doi: 10.1111/fcp.12322

PROCEEDING

Visualization of temporal patterns in patient record data

Catherine Plaisant*
Human-Computer Interaction Lab, University of Maryland, College Park, MD, USA

Keywords
data visualization,
health patient record,
temporal patterns

Received 8 September 2017;
accepted 13 September 2017

*Correspondence and reprints:
plaisant@cs.umd.edu

ABSTRACT

Visualization contributes to a variety of tasks, from reviewing individual patient records to helping researchers assess data quality, find patients of interest, review temporal patterns and anomalies, or understand differences between cohorts. We review some of visualization techniques developed at the University of Maryland.

INTRODUCTION

Event analytics tools can produce meaningful displays of common patterns, interesting rare events, and troubling anomalies. The University of Maryland ongoing work with temporal event records has produced powerful tools for analyzing and exploring patterns of event data. Lifelines were a pioneer in presenting a visual representation of a single record. We later introduced simple operators (align, rank, filter, and group by) to manipulate and visualize collections of records. Our research showed that such temporal queries require users to refine their queries iteratively after seeing the results. Next, we introduced a method for summarizing all sequences found in the collection of records. All are combined in EventFlow [1].

EVENTFLOW FOR VISUALIZATION OF TEMPORAL PATTERNS (http://hcil.umd.edu/EVENTFLOW/)

EventFlow provides a way to (i) visualize and review the data from individual records and their event sequence; (ii) search for temporal patterns of interest, using a powerful graphical interface; (iii) summarize all the event sequences, their timing and prevalence, and find anomalies; (iv) perform data transformations to reveal useful patterns that answer questions you have; (v) select cohorts of interest for further studies.

The interface of EventFlow consists of three main components: interactive controls and legend, overview, and timeline with search control (*Figure 1*).

Figure 1 illustrates data representing men diagnosed with cancer (we use a small sample to explain the design). On the right, the timeline shows details of individual records. Each patient is shown on a separate timeline. Triangles represent events. The legend shows all event categories, enabling users to change the color and order of the categories. The records have been aligned by the cancer diagnosis date (green). Users would need to scroll to see all the records. In the center, the overview aggregates groups of records with the same sequence of events into horizontal (gray) stripes that include colored vertical bars representing each event. Within each horizontal block stripe, the height of the vertical bar is determined by the number of patients in the group and the horizontal gap between events is proportional to the average time between events. Reading from the left, we can see that all records start with a cancer diagnosis. We can then see the different sequences of treatment with luteinizing hormone-releasing hormone (LHRH) (purple) and radiation therapy (brown). The most common first treatment in this group is the LHRH. The second most common is radiation therapy, and we can see that it

