disappears.

2. Cluster 33 (Section 3.5)

24 out of 66 deaths in a particular car bombing turn up in the L2 sample – 12 households with 1 death each and 6 households with 2 deaths each.

Pictures and descriptions of this bombing show that it ripped through a crowded marketplace and did not damage a row of houses (If lots of people had been killed inside

their homes that could explain how so many victims could live in a string of contiguous households.)

Were 24 neighbors out shopping together while all holding hands?

L2 reports that interviewers **did not try to confirm any of these deaths through death certificates**.

3. Cluster 34 (**Section 3.7**) - (The story here is fairly complicated: not easy to master.)

a. **Unusually violent cluster** (Always bear in mind that we are talking about just 40 contiguous households)

> - 35 violent deaths – the next most violent cluster has only 25 violent deaths.
> - At least 17 separate incidents affecting 22 households
> - **At least 18 deaths from air strikes in at least 5 incidents affecting 7 households (there is no death-certificate confirmation on any of these deaths)**
> - **At least 5 deaths in at least 4 separate bombings (at least 3 car bombings) affecting 5 separate households. This by itself cluster make cluster 34 even more likely than cluster 33 to contain fabricated data. What are the odds that victims of at least 3 separate car bombings would live so close together?**
> - A further 10 shootings, a 3-year-old crushed by a US military vehicle and a death in a "fight".

b. **The corroborating evidence scant** – Since women and children, and potentially compensation from the US government, are involved there are strong motivations for people to get these deaths confirmed.

> - L2 reports that there are only 22 violent deaths in the entire L2 sample for which interviewers asked for deaths certificates and none were produced – **19 out of these 22 cases are in Cluster 34.** In a further 7 cases in Cluster 34, L2 reports that the field teams did not even ask for death-certificate confirmation. **Only 9 out or 35 violent deaths in this cluster are reported to be confirmed by death certificate.**
> - Little media corroboration for either air strikes or car bombs – **IBC never has an air strike in this governorate (Nineveh) in the same month as any of the claimed air-strike deaths** in Cluster 34 - IBC does not have a car bomb in Nineveh for the month of 1 of the claimed car-bomb deaths.

c. **There is a strong tilt against the US consistent with the hypothesis that interviewers and/or respondents were fabricating deaths to make the US look bad.**

> - **24 of the 35 violent deaths are attributed to US military**
> - In Cluster 34 the **US** is claimed to have been **much dirtier both than other agents and than the US everywhere else in Iraq** – the US is claimed to have

killed 9 out of the 10 **females (90%)** and 12 out of the 12 **children (100%)** in Cluster 34 compared to 3 out of 18 **females (17%)** and 7 out of 14 **children (50%)** everywhere else in Iraq according to the L2 data.

d. **There is string of 9 household ID's out of 11 reporting killings by US**, suggesting possible manipulation of survey.

4. Irregular death-certificate patterns **(Section 3.6)**

a. **Death certificates not requested much more frequently for violent deaths than for non-violent deaths.** Maybe interviewers often did not ask for death certificates when they thought that deaths were being fabricated.

b. **Patterns of confirmation through death certificates when asked very different for non-violent deaths than for violent deaths**.

- Failures to confirm for non-violent deaths are spread over 8 governorates
- All failures to confirm for violent deaths are in a single governorate, Nineveh (of these 19 out of 22 are in Cluster 34)

c. **The run of 180 confirmations in 180 attempts outside Nineveh is very improbable** judged against various standards such as:

- the L1 death-certificate confirmation record.
- the L2 death-certificate confirmation record for non-violent deaths.

d. **The L2 death-certificate results rely on a very delicate balancing act:**

- on the one hand, authorities must issue death certificates very reliably and households are able to produce them when asked.
- on the other hand, the central authorities must keep copies of these issuances only rarely or the central authorities are engaged in a massive fraud to hide the true number of deaths for which they have issued certificates.

e. Interview times for L2 were so short (40 interviews in a day) that it is **hard to believe that respondents could produce death certificates so quickly even if they had them.**

5. Claimed work schedule is improbable (**Section 3.3**)

I mainly refer to **the paper of Madelyn Hicks** that goes through in some detail what the interviewers needed to accomplish and the time frame they had to accomplish these things.

