AWARD NUMBER: W81XWH-15-2-0016

TITLE: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

PRINCIPAL INVESTIGATOR: Dr. Jesus Caban

CONTRACTING ORGANIZATION: The Geneva Foundation
Tacoma, WA 98402

REPORT DATE: May 2018

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE ***(DD-MM-YYYY)*** May 2018	**2. REPORT TYPE** ANNUAL	**3. DATES COVERED** ***(From - To)*** 15 Apr 2017 – 14 Apr 2018
4. TITLE AND SUBTITLE An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER W81XWH-15-2-0016
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Dr. Jesus Caban Email: mvanwandelen@genevausa.org		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Geneva Foundation 917 Pacific Ave, Suite 600 Tacoma, WA 98402		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012		**10. SPONSOR/MONITOR'S ACRONYM(S)** USAMRAA
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We propose to design, develop, and validate an interactive visualization framework that physicians assessing TBI/PTSD patients with comorbid symptoms can use to explore and analyze clinical data and that researchers can use to hypothesize new research questions. The primary aims of this project are to (1) extend our interactive visual analytic framework which combines multiple clinical measurements to allow it to be used to explore large collections of clinical data and (2) validate the effectiveness of such visualization systems among clinicians that treat service members.

15. SUBJECT TERMS
Data Visualization, Health Information Technologies, TBI/PTSD, Open Source Tools

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	62	**19b. TELEPHONE NUMBER** ***(include area code)***

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS:

Example Table of Contents:

INTRODUCTION: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

1. **KEYWORDS:** Data Visualization, Health Information Technologies, TBI/PTSD, Open Source Tools

2. **ACCOMPLISHMENTS:** The overarching goals of this project are to (1) address the gap between the acquisition of clinical measurements and the diagnosis step by providing an institutive, flexible, and customizable interactive data visualization framework and (2) validate the system among clinicians treating service members diagnosed with TBI / PTSD.

During the first year of the award, a significant amount of work was accomplished. The work accomplished during the first two years was extended to continue to reach each of the tasks and deliverables of the award. The work accomplished during the last year can be summarized as:

1. [Aims 1.2, 1.3, 1.7] Performed an in-depth literature review of existing work related to temporal trajectory analysis that can be used to illustrate clinical data for TBI/PTSD patients. The visualization and analysis of temporal event data is a widely studied topic. In our literature review, we noticed two series of general studies, one that focused on lifelines and another one that was more focused on flow visualization. Table #1 summarizes our findings:

TABLE 1: Who, What, When, Why, and How of Trajectory Analysis

	who (tool name)	when (condition)	why (goal)	what (functionality)	how (method)
1	Care Pathway Explorer [4]	temporal event sequence data	Utilize historical EMR data to extract common sequences of medical events such as diagnoses and treatments, and investigate how these sequences correlate with patient outcome	1. Show an overview of the frequent patterns 2. Examine the frequent patterns and select specific patterns of interest 3. Compute the patient subsets that match the physicians specified subtraces. 4. The Frequent Pattern Analytics mines frequent patterns and displays them in the visualization.	1. Frequent sequence mining algorithm 2. Bubble chart for overview visualization and Sankey diagram for flow visualization
2	Careflow [7]	temporal event sequence data	Help doctors devise a care plan for their patient.	1. Mining care plans from data 2. Visualizing care plans	Sankey diagram for flow visualization
3	Outflow [6], [10]	temporal event sequence data	Provide important insights into how diseases evolve over time and help clinicians understand how certain progression paths may lead to better or worse outcomes.	1. Aggregate multiple event sequences 2. Display the aggregate pathways 3. Summarize the pathways corresponding outcomes 4. Allow users to explore external factors	Sankey diagram for flow visualization
4	Decisionflow [5]	high-dimensional temporal event sequence data (e.g., thousands of event types)	Help analysts and epidemiologists to study data from groups of patients to understand what factors may influence a particular outcome	1. Issue a query to retrieve subsequences of interest. 2. Aggregated to the matching data construct a DecisionFlow Graph, G. 3. G is then analyzed to extract statistics and visualized. 4. Interaction allows exploratory analysis.	1. Milestone demotion algorithm 2. Horizontal layout algorithm of milestone nodes
5	visual analytics technique [11]	temporal event sequence data	Help to understand the patterns of events observed within a population that most correlate with differences in outcome	1. A visual query module to interactively specify episode definitions 2. A pattern-mining module to help discover important intermediate events within an episode 3. An interactive visualization module that help uncover event patterns that most impact outcome and how those associations change over time.	1. Visual query capabilities 2. Pattern mining techniques 3. Interactive visualization techniques
6	LifeLines [12]	personal medical history records	Design appropriate visualization and navigation techniques for presenting and exploring personal medical history records.	1. Present a personal history overview on a single screen 2. Provide direct access to all detailed information from the overview with one or two clicks of the mouse 3. Make critical information or alerts visible at the overview level.	The medical record is summarized as a set of lines and events on a zoom-able time-line
7	LifeLines2 [8]	multiple records of categorical temporal data	Find hidden patterns contained in EHRs (Electronic Health Records) and other temporal dataset		timelines with the same absolute time scale
8	PatternFinder [13]	multivariate and categorical data (over 26,000 medical events for 950 patients)	Search and discovery of temporal patterns within multivariate and categorical datasets	1. Visual temporal query languages 2. Query result visualization	1. Define a temporal pattern as a sequence of events separated by time spans so that it can be queried by events and time spans components. 2. Multiple timelines for query result.
9	ActiviTree [14]	large and complex event-based temporal data	Systematic identification of sequences in social science activity diary data.	1. Enable the users to create a sequence of activities 2. Show a sequence in the context of the individuals daily life	1. The sequence exploration algorithm is steered using an interactive visual interface 2. The currently explored query sequence is shown with linked view

Table 1: Summary of papers reviewed related to temporal trajectory analysis.

