Reviewer's report

Title: CDAPubMed: a browser extension to retrieve EHR-based biomedical literature

Version: 2 Date: 19 October 2011

Reviewer: Aurelie Neveol

Reviewer's report:

CDAPubmed: a browser extension to retrieve EHR-based biomedical literature
David Perez-Rey, Ana Jimenez-Castellanos, Miguel Garcia-Remesal, Jose Crespo and Victor Maojo

BMC Medical Informatics and Decision Making Software

This paper presents a software tool that facilitates PubMed searches launched directly from EHRs in HL7-CDA format. Overall, this is a useful tool, made publicly available for free. The compatibility of the tool with Firefox 3.x (and not later versions) is a limitation that is not mentioned in the paper. The evaluation and comparison of the tool to other work/tools is also limited.

1. Does the software address a novel task? Alternatively, if there is already software available that performs this task, does the software outperform it in terms of speed, reliability, efficiency, or breadth of application?

The software addresses a similar task to the InfoButtons developed by Cimino et al. at Columbia University in the past decade. However, to my knowledge this is first tool of the kind to be freely available as a Firefox plug-in

2. Is it easy to use?

The software seems fairly easy to use (except where qualifiers are concerned – see # 6). One concern is the compatibility with Firefox 3.x and not later versions. The current version of Firefox is 3.7 – I had to replace my current version of Firefox with an older one (which may not be maintained by Mozilla in the future) in order to install the software. This might be an inconvenience for users requiring a recent version of Firefox for other needs, since I don’t think it’s possible to have two different versions installed concurrently.

3. Does it satisfactorily address the task or application the authors intend?

The software does address the task as described, i.e. it assists in querying PubMed based on keywords found in EHRs in HL7 format.

4. Is the software freely available for non-commercial use (note that this is a condition of publication)? And is the availability of the software and any restrictions on use clearly stated in the manuscript?
The software freely is available at the URL listed in the paper. The paper does state that the software should be used with HL7-CDA documents. However, the main restriction i.e. compatibility with an ancient version of Firefox is not mentioned in the paper. It is mentioned in the user manual freely available at the download URL.

5. Does the manuscript clearly describe the problem the software is designed to address

Yes, it does.

6. Does the manuscript clearly describe how the software is implemented?

Relevant keywords for biomedical literature search are extracted from EHR free-text by matching “sentences and words” (p. 4) extracted by the OpenNLP package to a list of MeSH Descriptors or Entry Terms. These terms are then used to build queries submitted to PubMed. Very little detail is provided on the keyword extraction process. For example, I did not understand how the authors used MeSH qualifiers. I tested the software on some of the sample EHRs provided by the authors and I was not able to generate a query with qualifiers, even when expressions such as “history of violence” were highlighted (en003.xml). The automatic query generated was (history [mh]) AND (violence [mh]) when violence/history[mh] would clearly have been more appropriate.

7. Does the manuscript clearly describe how the software performs and its advantages / limitations over existing applications?

The evaluation was performed on 17 EHRs. This is a small number of documents, but it could be considered acceptable for a proof-of-concept. However, the issue is with the evaluation itself. For the indexing part of the process, the authors state that “Keywords from each HER of the dataset (...) were compared to a manual identification performed by an expert”. Additional information would have been useful on the guidelines given to the expert for identifying the gold standard keywords. Was the expert shown the algorithm results for validation? What was the expert’s expertise? MeSH indexing? Biomedical terminology? Was the expert an MD? A linguist? The performance shown in table 1 seems very high. It would have been helpful to compare these results with either a baseline (e.g. plain string matching) or a state-of-the-art tool (e.g. MTI by Aronson et al.). For the retrieval part of the process, the evaluation is basically limited to the observation that the queries that can be generated with the tool are more specific than a query using only the keyword related to the “present condition” of the EHR. This is inherent to the process of query construction, so really this “evaluation” is not doing anything. It would have been more useful to measure user satisfaction (usability of the system for HER related information search) or the relevance of the documents retrieved (user-assessed relevance with respect to the EHR context).

8. Does the manuscript state the software’s operating requirements
The description of operating requirements is incomplete (see # 4)

9. Are the discussion and conclusions of the manuscript well balanced and adequately supported by the data?

The discussion of the most relevant aspects to this work is very thin. InfoButtons are not discussed at all (they are only briefly mention in the background section), although there is ground for a rich and interesting comparison. The discussion mentions other MeSH indexing tools such as MTI without any substantial comparison of performance or any other aspect. It is limited to a bare acknowledgment of the existence of such tools that “could be applied in future implementations” (p. 6). The conclusion is that the authors aimed at developing the software, and they did it. That much is true.

10. Do the title and abstract of the manuscript accurately convey what has been found?

Yes.

11. Is the writing acceptable?

Yes.

Major Compulsory Revision:

• Improve the evaluation e.g. by comparing the keyword extraction to a baseline or state-of-the-art method and by assessing user satisfaction and the relevance of retrieved documents.

• Improve the discussion of related work and the comparison to InfoButtons and MeSH indexing tools.

Minor Essential Revisions:

• The issue of the compatibility to Firefox 3.x (and not later versions) should be highlighted, e.g. on page 3 in the section “Architecture and Technologies”

• The authors refer repeatedly to the “MeSH taxonomy”. “MeSH thesaurus” would be more appropriate.

• On page 3, the authors state that there are “172,000 synonyms and 24,000 MeSH terms”. This is a vague and imprecise statement. A better alternative would be: “MeSH 2011 contains 177,000 Entry Terms and 26,142 Main Headings”.

• In general it is useful to specifically indicate which version of MeSH is being used in experiments (MeSH is updated every year). When providing the results of a PubMed query (e.g. page 5) it is also useful to indicate the date that the search was performed on (PubMed is updated daily).

• On page 4, it is unclear what the authors mean by “The [mh] suffix specifies that the selected keyword is the respective MeSH term not the synonym present in the EHR”. First, a search on a MeSH Main Heading (e.g. diabetes type 1 [mh]) and its Entry Term (e.g. diabetes mellitus, insulin-dependent [mh]) yield the exact
same results. Second, the [mh] field indicates that the search will be limited to MEDLINE citations indexed with the relevant MeSH term (as opposed to, e.g. PubMed records containing the term phrase in the title or abstract).

**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