Reviewer's report

Title: Managing Re-identification Risk for Small Geographic Areas

Version: 1 Date: 30 August 2009

Reviewer: Christovam Barcellos

Reviewer's report:

The manuscript examines the risk of disclosure of health data in small areas and suggests general criteria for data dissemination, considering this risk.

The theme is highly relevant to public health, once there is a growing use of geospatial data and tools for disease surveillance and health service location planning.

The text is clear and states the major ethical and technical issues involved in data dissemination in small areas. The methodology is adequate to evaluate the risk of re-identification of based on uniqueness criteria.

Major Compulsory Revisions

The model results show that this risk depends on the population size and the combinations of variables (quasi-identifiers). Therefore, it will be useful for readers outside Canada to know basic information about FSA: Median population and median area (km2), as well as their variation range (max, min, quartiles). The authors did not consider rural areas in the analysis. However, it would be useful to know what is the proportion of rural FSA among all national FSA. Is the FSA boundaries perfectly nested to census tracts?

The authors indicate that data suppression or aggregation would be the only alternatives avoiding the disclosure of individual records in small areas (e.g., pg 4, pg 14). Since this risk is a function of the number and combinations of quasi-identifiers, an additional alternative should be to restrict the availability of variables.

The model considers area population size, but the effect of population mixture should be also discussed in theory. For instance, the probability of re-identification in small areas inhabited by minorities is supposed to be higher than in more homogeneous areas?

Minor Essential Revisions

The sentence “In total there were 342,606 observations” seems incomplete. Observations of what?

The first section of Discussion is a summary, in which the introduction is repeated. It can be suppressed.

Replace the sentence punctuation (comma instead of period) “Based on such calculations. She concluded that…” for “Based on such calculations, she concluded that…”
**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.