All scientific papers reviewed were put into specific categories. Papers #1 to #5 were designed for visualization of aggregated data. An important step in these studies was the mining or aggregating the temporal data to find patterns and then employing visualization techniques such as the Sankey diagram to illustrate the longitudinal aspects of the data. Papers #6-#8 were designed for visualization of individual data points. Lifelines was designed for personal patients clinic history visualization, although the other two methods Lifelines2 and PatternFinder are designed for multiple records, they visualize the multiple records separately. Finally, paper #9 (ActiviTree) was a method developed to allow users to explore the event sequences rather than aggregate patterns form the multiple temporal sequences.

2. [Aims 1.3, 1.7, 1.8] Designed the *Necklace* interface. In order to help users explore the trajectory of medical claim data and compare trajectories and costs, we designed *Necklace*. The Necklace interface consists of three main design ideas: (1) *guided interaction*; (2) *position based global pattern finding*; (3) *multivariate comparison*. See Figures #1 and #2 for more information.

   There are two kinds of rings that are part of the Necklace interface: root ring and branch ring. The initial display only shows a root ring, which contains the information of all the transformation and relationships in the dataset. As users interact with the interface, Necklace will show more and more branch rings. Each branch ring has a parent ring and a parent trajectory, and contains a parent event E and the either the incoming events of E or outgoing events of E. Thus a trajectory is composed of the parent events of a series linked rings.

   Each ring contains four components: nodes, groups, links and chords. Each node on the ring represents a diagnosis event, and those events on branch rings are not individual events, they present events in a trajectory. For instance, assuming there is a branch ring R which shows its parent event E_3 , and those outgoing events of E_3 , and the parent trajectory of R is $E_1 \rightarrow E_2 \rightarrow E_3$, so an event E_4 on R represents the event E_4 in the trajectory $E_1 \rightarrow E_2 \rightarrow E_3 \rightarrow E_4$. And using the ICD 9 Code, we clustered those nodes on a ring into groups. And each link inside the ring represent a transformation relationship between two events, and links from one group to another group are gathered into a chord.

   The initial heights of nodes are similar and link nodes encode the incident rate of the corresponding event, which means that the more links, the higher the node will be. In addition, in each node there is a histogram for users to show more information. A ring shows abstract information when it occupies a small area, the detailed nodes information and links information will be hidden if the area is not big enough. In order to obtain detailed nodes information, users can click on outer contour of a group to expand this group to a higher level for more space, and they can also click on the inner contour of a group to shrink to a lower level. One of the advantage of ring is that it has a powerful ability to show large scaled event since a ring can show thousands of event at a time.

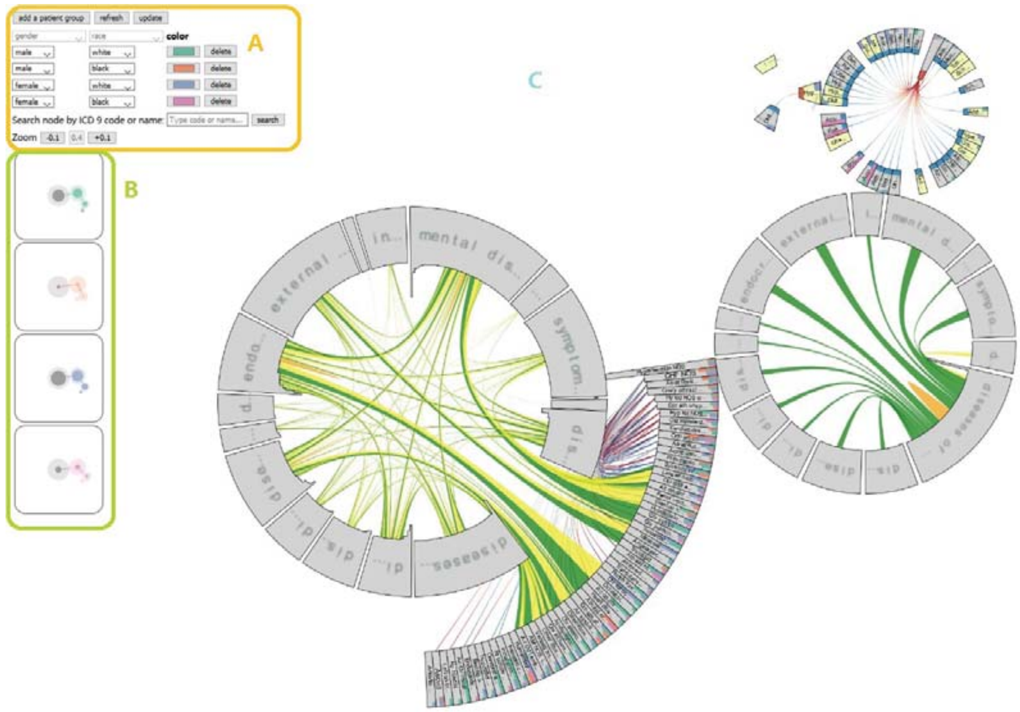


Figure 1: Screenshot of the Necklace system. Necklace visualizes temporal events trajectories, this screen shot shows CMS data, which contains 222 patients and 668 claim records. Users can explore the diagnoses trajectory by directly interacting with the visual objects shown in C, and compare trajectories of different patients cohorts with the overview graphs shown in B. And Necklace also offers a user panel, which is shown in A, for users to edit patient cohorts, search for diagnosis node and zoom the display.