- **40 interviews in a day** – L2 authors have claimed that teams of 4 broke into sub-teams of 2, each doing approximately 20 interviews in a day lasting 20 minutes on average (although there is no evidence that field teams recorded start and end times so this 20-minute figure seems to have been pulled out of thin air)
- One must consider standards of hospitality in the Middle East – interviewers might easily be **invited in for tea**, etc.
- Interviewers would need to **build trust, obtain informed consent, give people chances to tell sensitive stories**, etc.

I extend the Hicks analysis to consider:

- the time it would take to **finish the sampling procedures (Section 2, pages 9-14)** at each locality which considerable.
- the logistics of moving from cluster to cluster - **52 clusters of interviews in 52 days, traveling all over the country on bad roads in an insecure environment.**

6. Improbable response rates and lack of security problems in moving from cluster to cluster **(Section 2, pages 14 – 17)**

a. **Failure to make contact with selected households 0.9% of the time seems unreal, especially considering the extreme time pressure on the field teams (40 interviews in a day).** A number of Iraq surveys do report very high contact rates but a no-contact rate of, for example, 3.4% as reported by the IFHS, is statistically very different from L2's 0.9% no-contact rate. **I compute very low probabilities for the L2 contact rate using a variety of contact probabilities based on other surveys.** For example, assuming that the IFHS rate is the true one and that the success or failure of visits are statistically independent of each other, **the odds against the L2 no-contact rate are more than 500,000 to 1 against.**

b. The IFHS reports a number of failed cluster visits due to security problems. These failures are frequent in Baghdad (31 out of 96) and Anbar (71 out of 108). L2 reports just one failed visit, strangely in relatively peaceful Wassit. Assuming that the IFHS experience reflects the reality, **the near-perfect record claimed by L2 is extremely improbable**. The odds against 12 successes in 12 attempts, as claimed by L2, **in Baghdad are 108 to 1 against** and the odds against 5 successes in 5 attempts **in Anbar are 214 to 1 against**.

There is **a connection with the independently suspicious Cluster 33 which was in Baghdad** – a particular suspicious cluster in Baghdad combined with an improbable claim to have surveyed 12 clusters in Baghdad without every having a failed visit.

7. Presence of a large number of factors that have been identified by survey researchers as risk factors for fabrication **(Section 3.2)**

I use a document of the American Association for Public Opinion Research and the American Statistical Association that give a nice check list on risk factors. Of these, particularly serious factors include:

a. **Inadequate supervision:**

- no supervision by the US-based authors who were never in Iraq when the survey was conducted.
- no use of recontact methods either to check that interviews were actually done or to check that interviews were done properly
- no collection of demographic information and, therefore, no use of such information to check the veracity of collected data
- no analysis of data matching annonomized interviewer IDs with interview results – the US based authors never even had access to such data.

b. **Poor quality control** – For example:

- Data that are sometimes missing include household sizes (13 times), months in which deaths occurred (57 times), and the number of males and females in each household (55 times).
- Data on the number of males and females in each household are missing for an entire cluster.
- 14% of the time, failure of the identity,

Household size 2006 = Household size 2002 + births – deaths + in migration – out migration

c. **Excessive workload** – this factor is extreme.

d. **Offsite isolation of interviewers from the parent organization**

This material reinforces most of the other arguments in the paper. Interviewers were working within an environment that was very conducive to fabrication. They then produced data that show a number of signs of fabrication.

More concretely, poorly supervised field workers under extreme time pressure reported, improbably, 12 successful cluster visits in 12 attempts in Baghdad. In one of these clusters, Cluster 33, they reported 24 victims of car bombings. They did not attempt to verify any of these deaths through death certificates. **A variety of independent factors all point in the direction of fabrication.**

8. Fabrication and distortion of related evidence and non-evidence **(Section 3.8)**

This section is a shocker. It covers a long string of claims made in the paper that are simply not true, all going in the direction of "confirmation" of L2.

