Visual Exploration of Temporal Data in Electronic Medical Records

Josua Krause, Narges Razavian, PhD, Enrico Bertini, PhD, David Sontag, PhD
New York University, New York, NY

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.

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. There-fore, 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.

Contact Information

Josua Krause (josua.krause@nyu.edu)
Narges Razavian (razavian@cs.nyu.edu)
Enrico Bertini (enrico.bertini@nyu.edu)
David Sontag (dsontag@cs.nyu.edu)
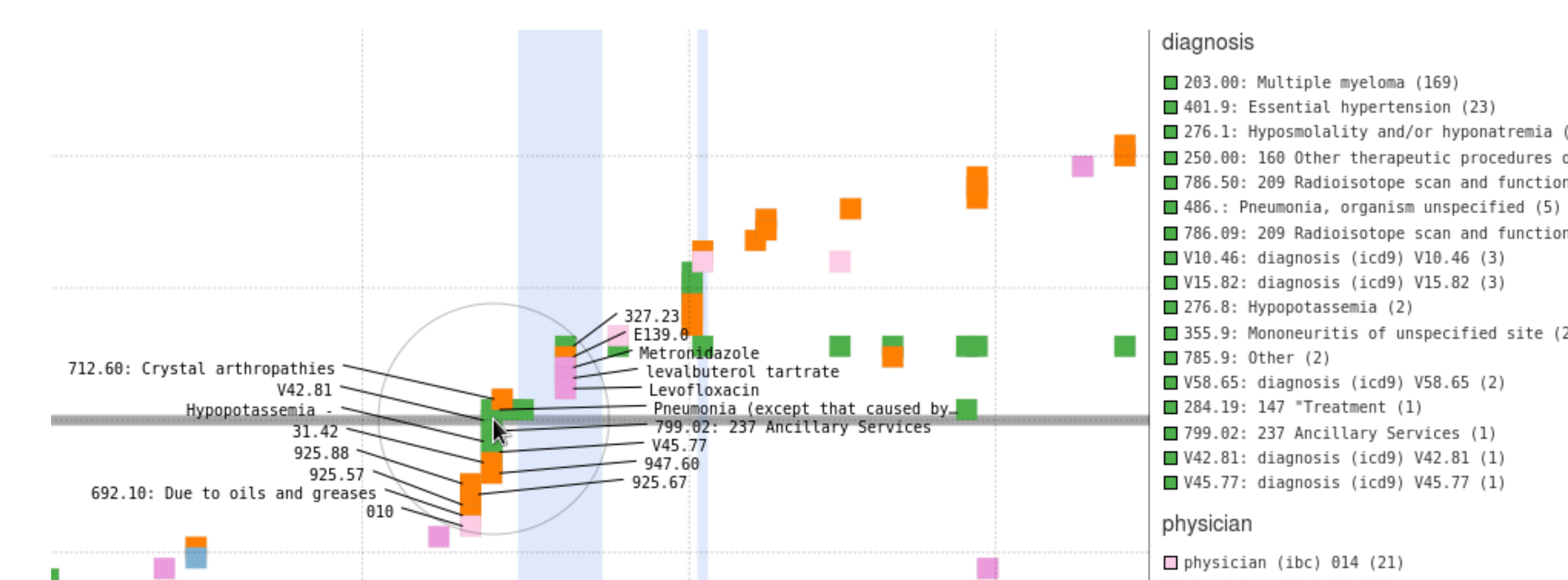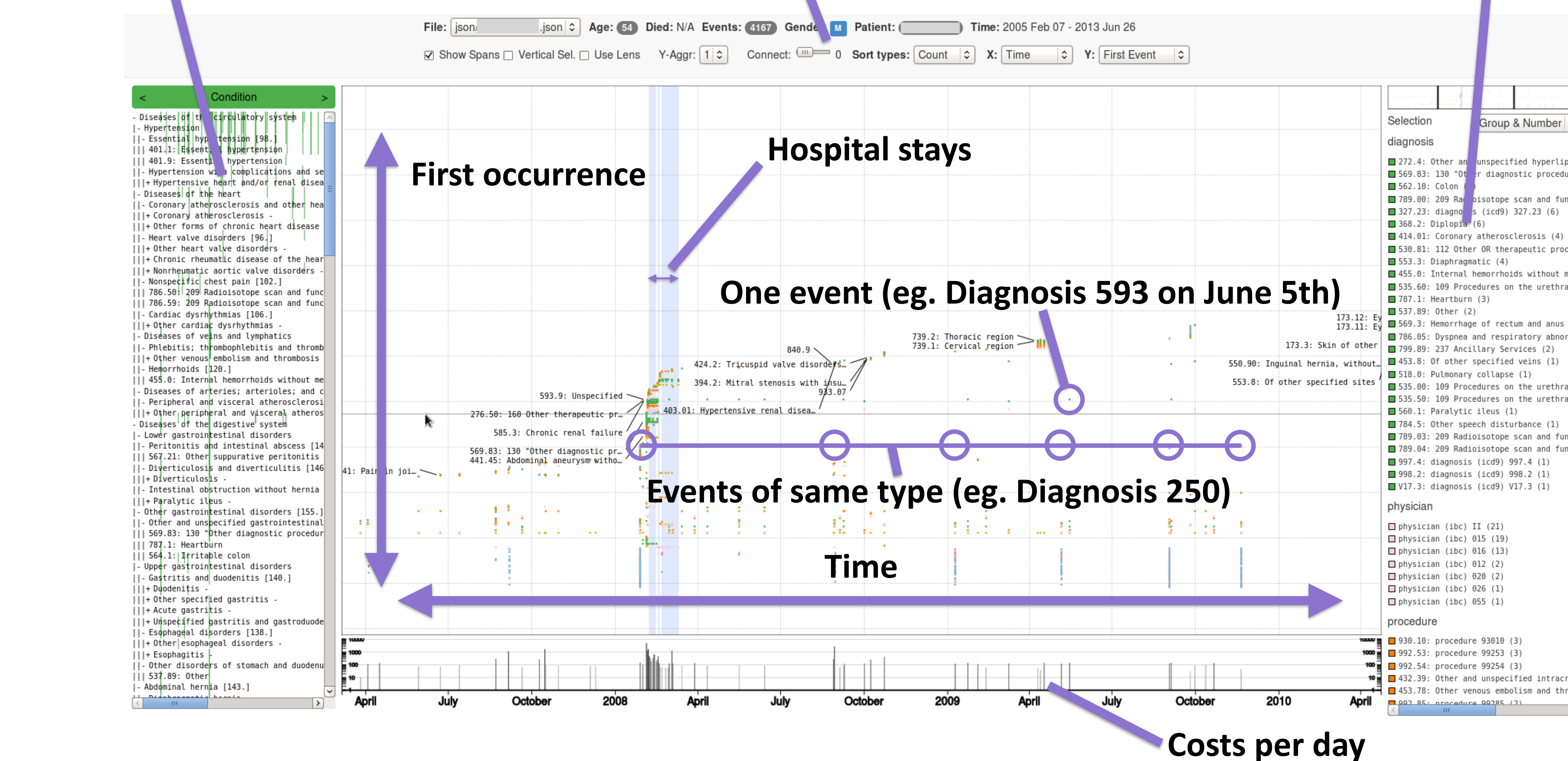Website: https://github.com/nyuvis/patient-viz
