

ABSTRACT

Automatically tracking diabetes using information in physicians' notes

RS Bhatia, S McClinton, and RF Davies

Cardiology Research, University of Ottawa Heart Institute, Ottawa, Ontario, Canada
E-mail: rbhatia@ottawaheart.ca

Objective

This paper describes DMReporter, a medical language processing system that automatically extracts information pertaining to diabetes (demography, numerical measurement values, medication list, and diagnoses) from the free text in physicians' notes and stores it in a structured format in a MYSQL database.

Introduction

Patient consultations recorded as voice dictations are frequently stored electronically as transcriptions in free text format. The information stored in free text is not computer tractable. Advances in artificial intelligence permit the conversion of free text into structured information that allows statistical analysis.

Methods

DMReporter uses machine learning and natural language processing tools to do information extraction. Numerical measurement values extracted include blood pressure,

weight, LDL, HDL, total cholesterol, HbA1C, fasting glucose, glucose (unspecified), and creatinine. These values are present as multiple readings, potential target values, values over a period of time, varied location, values reflecting family history, changes in value versus absolute value, and so on. The methods developed in Bhatia *et al.*¹ extract and disambiguate these values while producing informative label-value pairs. The system extracts 290 medications in six categories using regular expressions and edit distance algorithm. The diagnosis detection uses negation detection in the sentences that mention diabetes using the NegEx algorithm.² The dataset used by the program consists of

Diabetes Report Card University of Ottawa Heart Institute/The Ottawa Hospital			
Patient Info		LAB Data	
Patient ID	00000001	Blood Pressure	May 11, 2007 140/95
Problem	1.04	Cholesterol	May 11, 2007 4.1
HPI	2.042	Fasting Glucose	May 18, 2007 5.4
Gender	Female	HDL	Jan 19, 2008 1.46
Age/Last Visit	46	LDL	Jan 19, 2008 2.2
Measure		Medication	
Measure	Score	Medication	May 11, 2007 11 mg
Alcylure	3	Metoprolol	May 11, 2008 100 mg
Blood Pressure	0	Toradolol	May 11, 2008 40 mg
LDL	0.5	Yasmin	May 11, 2007 70 mg
HbA1C (glucose)	1		
Wt	0		
Total	2.5/5		

Figure 1 Diabetes report card.

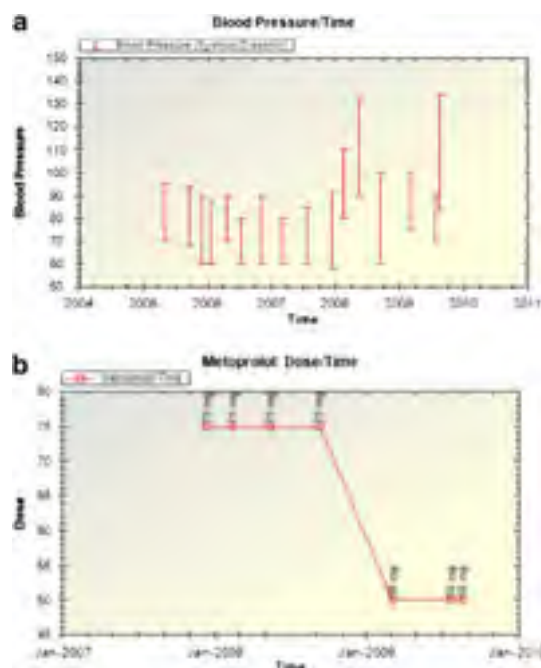


Figure 2 (a) Blood pressure recorded over time for a patient. (b) Metoprolol doses recorded over time for a patient.

81 932 transcribed outpatient notes collected from 30 459 patients over 5 years.

Results

Currently, the system automatically extracts diabetes diagnoses, medications along with dosage and frequency information and nine numerical lab values. It produces a report card for every patient (Figure 1), assigns a score out of five representing treatment and health, and tracks each variable extracted over the entire recorded period of patient history (Figures 2a and b) in addition to population reports.

Conclusion

DMReporter allows monitoring of diabetic patients' laboratory values and medications over the entire documented history of the patient. The solutions developed for this project can be applied to documents from other institutions and to other diseases. It demonstrates the ability to

develop patient-chart abstractors within a specified restricted domain.

Acknowledgements

This paper was presented as an oral presentation at the 2010 International Society for Disease Surveillance Conference, held in Park City, UT, USA on 1–2 December 2010.

References

- 1 Bhatia R, Graystone A, Davies RA, McClinton S, Morin J, Davies RF, *et al.* Extracting information for generating diabetes report card from free text in physicians notes. NAACL HLT 2010 Second Louhi Workshop, 2010, pp 8–14. Available at <http://aclweb.org/anthology-new/W/W10/W10-11.pdf>.
- 2 Chapman W, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. Evaluation of negation phrases in narrative clinical reports. *Proc AMIA Symp* 2001, pp 105–14. Available at <http://www.ncbi.nlm.nih.gov/sites/ppmc/articles/PMC2243578/>.