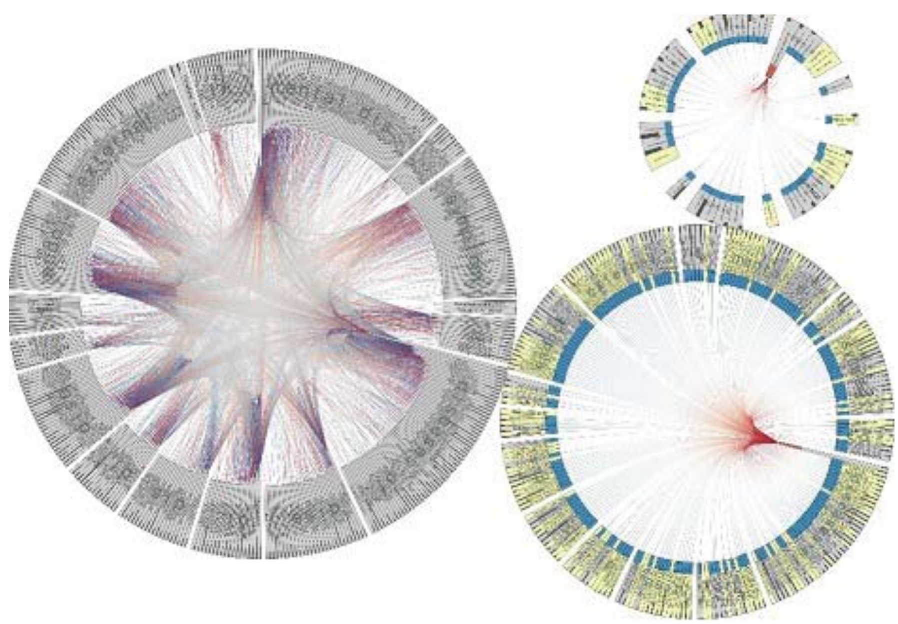


Figure 2: Screenshot of the Necklace system. Necklace visualizes temporal events trajectories, this screen shot shows CMS data, which contains 222 patients and 668 claim records.

3. [Aims 1.3, 1.7, 1.8] Developed, extended, and continued to enhance the *Patient Timeline* tool based on comments and feedback received from different clinical staff members. The *Patient Timeline tool* was developed to assist physicians explore the longitudinal medical data of a patient. In a visually appealing way, the Patient Timeline starts by displaying a tree that provides a quick glimpse of the data available for a patient. Icons and their size show if the patient has a certain type of data in their history and how much or how important that data is. The year nodes can be expanded to show another level of the tree, showing the patient's monthly data for that year. To the right of the patient's tree is a summary panel. This summary panel shows a brief text summary of the patient, their lab history, their diagnosis history and their medication history. The Labs, Diagnoses, and Medications tabs display their information in analytical models that make the data easier to digest. The Labs, Diagnoses, and Medications tabs update when a new tree node is clicked. Furthermore, three patients' tree can be seen by selecting the Three Trees display from the dropdown button at the top of the page.

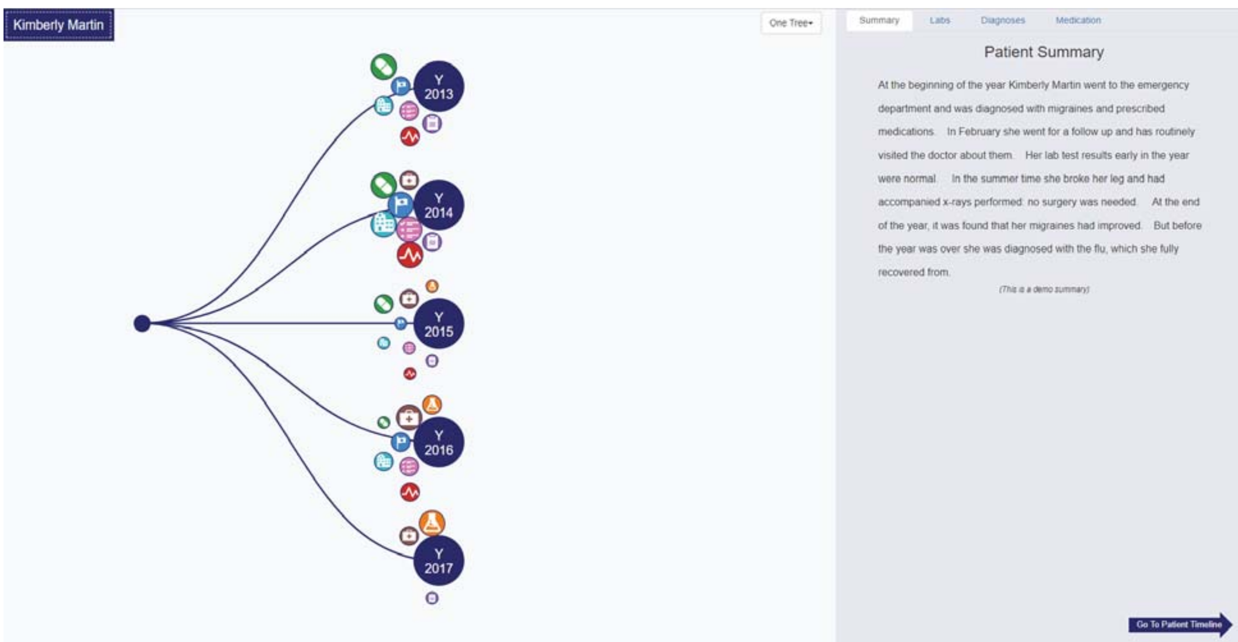


Figure 3: The Patient Timeline starts by displaying a tree that provides a quick glimpse of the data available for a patient.

In the more robust portion of the Patient Timeline, the nodes are presented to us on a timeline as shown in Figure #3. Here, individual days of the patient's data can be explored in panels. Within the panels there are tabs for Labs, Medications, Vitals, Notes, Diagnoses, Procedures, Radiology Note, and Chief Complaint for a single day, assuming the patient has that type of data available for that day. To the left of the timeline, we have a filter that allows the user to filter nodes based on Provider Type, Provider Name, DMIS/MTF, MEPER4, Encounter Type, Data Type, and/or Date. To the right we have a smaller version of the, previously presented, summary panel.

Figure 4: The nodes are presented on a timeline and individual days of the patient's data can be explored in panels.

4. [Aims 1.3, 1.7, 1.8, 1.9] In order to also satisfy the requirements from research clinicians interested in exploring clinical data but also understanding what data is available, we designed the *Database Search* tool. The *Database Search* tool was designed to facilitate the searching and exploration of databases. When the user clicks the Run Query button, the Results panel fills with statistics of their selected tables and variables. In the Results panel, they can see the Count, Missing, Mean, Standard Deviation, Zeroes, Minimum, Maximum, and Histogram for their chosen variables. Once the user has selected all the data they would like to put a request in for, they can click the Export button, followed by the 'Copy to clipboard' button to copy their selections to their computer's clipboard. This saves the user from having to individually copy and paste, or type up, the tables and variables they would like to put a request in for.