This section is particularly important because it **documents multiple instances of information falsification.** This section alone **should be sufficient to prompt a withdrawal of the paper.** The pattern documented in this section also **reinforces doubts about the L2 data itself.**

Highlights include:

a. Terrible distortion of figures from the US Department of Defense (DOD) (pages 39, 43-44)

- **Misrepresents casualties (killings plus injuries) as killings**
- **Misrepresents casualties of civilians and combatants as killings of civilians.**
- **Inflates the DOD figure by 67%** through manipulation of time periods
- **Creates a DOD figure for a time period not covered by the data**
- **Compares the growth in DOD** *cumulative* **figures with L2 rates to create an illusion that L2 trends match DOD trends.**

b. **Compares IBC cumulative figures with L2 rates to create a false impression that IBC trends match L2 trends (pages 43-44)**.
.
c. **Ignores many pieces of evidence**, most importantly the **ILCS**, that cast doubt on L2. I include the monitoring of the UN for 2006 (UNAMI) and the Iraqi Ministry of Health (MoH) as sources that at least need to be considered. **L2 exceeds UNAMI by a factor of 12 and the MoH by a factor of 15 (page 40).**

d. **Misleads readers into thinking that Interior Ministry figures are 75% higher than IBC** figures, and hence IBC can be dismissed, when in reality IBC figures are 50% higher than Interior Ministry figures **(page 40).**

e. Launches another **spurious assault on IBC,** claiming that **seven separate studies** all show that the "passive surveillance" methods of IBC undercount conflict deaths by a very wide margin. **None of these citations actually support the claim (pages 45-47).**

f. **Fails to mention that three studies, including their own L1 and one by colleagues at Johns Hopkins, suggesting that IBCs methods capture a much higher fraction of violent deaths in Iraq than L2 claims (page 47**).

g. **Cites as confirming evidence two "surveys" that seem to be nothing more than internet rumors**- The "Iraqiyun" and "People's Freedom Party" surveys **(pages 41 and 42).**

h. **Claims, misleadingly, that L1 and L2 confirm each other (page 40).**

9. Non disclosure of basic parts of the methodology that no survey with scientific pretensions can withhold. (**Section 2, pages 6-19**)

Important examples of non-disclosed information include:

a. **Exact wordings of questions, i.e., a questionnaire.**

b. **Details of the sample design** - I document a long string of contradictory statements on how the sampling has been done.

c. **Data matching annonomized interviewer IDs with survey results.**

Again, **these instances of non-disclosure enhance the importance of other evidence of fabrication and falsification.** A number of suspicious patterns have been identified yet the L2 researchers refuse to disclose information that might shed light on these patterns.

10. The three point graphic (**Section 3.1**)

This is evidence that the figure of 600,000 excess deaths was extrapolated from earlier studies. Here are the key points:

a. In a letter to the *Lancet* the L2 authors cited a study done by Les Roberts and others in the DRC that found that 1.8% of the population was violently killed in 32 months (33 months according to the letter and even the cited paper sometimes says 33 months but the correct figure is 32 months.) This is cited as making it plausible that 2.3% of the population in Iraq might have been violently killed in 40 months of conflict as claimed by L2.

b. The L2 authors frequently cite, including in L2, a Kosovo finding that 0.8% of the population was in 17 months of conflict.

c. **It turns out that these three data points [(17, 0.8), (32, 1.8), (40, 2.3)] are in near-perfect alignment.** (This is robust to reasonable variation. For example, little changes if you replace the true figure of 32 with the L2 authors' figure of 33).

d. Mark van der Laan calculated **the probability of such alignment is 0.036, i.e., odds against of 28 to 1.**

This graphic is statistical evidence of falsification by the L2 authors themselves but the odds against the alignment are not high enough for this to be definitive.

11. Political motivations (not in the paper)

a. Les Roberts was running for Congress on an anti-war platform.

b. Both L1 and L2 came out just before US elections and both Roberts and Burnham have acknowledged political timing of both (although Burnham now denies that he said this)

c. Soros funding was not disclosed (Soros funding is (wrongly) viewed as completely discrediting by many conservatives in the US)

Political motivations should not be, but often are, treated as automatically discrediting. This is why I left them out of my paper. But in a full discussion **it is clear that political motivations reinforce other evidence of fabrication and falsification.**