Natural language processing and inference rules as strategies for updating problem list in an electronic health record

Fernando Plazzotta[a], Carlos Otero[a, b], Daniel Luna[a], Fernan Gonzalez Bernaldo de Quiros[a]

[a] *Health Informatics Department, Hospital Italiano de Buenos Aires, Argentina*
[b] *Department of Medical Informatics & Clinical Epidemiology, Oregon Health & Science University, Portland, OR, USA*

Abstract

Physicians do not always keep the problem list accurate, complete and updated. **Objective:** *To analyze natural language processing (NLP) techniques and inference rules as strategies to maintain completeness and accuracy of the problem list in EHRs.* **Methods:** *Non systematic literature review in PubMed, in the last 10 years. Strategies to maintain the EHRs problem list were analyzed in two ways: inputting and removing problems from the problem list.* **Results:** *NLP and inference rules have acceptable performance for inputting problems into the problem list. No studies using these techniques for removing problems were published* **Conclusion:** *Both tools, NLP and inference rules have had acceptable results as tools for maintain the completeness and accuracy of the problem list.*

Keywords:
Electronic health records; medical records; medical records, problem-oriented; problem list; natural language processing; inference rules.

Introduction

The problem-oriented electronic health record has become one of the most developed clinical documentation systems in medical informatics. While the advantages of a problem list are known and have been published in numerous studies, physicians do not always keep the problem list accurate, complete and updated.

The difficult with problem lists encompasses not only the omission of problems in the list, but also, eliminating problems once the patients no longer suffers from a disease. A study about this topic at Hospital Italiano de Buenos Aires found that only 1% of the problems were eliminated when the condition was no longer present.

The aim of this paper is to analyze natural language processing (NLP) techniques and inference rules as strategies to maintain the completeness and accuracy of the problem list in electronic health records.

Materials and Methods

We conducted a search in PubMed. The terms used for the search were the combination of the term "problem list" and one of the following MeSH terms: natural language processing, electronic health records, medical records, and medical records, problem-oriented. We also use the term "inference rules" in combination with the term problem list. From the articles we obtained, we review their reference lists and related articles to identify other studies. The results were evaluated according two criteria: Strategies for inputting problems into the problem list, and Strategies for removing problems from the problem list. The review examined the publications from 2002 to the present. We not review conference proceedings not indexed in PubMed, or books. Also, this was not a systematic review

Results

The terms "problem list" and "inference rules" do not have a corresponding MeSH term that defines this concept, so the search was complex and tedious.

Strategies for inputting problems into the problem list: There are several reported cases using NLP with an acceptable level of success. The work of Meystre et al showed that NLP strategies achieved a high rate of health problem recognition, especially when the study was performed with a controlled set (a specific specialty such as cardiology, for example, or when the study was conducted in intensive care units). The Meystre team evaluated different NLP techniques, with similar results. Similarly, both Long and Cao et al, achieved good results using NLP in the recognition of problems in free text when detecting diagnoses on discharge summaries. One limitation for updating problem lists using NLP tools is that in most cases, the studies were conducted with a limited set of health problems. Another limitation, is the language in which the tool was developed. Although there are reports of the use of this NLP tool in English and German, these strategies are not validated in other languages, so validity must be test in other settings.
Inference rules are systems that define a set of data that allows identifying special conditions in a health information system (HIS). Wright et al has published their experience using this tool. They work with a set of inference rules for the detection of 17 common problems in ambulatory practice. One limitation with inference rules as a strategy to input a problem into the list is that the rules cover a limited universe of problems. Other limitations are that you need to create a different rule for each condition. However, this strategy has the advantage that it need not be translated and can be used in different languages.

Strategies for removing problems from the problem list: This section discusses strategies dealing with the elimination of problems in the problem list when the patient has no longer this disease or condition. We found no articles discussing this issue.

Conclusion

Natural language processing and inference rules have had acceptable results as tools for incorporating health problems into a problem list, mainly using limited sets of data. Further studies are needed to validate these rules in other areas and to extend the tools to a more comprehensively

Address for correspondence

Dr. F. Plazzotta || fernando.plazzotta@ hospitalitaliano.org.ar