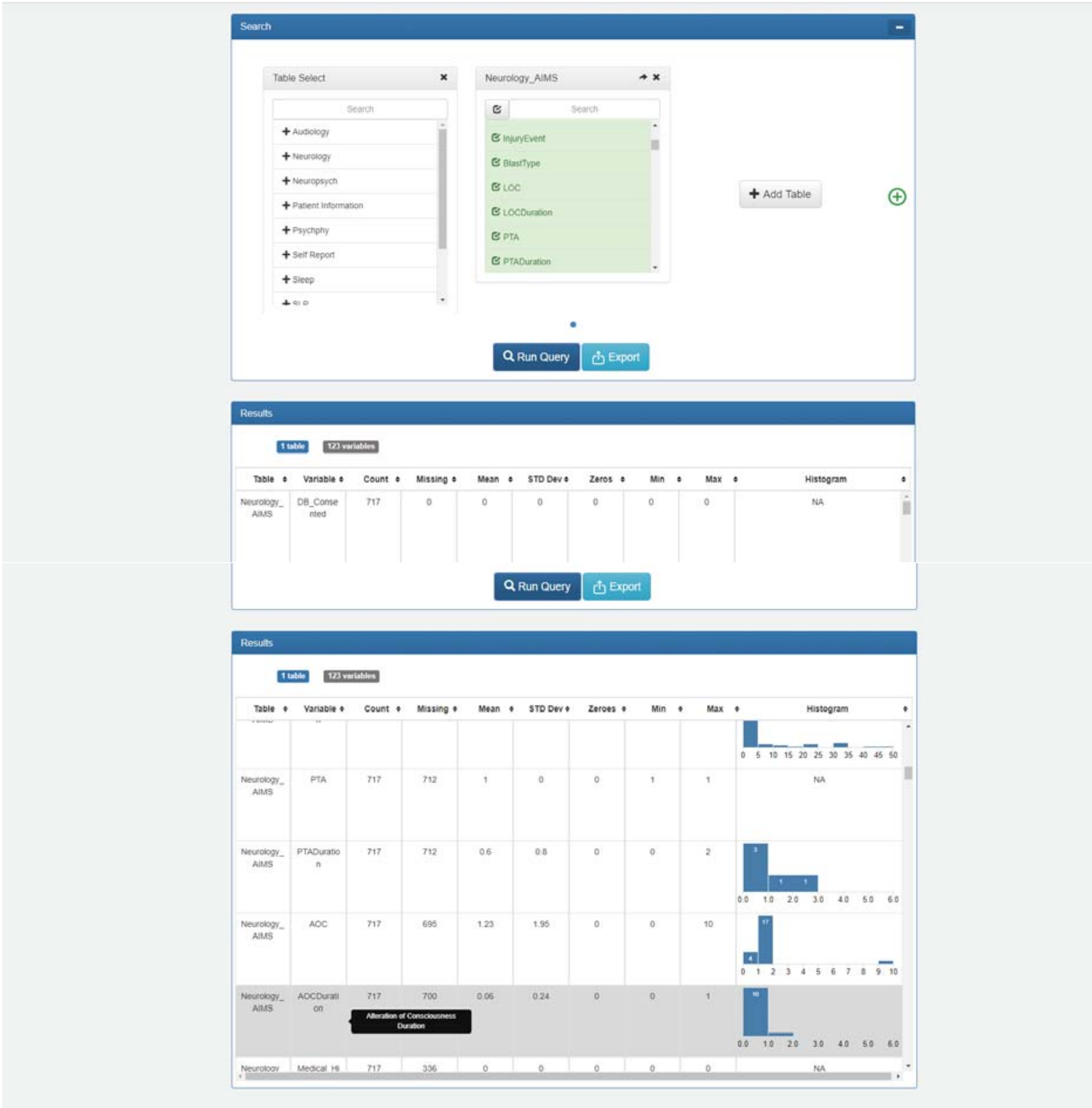


Figure 5: Screenshots of the Database Search tool designed to better explore clinical data that is available to research purposes.

5. [Aims 1.3, 1.7, 1.8, 1.9] A new module was prototyped to visually analyze clinical notes. The tool automatically highlights clinical concepts within unstructured provider notes. The new module can be used to compare clinical notes and highlight changes over time. In addition, the new module can be used to identify commonalities between different EHR notes.