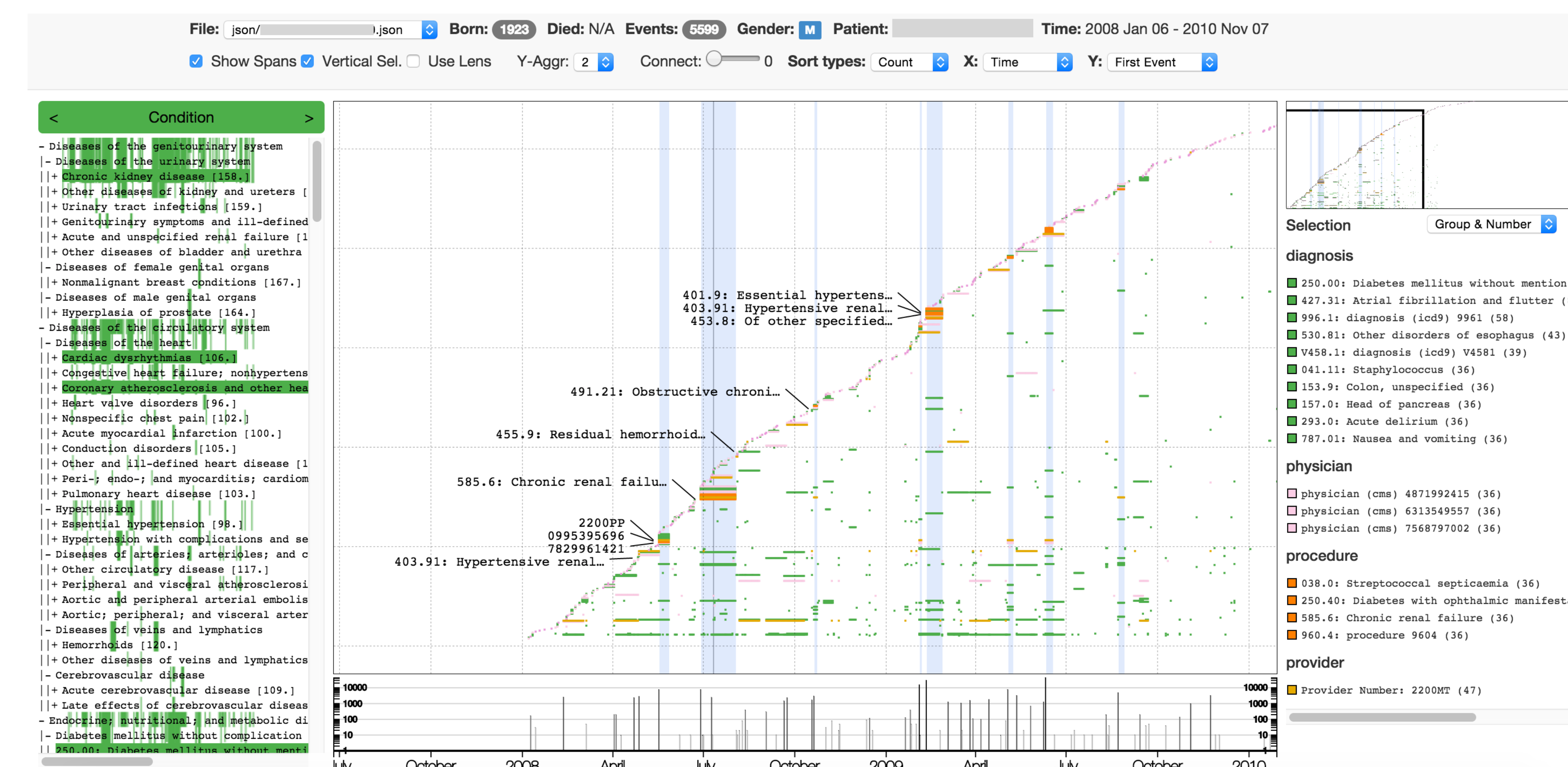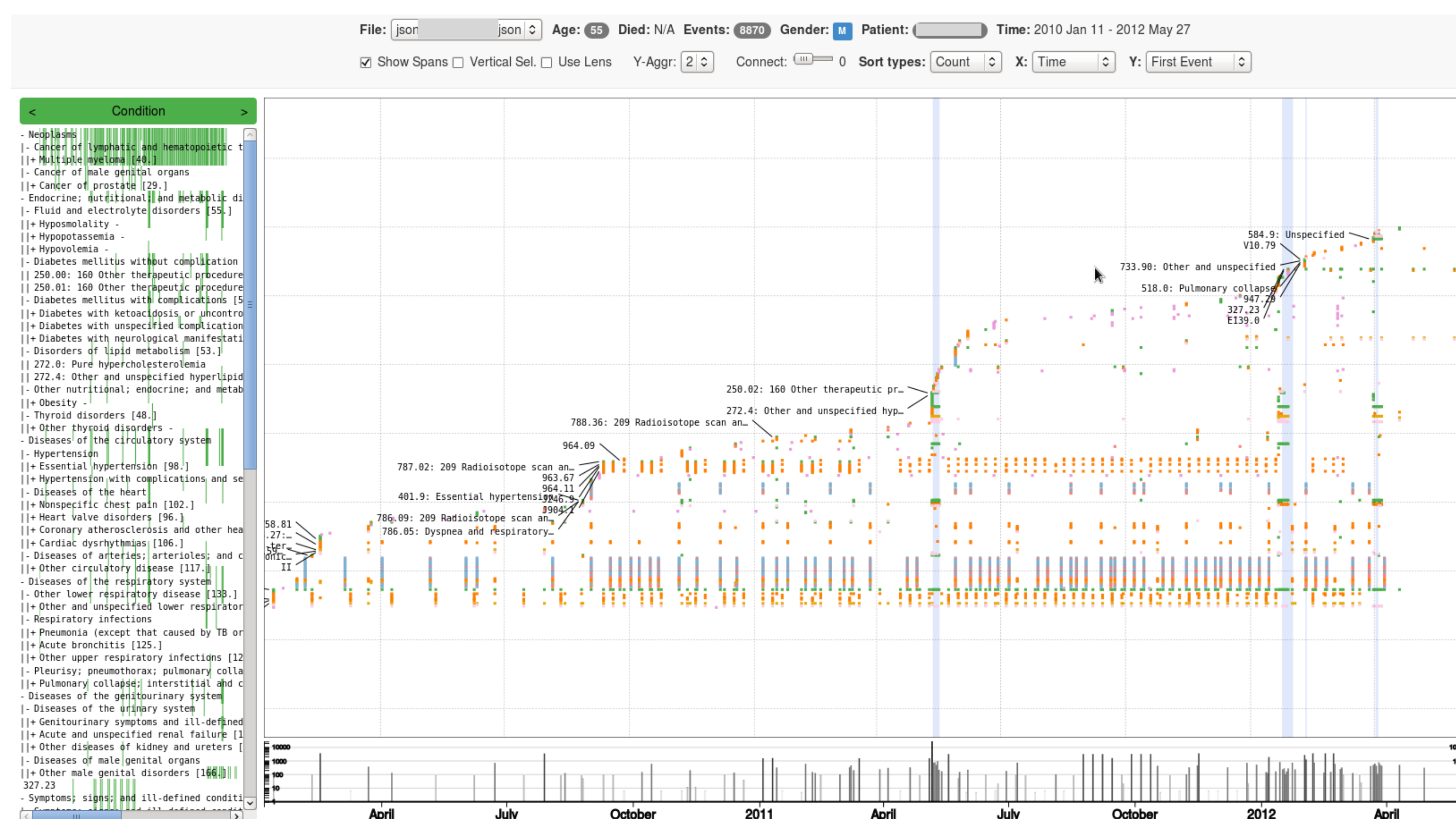
Steep ascends in the timeline caused by many new event types are an indicator for an incident happening in the live of a patient (left). The CMS provided data (right) merges three patients for anonymization purposes which mitigates the impact of incidents letting steep ascends mostly disappear.

A user can use a lens to get information about interesting events. The horizontal granularity of single events is by day and the vertical position is determined by the first occurrence of the particular event in the history of the patient.

Vertical selection can be used to see all events of a particular day. Here a set of lab-tests was performed. Out-of-range values are depicted in red.

Visual Design

Our design is similar to LifeLines [3] 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 [4] which requires even more vertical real estate per type.

Every event is represented as rectangle whose color indicates its type, *ie.* a diagnosis (green), a performed procedure (orange), a laboratory test result (blue), a prescribed medication (purple), a physician (pink), or a hospital (brown).

Conclusions

The tool proved to be helpful for analyzing the output of machine learning algorithms in predictive healthcare analysis. It provides a quick way to identify the problems of an automatically generated model by looking at patients that were classified wrong or turn out to be outliers in the cohort. Additionally, the tool enables clinicians to quickly "tell the story" of the patient using just the administrative data often found in insurance claims. As future work we plan to work more closely with physicians and investigate how we can improve our tool to integrate more with their tasks.

References

1. CMS medicare claims synthetic public use files. 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. www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp, 2012. [Online; accessed 16-July-2015].
3. 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.
4. 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.