

Summary of Discussion for

Modeling emergency department visit patterns for infectious disease complaints: results and application to disease surveillance

by Judith C. Brillman, Tom Burr, David Forslund, Edward Joyce, Rick Picard, and Edith Umland

BMC Medical Informatics and Decision Making 2005, 5:4

<http://www.biomedcentral.com/1472-6947/5/4>

presented by Christopher Maier for

INLS 279: Bioinformatics Research Review

2006-03-29

This article focuses on the timely issue of rapid outbreak detection in public health surveillance systems. Following the terrorist attacks of 2001, interest in quick notification of health officials regarding suspected bioterrorism strikes has grown dramatically. This field offers many opportunities to information scientists, as public health data is very dense and information-rich. The article investigates the use of hospital emergency department chief complaint (CC) data in syndromic surveillance, as opposed to more traditional surveillance methods using diagnosis or discharge data.

While the article was interesting and yielded good discussion, it unfortunately glossed over many of the aspects of the system that are of interest to information scientists. The group had hoped for more discussion concerning the particular methods used to parse the largely free-text CC data, particularly since this a research area of a department professor (Dr. Stephanie Haas), as well as a recent graduate of the SILS Ph.D program (Dr. Debbie Travers). There are many issues surrounding the parsing of CC data, including synonymy, misspellings, alternate phrasing, and acronym expansion. How this non-trivial parsing task is accomplished can have a profound effect on the functioning of the system.

The point that the text parsing functionality of the system was not discussed is almost moot, however, as there are only a small number of categories a CC can be placed into. As such, there is quite a bit of leeway concerning the parsing. However, the fact that there are only seven rather broad categories could potentially mask useful information. If there were finer categories for classification, the general information the system currently generates could certainly still be recovered (perhaps through the use of a CC ontology), but a much more detailed picture of a potential outbreak scenario could be formed. Such a picture is not possible with the current limited categorization scheme.

Another concern of the group stems from the way CCs are categorized. From the wording of the article, it was unclear as to whether the potentially multiple CCs a patient has would each be individually categorized, or if the whole collection would be assigned to a single category. Clearly, this is an important issue that can impact surveillance profoundly. Related to this point, it seemed that while the system was able to detect outbreaks of syndromes confined to a single CC category, there was no functionality that enabled it to notice patterns of symptoms that crossed category boundaries. This would

be an incredibly helpful tool to have, particularly with bioterrorism surveillance, as these agents may cause diseases with multiple disparate symptoms. The group postulated the idea of creating disease profiles, snapshots of disease progression in terms of CC counts in all categories at various timepoints in the progression of an outbreak. This would likely be very sensitive, but the creation of such profiles would be quite difficult, from an ethical standpoint at the very least!

Another favorite issue of the group (mentioned only in passing in the article) is that of effective user interface design. The electronic patient record system in use at the University Hospital featured a drop-down menu for entering CC information, but most users favored the free-text entry facility for its ease of use. The group wondered about the implementation of the menu. Could multiple CCs be selected, or only a single one? How many separate mouse movements were required to enter all the necessary information? How well did the controlled list of CCs reflect the actual practices of the medical staff? This single aspect of the system could provide some very interesting research opportunities.

The paper mentioned ICD-9 codes for diagnoses, leading the group to wonder about the existence of CC controlled vocabularies or ontologies that could be used in systems such as this. The implementation of a consistent terminology for CCs on a nation-wide basis would be an incredible boon for the syndromic surveillance community, as well as for the public health and epidemiology communities in general. The difficulties concerning such an undertaking would be quite impressive as well.

On the whole, the paper provided much thought-provoking discussion. More attention to topics of interest to information scientists would have been appreciated, naturally, though the group understands that this is not the target audience of the paper. The writing of the paper, as well as the labeling and discussion of several figures left much to be desired, however. Many concepts were unclear and received somewhat cursory attention. The brief discussion and non-existent labeling of Figure 5 in particular was very disappointing.