Visual Inspection of Longitudinal Electronic Medical Records

Josua Krause
New York University
josua.krause@nyu.edu

Narges Razavian
New York University
narges.sharif@gmail.com

Enrico Bertini
New York University
enrico.bertini@nyu.edu

David Sontag
New York University
dsontag@cs.nyu.edu

ABSTRACT

Electronic medical records (EMRs) and administrative data contain a large amount of distinct events, like diagnoses, laboratory tests, etc., making it difficult to "tell" the story of a patient. We propose *patient-viz* to address this issue by using a visual representation of the data. This allows us to provide the large number of distinct event types and additional information like costs and hospital stays in a manageable form. Using both an anonymized public and an unaltered private dataset we explore the usefulness of our tool.

Keywords

Longitudinal event display, zoomable user interface, electronic medical records.

1. INTRODUCTION

Longitudinal studies and insurance claims data generate a large amount of electronic medical records (EMRs) and administrative data. This data contains a large number of distinct events throughout the observed time window of the life of patients. Understanding, interpreting, and finding relations in those records is a challenging task that is hard to achieve using a tabular or similar representation. Therefore, visualization is needed to e.g. better understand predictive models built on top of the data, impact of comorbidities, progression of chronic diseases, or contributors of health care costs.

Our proposed tool *patient-viz* has a visually rich design aimed to make the huge amount of administrative data manageable for data scientists and medical doctors. The goal of the tool is to provide a quick overview of one patient which can be further explored to inspect detailed information. The input can be any temporal event data with a large number of differently typed events. We test our tool with online accessible semi-synthetic data provided by CMS [1] and privately collected data from a major US insurance company.

VAHC '15 October 25th, 2015, Chicago, IL, USA

DOI: 10.1145/1235

2. VISUAL DESIGN

Our design is similar to *LifeLines* [4] with additional information shown in overlays and using a smaller vertical footprint which allows for a larger number of distinct types to be shown at the same time. Another notable similar work is the tool *VISITORS* [3] which requires even more vertical real estate per type.

The user interface is split into four components: the information panel, the timeline view, the type view, and the selection view. The information panel shows general information about the currently viewed patient and offers controls over the other views. The timeline view shows all events that happened in the observed time of the patient. Every event, represented as colored rectangle, can be a made diagnosis (green), a performed procedure (orange), a laboratory test result (blue), a prescribed medication (purple), a physician (pink), or a hospital (brown). Same types have the same vertical position. The vertical order signals when a particular type first happened in the timeline, earlier being closer to the bottom of the view, while the horizontal position of an event encodes the day the event occurred. Most event types obey a hierarchy, e.g. the CCS (Clinical Classifications Software [2]) hierarchy for diagnoses and procedures or a simple granularity hierarchy for physicians from identifying the actual person to its profession. Those hierarchies can be used to group together types on different levels of granularity to simplify the view or reclaim vertical space. Hospital stays of a patient are indicated by blue vertical spans behind the events during the time. The histogram below the timeline shows the costs of the claims of a day on a logarithmic scale.

Event labels are either shown in a lens around the mouse or selected by an interestingness measure. For this the timeline is split into time spans defined by steep ascends caused by many new event types. Within each time span the most expensive events are labeled.

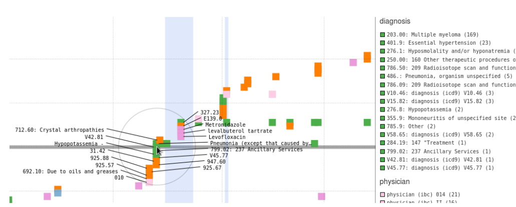

Figure 1: Using the lens to label events close to the cursor. The events shown on the right are selected.

Figure 2: An overview of the tool. The center shows events happening in the time-line with each row containing only events of the same type. At the bottom the cost of one day is shown. The top allows to change axes, the type aggregation level according to the CCS hiearchy [2], connect subsequent events, as well as change selection modes (vertical vs. horizontal), and shows general information about the patient. The lists left and right of the main view show all types and the current selection respectively.

The list on the left hand side shows all events of a given type. The background of each item has vertical lines indicating the occurrences during the observed time window. Users can browse the type hierarchy in the list, expand and collapse individual levels which gets reflected on the timeline, and select types to find them in the main view.

Events in the main view are selected while moving the mouse. Clicking freezes the current selection. The mode of selection can either be by day or by type. The selected events are shown in a list representation on the right hand side.

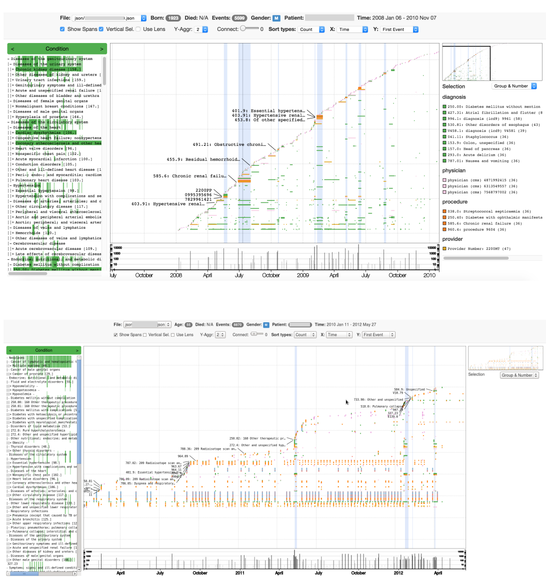

Figure 3: Comparison of the public data provided by CMS [1] (top) and private data (bottom). Notice how steep ascends are mitigated due to the anonymization of the public data.

3. DISCUSSION & FUTURE WORK

An interesting insight in our data can be seen in Figure 3. Steep ascends in the time-line caused by many new event types are an indicator for an incident happening in the live of a patient. The CMS provided data merges three patients for anonymization purposes which mitigates the impact of incidents letting steep ascends mostly disappear.

Trying to fit as many event types as possible into the view limits the amount of detailed information that can be provided by our tool. For example, currently we indicate actual lab-test values by shade instead of using a more sophisticated approach. One solution would be to expand some types of interest to provide more detailed information if available.

As further future work we plan to work more closely with physicians and investigate how we can improve our tool to integrate more with their tasks. Our tool is also provided as open source project on Github: https://github.com/nyuvis/patient-viz.

4. REFERENCES

[1] CMS medicare claims synthetic public use files. http://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/SynPUFs/, 2008–2010. [Online; accessed 16-July-2015].

[2] HCUP. Clinical Classifications Software (CCS) for ICD-9-CM. https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp, 2012. [Online; accessed 16-July-2015].

[3] D. Klimov, Y. Shahar, and M. Taieb-Maimon. Intelligent visualization and exploration of time-oriented data of multiple patients. *Artificial Intelligence in Medicine*, 49(1):11 – 31, 2010.

[4] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. Lifelines: Visualizing personal histories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '96, pages 221–227, New York, NY, USA, 1996. ACM.