Fundamental & Clinical Pharmacology

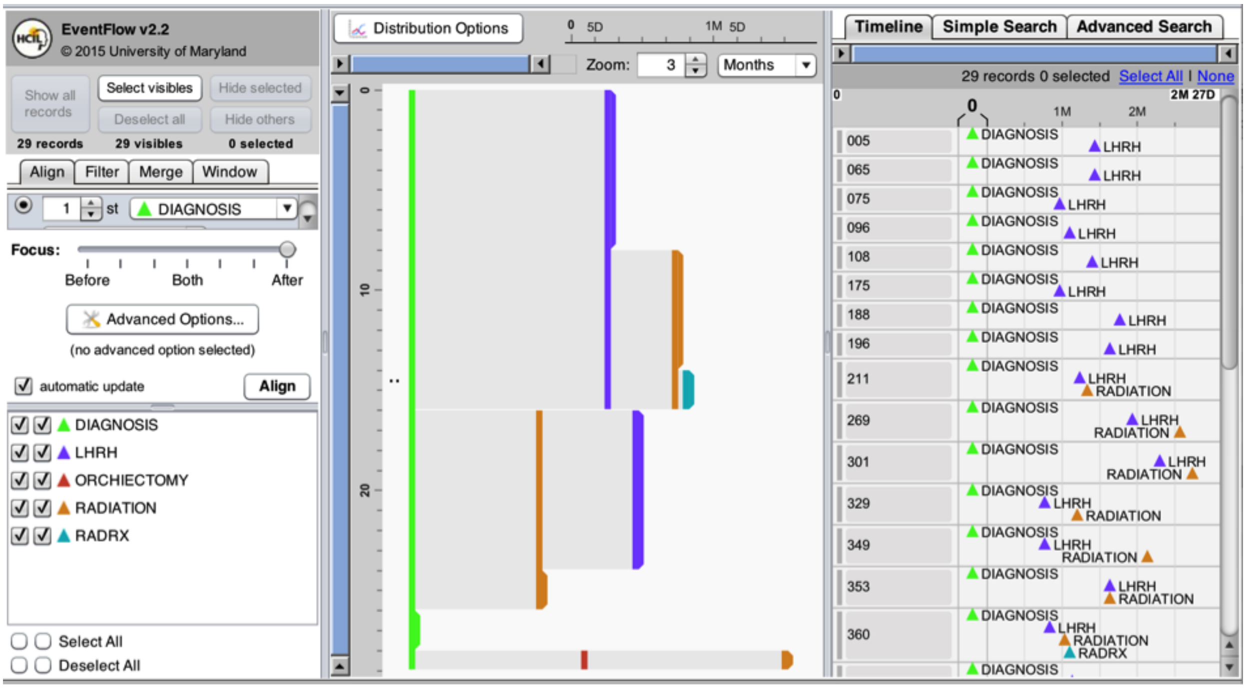

Figure 1 Sample data in EventFlow with diagnosis and treatments.

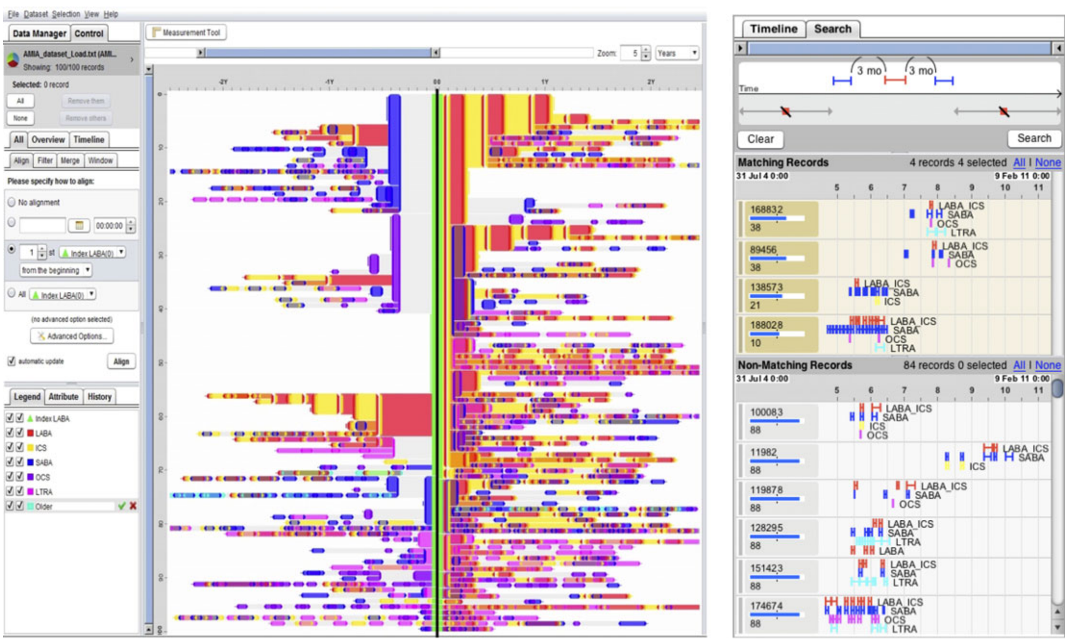

Figure 2 Analysis of prescription patterns of asthma medication (a case study in collaboration with the Army PharmacoVigilance Center – see [2] for more details).

occurs earlier on average than LHRH as the distance from green to brown is shorter than the distance from green to purple.