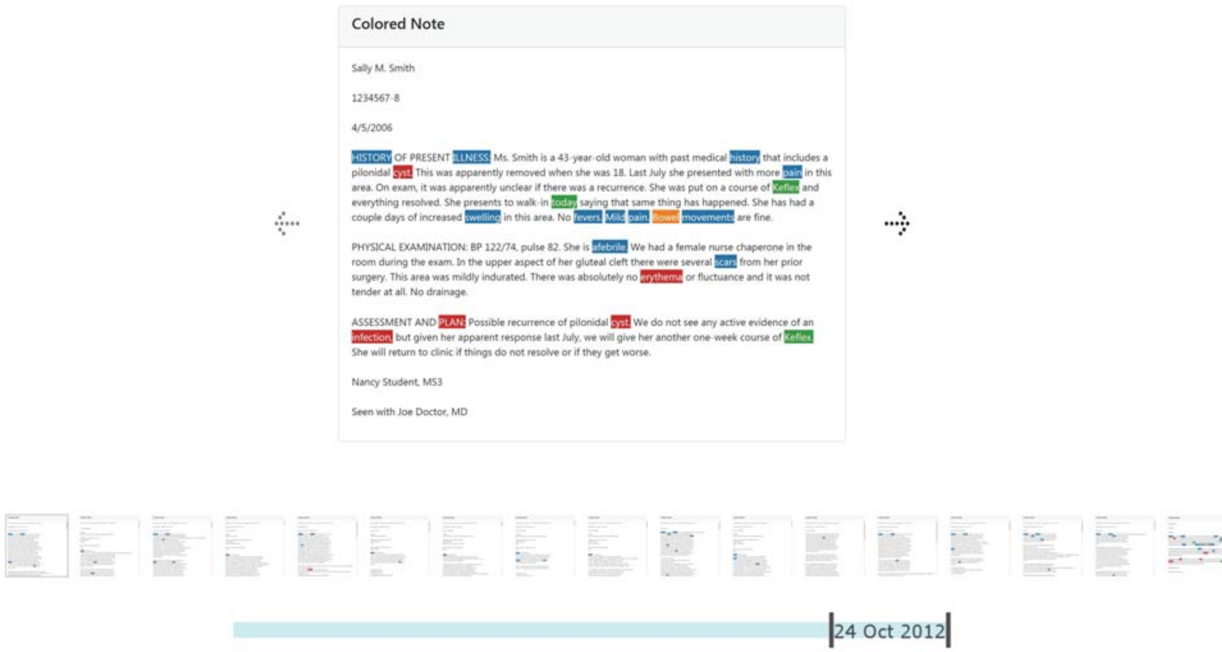


Figure 6: Screenshots of the module under development to analyze / compare clinical notes.

6. [Aims 1.3, 1.7, 1.8, 1.9] In order to better understand and validate the effectiveness of different visualization tools, we have developed a method to capture how users interact with the different systems. Given a clinical dataset and a visualization tool or dashboard, we designed an application to compare the correlations, patterns, and flows different groups follow to reach an answer for a specific clinical question.

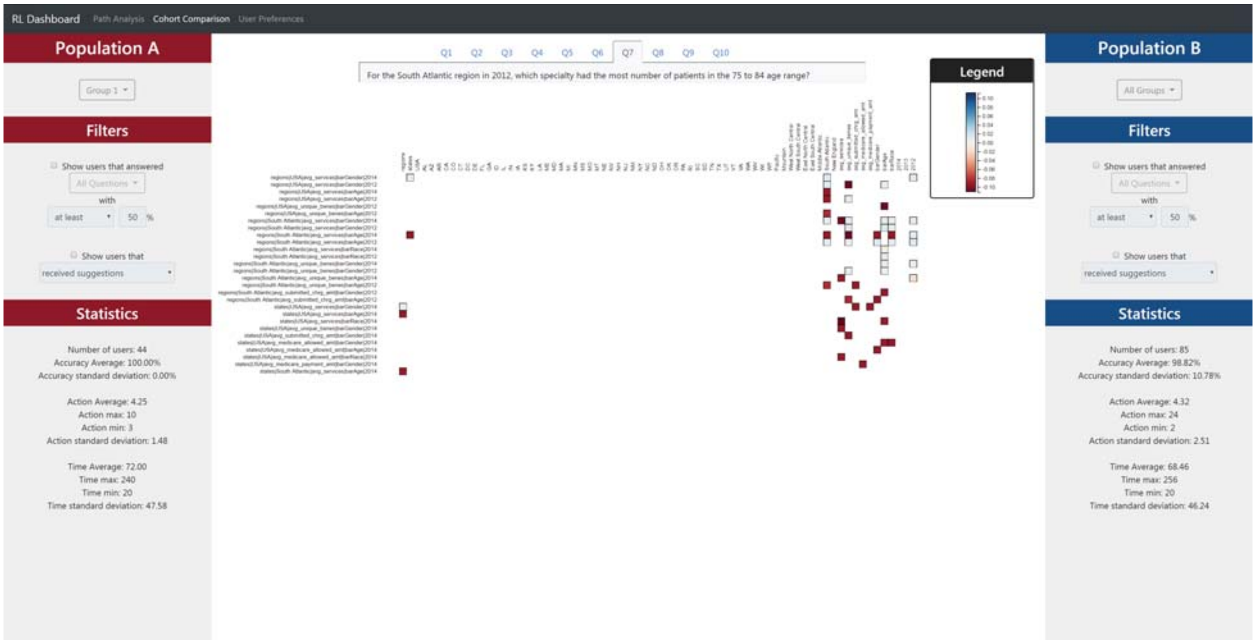


Figure 7: Screenshot of a tool designed to explore and understand how users interact with different visualization tools.

What were the major goals of the project?

The first major goal of this project was to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal (year 2) is to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

The first major goal of the second year was to iterate on our work for summarizing large, longitudinal datasets by expanding on our previous work and developing a novel visualization method for analyzing a large dataset of events. In addition, we validated the effectiveness of our visualization techniques throughout the year. The second major goal of the second year was to build a system that summarizes the various elements of a patient's EHR data. Through these two goals, we have addressed the many difficulties associated with clinical variables and data from both a big data and a singular patient-provider perspective.

The first major goal of the third year was to continue to enhance the existing tools from previous years by adding new visualization techniques, new validation approaches, and new data. That was accomplished by continuing the perform a literature review, develop the Necklace interface, continue to enhance the patient time-line framework, develop an effective interface for research clinicians to search at data, and by creating a systematic approach to analyze the patient interaction data. During the next year the priority will be to continue to enhance those tools and make them available to other researchers, investigators, and organizations.

What was accomplished under these goals?

During the first year, four different systems were prototyped and developed to perform visualization of tabular, hierarchical, and longitudinal data. First, **VisXplore** was enhanced to become a clinical data visualization system to perform group or single-subject analysis of multivariate tabular, hierarchical, or temporal clinical data. Second, **CoFlow** was developed as an interactive multi-view and exploratory visualization tool designed to analyze longitudinal EHR data. Third, a **graph-based visualization technique** was developed to visually explore the frequency of patients going from one specific clinical diagnosis to other diagnosis. Finally, a **visual summarization approach** was created and tested with thousands of mTBI patients. Each of the tool has a corresponding draft paper describing the design and techniques. See attachments.

