Reviewer's report

Title: Linkage of health and aged care service events: comparing linkage and event selection methods

Version: 4 Date: 7 May 2008

Reviewer: Catherine Quantin

Reviewer's report:

This paper was thoroughly revised by the authors and the new version is much clearer. They have answered to all major points.

However, these five following new major points still need to be corrected:

1- page 11, 1st sentences: “showed that all three of the above scenarios had occurred, with no one single cause dominating. Taken together, the above results confirm the use of the N linkage as the reference linkage for this study”. These results should be moved from the methods section to the results section.

2- page 11: the definition of the Event-based (E) linkage is misleading: “limited demographic information in conjunction with event dates and event characteristics”.

In fact, in page 7, the authors precise that “characteristics as corresponds to date of birth, sex and geographic region of usual residence of the person involved”. These characteristics seem to refer more to personal characteristics than to event ones. As a consequence, Event based linkage does not use only event based data but a mixture of event-based and personal based characteristics. The title of this linkage strategy may be confusing.

page 16 and 17: the part “Constrained strategy” is too long, and should be reduced (with the detailed lists moved to an appendix) and clarified. For example, in page 16 line 5, “partitioning the RAC and hospital data sets into subsets minimised collisions”: an explanation is needed for this point and for the two other points of the same sentence (greater differences and link priorities).

3- page 16: You should add one sentence to explain the aims of the Websphere software (for people who are not familiar with this software) and give at least a reference.

4- Table 3: a legend should be added to give a detailed definition of each type of concordance.

5- Tables 6, 7: these two tables are very difficult to understand

Table 6: the table is too long and could be reduced (keeping PPV and sensitivity and deleting NPV, specificity and relative size). The main results of the deleted parts could be summarized in one or two sentences in the main text.
Table 7: this table could be totally deleted. The main results could be given in the text.

MINOR POINTS

a- page 16, line 19: Is “by” missing after “given”?
b- page 21: you should mention table 5 since the beginning of the presentation of the results

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.