Users can select other metrics such as the median time between events, and the distribution of values can be seen when the cursor hovers over the space between events.

Drug prescriptions can also be represented in EventFlow (*Figure 2*). Each prescription can be represented as an interval in the timeline. Simple interactive controls allow users to remove small gaps or merge intervals to create longer episodes. Because multiple interval events can occur concurrently, EventFlow renders overlapping intervals in the overview using the combined color of the two overlapping categories. The technique works best when limited to a small number of event categories (i.e., colors), but our experience suggests that being able to see overlaps of just two or three event categories is for many users a significant improvement over existing techniques.

The two views (overview and timeline) are coupled so that when users select an event sequence, all records with that sequence are selected on the timeline view and moved to the top of the timeline for review. The legend allows users to select or deselect which event categories they want to display on the overview or the timeline. After selecting records, users can remove selected records from the display – or save them. Reducing the number of event categories also dramatically simplifies the display to answer questions related to those events. Combining such simple interactive techniques allow users to rapidly narrow the focus of an analysis on records of interest.

Search and data simplifications

EventFlow includes two separate search interfaces. The basic menu-based search interface gives users easy access to either before and after relationships (Subsequence module) or during relationships (Overlap module). The Advanced Search allows users to specify more complex temporal features such as absolute time constraints and absence of events scenarios (e.g., men who did not receive LHRH within 6 months of diagnosis) [1]. Matching records are selected in the timeline display and moved to the top, while those that do not match remain at the bottom. This allows users to quickly see not only the records that match their query, but also records that did not match so they can check that the search behaved as expected.

A set of simplification operations allows users to focus on patterns of interest [1]. Those include graphical search and replace or category aggregation (possibly using hierarchical structures). Based on 20 case studies, we cataloged 15 strategies for sharpening analytic focus that analysts can use to reduce the data volume and pattern variety when dealing with larger number of records [3].

Medical researchers have applied EventFlow to analyze treatment patterns and outcomes in electronic health records, claims data [4], or even codes extracted from video observations of medical procedures.

HIGH-VOLUME HYPOTHESIS TESTING FOR COHORT COMPARISON (http://hcil.umd.edu/EVENTFLOW)

Cohort comparison studies have been traditionally hypothesis-driven and conducted with carefully controlled environments. However, researchers are now moving toward more exploratory methods and retrospective analysis of existing data. CoCo is a visual analytics tool that enables users to compare two sets of temporal sequence data, and aid researchers in understanding high-volume hypothesis testing results [5]. Using interviews and case studies with domain experts, we iteratively designed and implemented techniques dealing with differences in term of event prevalence, time, and frequency between two cohorts.

ACKNOWLEDGEMENTS

This short paper summarizes the work of many colleagues at the University of Maryland and numerous case study partners. The work was supported by partial funding from NIH, Oracle, Adobe, Leidos, and the State of Maryland.

REFERENCES

1 Monroe M., Lan R., Plaisant C., Shneiderman B. Temporal Event Sequence Simplification. IEEE Trans. Vis. Comput. Graph. (2013) **19** 2227–2236.
2 Plaisant C., Monroe M., Meyer T., Shneiderman B. Interactive visualization, in: Marconi K., Lehman H. (Eds), Big data and health analytics, CRC Press Taylor and Francis, Boca Raton, FL, 2014, pp. 243–262.
3 Du F., Shneiderman B., Plaisant C., Malik S., Perer A. Coping with volume and variety in temporal event sequences: strategies for sharpening analytic focus. IEEE Trans. Vis. Comput. Graph. (2017) **23** 1636–1649.
4 Bjarnadottir M., Malik S., Onukwugha E., Gooden T., Plaisant C. Understanding Adherence and prescription patterns using large scale claims data. Pharmacoeconomics (2016) **34** 169–179.
5 Malik S., Shneiderman B., Du F., Plaisant C., Bjarnadottir M. High-volume hypothesis testing: systematic exploration of event sequence comparisons. *ACM Trans. Interact. Intell. Syst.* **6**, 1, (March 2016), 23 pages.