During the second year, two different systems were extensively prototyped and developed to effectively summarize the various data elements that are present in Electronic Health Records (EHRs). First, a novel visualization method, **event summary diagrams**, and a corresponding system were built to enable for a large dataset of events to easily be understood through a top-down interactive exploration. This visualization was evaluated with a dataset of thousands of mTBI patients and shown to reduce the visual complexity and analytical capacity required compared to existing techniques. Second, a **timeline-based framework** for aggregating and summarizing EHRs was extensively researched, designed, and developed to overcome the challenges that exist in EHR systems where data integration is lacking and the disparate nature of data creates difficulties for clinicians. Through this framework, a clinician is able to view the entire history of a patient at multiple time scales and develop an understanding of the patient state over time. Each of these tools have a corresponding draft paper describing the design and techniques. See attachments.

During the third we continued to perform a literature review, developed the Necklace interface, continued to enhance the patient time-line framework, developed an effective interface for research clinicians to search at data, and by creating a systematic approach to analyze the patient interaction data. During the next year the priority will be to continue to enhance those tools and make them available to other researchers, investigators, and organizations.

What opportunities for training and professional development has the project provided?

“Nothing to Report.”

How were the results disseminated to communities of interest?

Some of the prototype visualization tools were demonstrated during the 2017 Workshop on Visual Analytics in Healthcare and at the 2017 American Medical Informatics Association (AMIA). Three additional papers describing the other systems are currently in draft mode. See attached documents.

What do you plan to do during the next reporting period to accomplish the goals?

Continue to enhance the different system, continue the validation process, and prepare the applications for distribution.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

The four prototype systems that have been designed have generated great interest among multiple providers, researchers, and administrators. Two senior individuals at the Defense Health Agency (DHA) have seen the systems and are interested in looking into how we can integrate some of those tools within the DHA enterprise enclave. In addition, widely recognized researchers from Johns Hopkins University (JHU) are interested in how to used our visualization techniques for population health.

What was the impact on other disciplines?

The impact of our work is touching multiple disciplines and research domains including clinical informatics, health IT, computer science, medicine, and population health.

What was the impact on technology transfer?

"Nothing to Report."

What was the impact on society beyond science and technology?

"Nothing to Report."

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

"Nothing to Report"

Actual or anticipated problems or delays and actions or plans to resolve them

The project and actions are a little bit behind schedule due to the challenges of finding qualified candidates that can obtain the credentials needed to work within a DoD facility.

Changes that had a significant impact on expenditures

No changes on expenditure.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

"Nothing to Report"

6. PRODUCTS:

- **Publications, conference papers, and presentations**

  Journal publications. Nothing to Report

  Books or other non-periodical, one-time publications. Nothing to Report

  Other publications, conference papers, and presentations

  - Filip Dabek and Jesus J Caban "VisXplore: Flexible Visualization System for Analyzing Complex Clinical Datasets", Workshop on Visual Analytics in Healthcare, ACM Digital Library, Oct. 2014
  - Filip Dabek, J. Chen, A. Garbarino, and Jesus J. Caban, "Visualization of Longitudinal Clinical Trajectories using a Graph-based Approach", Workshop on Visual Analytics in Healthcare, ACM Digital Library, Oct. 2015
  - Akshay Peshav, Jian Chen, and Jesus J. Caban, "CoFlow: Interactive Visual Exploration of Temporal Encounters in Electronic Health Records" [Draft]
  - Filip Dabek, Jian Chen, and Jesus Caban, "Visual Summarization of a Collection of Temporal Sequences using Adaptive Frequency Mining and Graph-based Event Modeling" [Draft]

- **Website(s) or other Internet site(s)**

"Nothing to Report"

- **Technologies or techniques**

The design and development of our different visualization tools have produced novel techniques including:

- Novel graph-based approach to visualize clinical trajectories
- New pixel-based visualization method that works as a look-ahead tool for patients
- Novel sequence modeling algorithm to summarize longitudinal trajectories

1. **Inventions, patent applications, and/or licenses**

"Nothing to Report"

2. **Other Products**

- The four different software tools

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS
What individuals have worked on the project?

Name: Jesus Caban, PhD
Project Role: PI
Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Jian Chen, PhD
Project Role: Co-PI
Contribution to Project: Dr. Chen has helped assist in meetings and researched existing visualization techniques in order to identify a new technique to deploy for the visual analytics framework.

Name: Elizabeth Jimenez
Project Role: Developer
Contribution to Project: Ms. Jimenez has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Filip Dabek
Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)
Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

"Nothing to Report."

What other organizations were involved as partners?

8. SPECIAL REPORTING REQUIREMENTS:

See attachments.

- W81XWH-15-2-0016 Year 3 Quarter Reports.pdf: copy of all the quarterly reports for year #1.
- W81XWH-15-2-0016 Year 3 Supplements.pdf: copy of all the papers and draft papers.

8. APPENDICES:

See attachment.

- W81XWH-15-2-0016 Year 3 QuadChart.ppt

Supplemental Material

DHA Defense Health Agency

NICoE NATIONAL INTREPID CENTER OF EXCELLENCE

PROTECT PROJECT SUSTAIN

An Interactive Visualization Framework to Support Exploration and Analysis of TBI/ PTSD Clinical Data

Jesus J Caban, PhD
National Intrepid Center of Excellence,
Walter Reed National Military Medical Center

Award #W81XWH-15-2-0016

Disclosure

The views expressed in this presentation are those of the authors and do not reflect the official policy of the Department of Army/Navy/Air Force, Department of Defense, or U.S. Government.

