Reviewer's report

Title: Evaluating the Risk of Patient Re-identification from Adverse Drug Event Reports

Version: 1 Date: 27 May 2013

Reviewer: Traian Truta

Reviewer's report:

The paper “Evaluating the Risk of Patient Re-Identification from Adverse Drug Event Reports" presents an in-depth analysis of how re-identification can occur when a broadcaster matches the available ADE report with an obituary database.

The paper has many strengths: the problem is well formulated and motivated, the re-identification risk is well introduced and the experimental section shows the validity and the usefulness of the approach.

However, the reviewer has several questions for the authors and several suggestions for improvements.

1. Regarding, Lemma 1, the proof is incomplete, and moreover, the last formula from the proof does not match (for n=2) the last branch of the lemma. In addition to that, I disagree with the second claim from page 10, this is true only if p=1, if p<1 the last unverified record may not be the correct match because of a mismatch in an earlier step.

2. In Section 4.4, the authors take the “conservative” approach with respect to the amount of external information known by the intruder/broadcaster. However the use of p (with a value < 1) is a type of optimistic approach. The “conservative” approach will be to use p=1. Considering that in reality some match verification will have probability 1, it is not clear why the authors underestimate this probability. To the minimum the authors should use p=1 in their experiments (an alternative will be to use p=1 in the entire paper).

3. In Table 4, the reported results seem to be wrong. The first row has the equivalence class size 253, while the row that adds Province in its quasi-identifiers has a larger minimum equivalence class size (263). Please correct. (Table 3 does not have this error).

Some minor errors:

- On page 4, one reference is missing (it is shown as [?])

- On Section 7, FKD should be FD (to match the name of the author as listed in the beginning).

Level of interest: An article of importance in its field
Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

'I declare that I have no competing interests'