Reviewer's report

Title: A Controlled Trial of Automated Classification of Negation from Clinical Notes

Version: 1 Date: 27 October 2004

Reviewer: Dipak Kalra

Reviewer's report:

General

This is a valuable investigation, performed and documented well, and this material IS highly suitable for publication.

-------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

None

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. The paper needs some tidying with respect to headings and the right text margin. I found one case of a double full stop at the end of a sentence near the top of page 2.

2. The first three paragraphs of the introduction is quite detailed. (Maybe this is meant to be the whole introduction, and the next heading is missing.) It includes some internal repetition and also repeats material from elsewhere in the paper. I would suggest a more introductory introduction (i.e. remove some bits of the method detail, and add a few words to explain some of the core concepts of the domain such as "concept-based indexing").

3. Page 4-5 - During the categorisation of encoded concepts into Positive, Uncertain or Negative, to what category is uncertain negation assigned (e.g. "probably not...")? I would recommend that the authors add an explanation on this, probably in the upper half of page 4. A sentence or two would suffice.

-------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

1. Study Design (page 3) describes the classification of the document content into sections, for which a set of headings is given. It is not clear if this sectioning process alters the way in which concept indexing subsequently occurs (i.e. does the section alter the probability of occurrence of different concepts?) or impacts on the way that SAVS investigates for negation. Although the tables of results include the section name, a few sentences could be added to the discussion section to explain how this is relevant to the negation
challenge addressed in this paper.

2. Is there any diagram or other easily understood representation of the authors’ “automated negation assignment grammar” that might be included in the paper. The design of this conceptual approach to negation is clearly fundamental to the success observed.

3. As a general remark, I think this paper would be of interest to a wide range of readers, but presently uses many terms taken from the author’s specialised branch of health informatics. Some of the terms that might not readily be understood by, say, a general health informatics student, could be complemented by simple examples (as the authors have nicely done in paragraph 3 of page 4). My suggestion is therefore to include a few more simple illustrative examples within the paper.

What next?: Accept after minor essential revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

None