Project Information

- **Organization / key partners / institutes:** The Geneva Foundation
- **Award #:** W81XWH-15-2-0016
- **Award Mechanism:** Cooperative Agreement
- **Solicitation:** USAMRMC BAA
- **Principal Investigator:** Jesus Caban
  - Key Sub-Awards: (Co-PIs)
- **Total Cost/Budget:**
- **Period of Performance:** 15 Apr 2015 – 14 Apr 2018[1]
- **Grants / Contract Officer Representative:** Gay Hayden
- **Grants / Contract Specialist:** Karen L. Petrore
- **Related government funding:** N/A

18

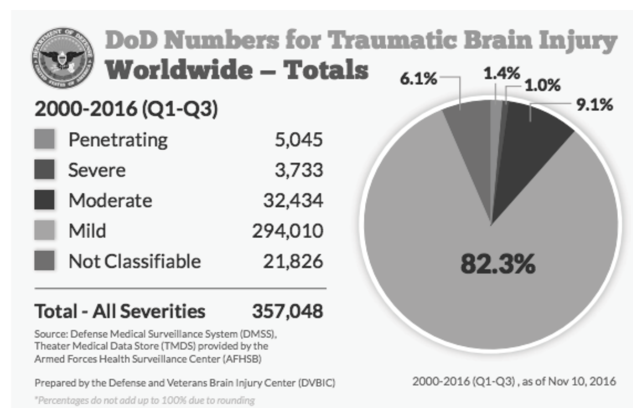
[1] Requested no-cost extension on Feb 2018

Introduction

- Traumatic Brain Injury (TBI) continues to be labeled a the signature injury of the wars in Iraq, Afghanistan, and other recent conflicts

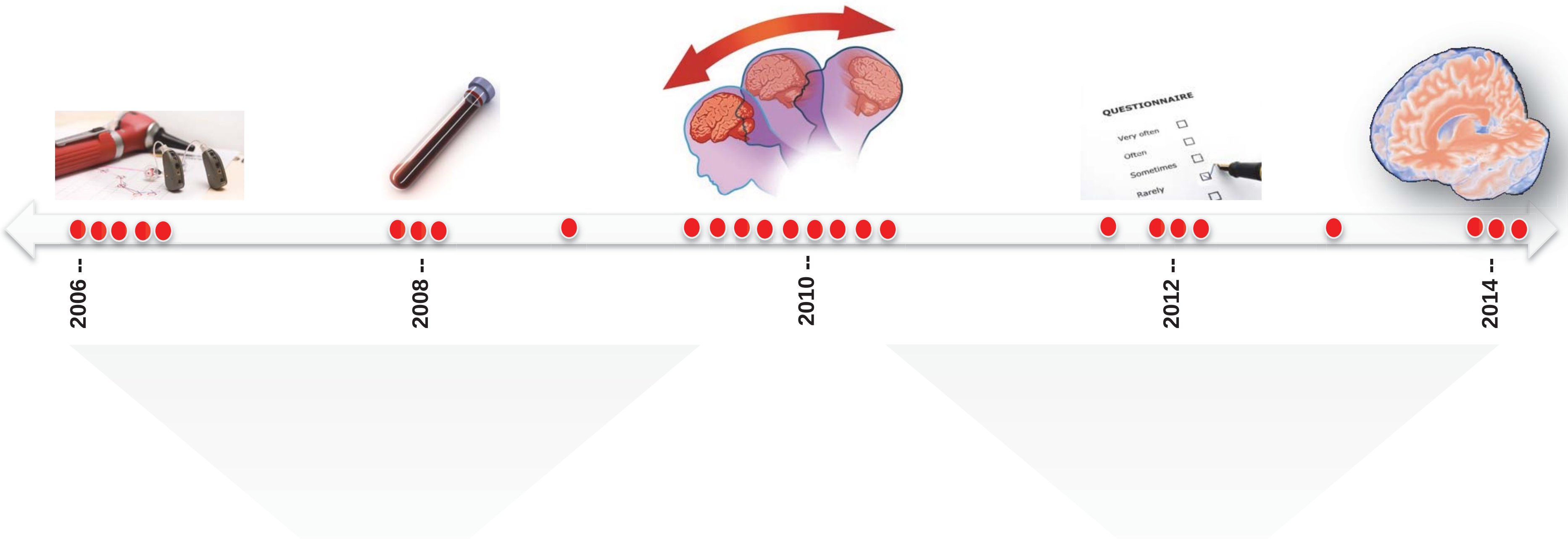
- Over 360,000 TBIs occurred in the DoD between 2000 – 20161
  - ~294,000 are cases of Active Duty mTBI

- mTBI may result in chronic symptomatology
  - Up to 25% of deployed service members report mTBI symptoms
  - Following mTBI, most people are expected to recover within a short period of time, however, neurologic, cognitive, and physical symptoms may persist for days to months or longer, requiring ongoing medical treatment



Natural History of TBI

- The National Intrepid Center of Excellence (NICoE) and Intrepid Spirit Centers (ISCs) core mission is to *improve the lives of patients and families impacted by TBI through <u>excellence</u> and <u>innovation.</u>*



Pre-mTBI
(e.g. Pre-existing Conditions)

Post-mTBI

History of mTBI

NICoE Informatics

Provide services, tools, data, and analysis that can have profound impact in the treatment, management, and overall understanding of TBI.

Solutions Development

- CMT / TBI Portal
- WIIR - TBI
- TBI Document Center
- NICoE Online Referrals
- NICoE Research App

Metrics and Dashboards

- NICoE Dashboards
- NCR/WRNMMC Dashboards
- TBI Dashboards
- Support NICoE Strategic Goals
- Health Outcomes

Data Management

- Data collection
- Data documentation
- Data requests
- Data quality control
- Data Lifecycle control

Informatics Research

- Data Analysis
- Data Modeling (stats & ML)
- Data Mining and Forecasting
- Proposal Development
- USUHS Informatics Partnership

HealthIT

- IT Support
- IT Sustainment
- Data backups
- TeleHealth support

NICoE Informatics

Provide services, tools, data, and analysis that can have profound impact in the treatment, management, and overall understanding of TBI.

