International Journal of Integrated Care ISSN 1568-4156

Volume 15, 20 October 2015
Publisher: Uopen Journals
URL: http://www.ijic.org
Cite this as: Int J Integr Care 2015; ETC Conf Suppl; URN:NBN:NL:UI:10-1-117169

Conference Abstract

From unstructured EHR text to data-driven clinical decision support

***Kasper Jensen**, PhD, Postdoctoral Fellow, Norwegian Centre for Integrated Care and Telemedicine, Tromsø, Norway*

***Knut Magne Augestad**, MD, PhD, Department of Surgery, Hammerfest Hospital, Hammerfest, Norway*

***Rolv-Ole Lindsetmo**, MD, PhD, Department of Gastrointestinal Surgery, University Hospital of North Norway, Norway*

***Stein Olav Skrøvseth**, PhD, Section Manager, Norwegian Centre for Integrated Care and Telemedicine, Tromsø, Norway*

*Correspondence to: **Kasper Jensen**, E-mail: kasper.jensen@telemed.no*

Abstract

Purpose: The structured content of the electronic health records are commonly used for both data-mining and clinical decision support. With some information stored in a structured manner, there is a vast underexplored potential in the free text that healthcare providers write about their patients. Unfortunately, the information in the unstructured texts of the records are much harder to access, because accessing the information in the unstructured text involves much more preprocessing. In this project we are developing a data-driven clinical decision support system that's builds as a live feed on top of the unstructured text.

Context: We believe that having a tool that can track the health of patients from they enter the system, empowering healthcare providers with data-driven decisions, will improve the outcomes for patients and that such a system holds a great promise to the future of healthcare. We envision such a tool to enable us to track the patient's health; from the very point of entering the hospital system, enabling data-driven decisions for healthcare providers all the way from diagnosis to treatment and follow-up, improving outcomes for patients, shortening overall treatment time, providing safer and more convenient treatment for patients, while lowering the workload on the system and overall costs. By providing optimal clinical-decision-paths for patients we may also learn new correlations between symptoms, diseases, unwanted side effects and detect those early on.

Methods: We use the medical subject headings terms to schematize the unstructured text in the health records. Norwegian compound words were mapped by splitting them into fragments no shorter then 5 characters, matching the longest compounds. We then used a Smith-Waterman algorithm to map text with spelling variations to the medical subject headings. Natural language processing was used to recognize and map terms in the unstructured text of the records to concepts. By using the time information of the records we create temporal events that we transform into patient pathways by systems-wide correlation, learning the correlations of the data.

Results and discussion: The electronic medical records contain a huge amount of patient related data, such as diagnosis, treatment procedures, drugs, symptoms and time. There have been some attempts to analyze disease progression pathways to reveal how disease progress over time such as the IBM PrognoSim [1], or the attempt to extract patient trajectories from the health records in the Danish health register [2]. We acknowledge that such clinical decisions support tools has to be made in a clinical relevant manner that helps the healthcare provider in their daily interactions with the patient. We believe that for such a data-driven clinical decision support tool to be useful it needs to not only show what kind of pathways exists, but it needs to be able to provide an confident estimate of the pathway is a given patient likely going to experience, such that relevant information for the decisions to be made, are made available at the right time and context.

Keywords

data-driven clinical decision support; electronic health records; natural language processing

---

References

1. K. Wongsuphasawat and D. Gotz, “Outflow: Visualizing Patient Flow by Symptoms and Outcome,” *computer.org*, Mar. 2011.
2. A. B. Jensen, P. L. Moseley, T. I. Oprea, S. G. Ellesøe, R. Eriksson, H. Schmock, P. B. Jensen, L. J. Jensen, and S. Brunak, “Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients,” *Nat. Commun.*, vol. 5, p. 4022, 2014.