Solutions Development

- CMT / TBI Portal
- WIIR - TBI
- TBI Document Center
- NICoE Online Referrals
- NICoE Research App

Metrics and Dashboards

- NICoE Dashboards
- NCR/WRNMMC Dashboards
- TBI Dashboards
- Support NICoE Strategic Goals
- Health Outcomes

Data Management

- Data collection
- Data documentation
- Data requests
- Data quality control
- Data Lifecycle control

Informatics Research

- Data Analysis
- Data Modeling (stats & ML)
- Data Mining and Forecasting
- Proposal Development
- USUHS Informatics Partnership

HealthIT

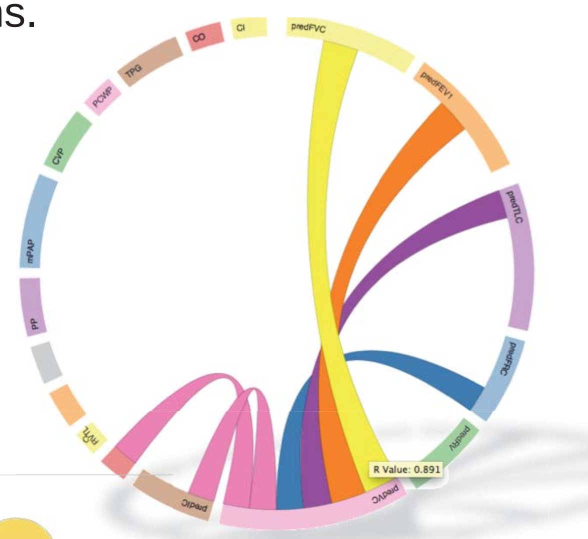
- IT Support
- IT Sustainment
- Data backups
- TeleHealth support

22

How can we explore clinical data?

Overview of the Research Project

- We propose to design and validate an interactive visual analytics framework that:
  - clinicians assessing TBI/PTSD patients can use to explore and analyze clinical data
  - researchers can use to hypothesize new research questions.



- Primary aims of this project are:
  1. design a visual analytic framework that combines multiple clinical measurements
  2. validate the effectiveness and usability of different visualization techniques

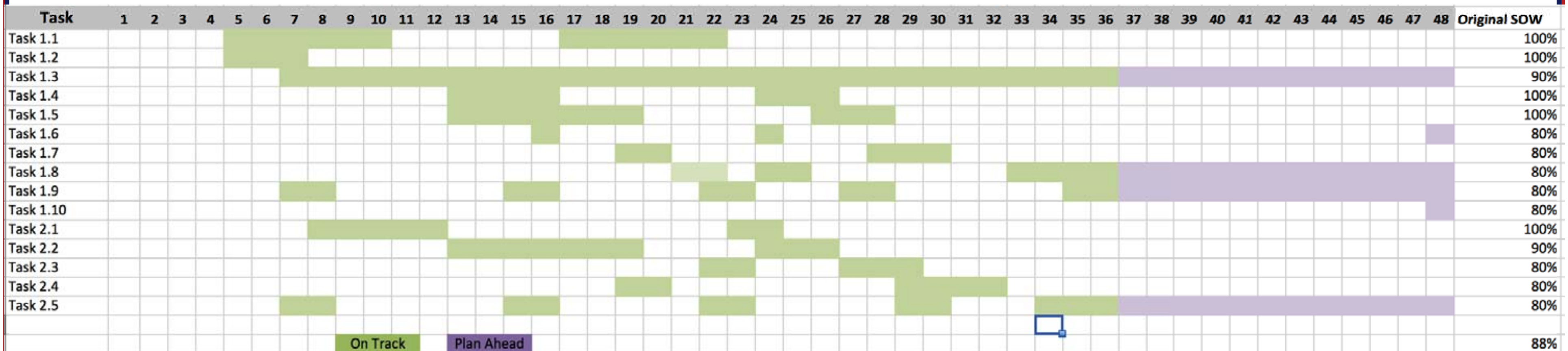
Research Question/Hypothesis & High Level Objectives

- **Objective 1:** The design and development of an intuitive framework to visually analyze and explore a large number of clinical variables.

- **Objective 2:** Perform a usability study to validate application with clinical and research staff treating service members diagnosed with TBI/PTSD

- **Objective 3:** Establish a research community. The software application will be shared with researchers and clinicians at different DoD and federal organizations as a validated software application to visually integrate and analyze large number of clinical variables.

Tasks & Milestone Update



Task	Timeline	Site 1	Site 2
Major Task 1: Design and development of clinical visual analytics framework	Months	NICoE	UMBC
Subtask 1: New software developer and research assistant will be hired and brought up-to speed about our research, the expectations, and deliverables of the project.	1-3	Dr. Caban	Dr. Chen
Subtask 2: Team will compare and evaluate existing data visualization libraries and software tools that could be used to develop a visual analytics framework to explore and analyze clinical data. In particular, Data-Driven Documents, ProtoVis, Flare, and Prefuse will be assessed for their feasibility of extending them to illustrate dependencies among many clinical variables. In addition, software applications such as Tableau, QlikView, and Spotfire will be evaluated.	1-3	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant
Subtask 3: Design a modularized visual analytics framework capable of loading, filtering, and illustrating any number of clinical variables.	3-18	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant
Subtask 3: Obtain previously collected multi-modal clinical data for 300 mTBI/PTSD subjects and load data into the visualization system. The data will be obtained from the NICoE at no cost to this project.	*9-12*	Dr. Caban	
Subtask 4: Perform preliminary validation of different visualization techniques	9-15	Dr. Caban	Dr. Chen TBD – Research Assistant
Subtask 5: Prepare and submit research progress report	12	Dr. Caban	Dr. Chen
Subtask 6: Update application based on preliminary validation with clinicians and researchers.	15-16	TBD - Software Developer	TBD – Research Assistant
Subtask 7: Update system based on usability study and suggestions received from users.	20-24	TBD - Software Developer	TBD – Research Assistant
Subtask 8: Share initial framework with different federal organizations including WRNMMC, MHS,	23-24	Dr. Caban	Dr. Chen

Major Task 2: Validation of visualization systems in clinical settings			
Subtask 1: Write and submit expedite IRB protocol to perform usability study	4-8	Dr. Caban	Dr. Chen
Subtask 2: Preliminary validation of the framework with small number of students, clinicians and researchers.	9-15		Dr. Chen TBD – Research Assistant
Subtask 3: Perform formal user study of the usability of the system. The system will be tested with over 20 clinicians and 40 students / researchers from Walter Reed and the University of Maryland System.	16-19	Dr. Caban	Dr. Chen TBD – Research
Subtask 4: Analysis of usability study.	20-22	Dr. Caban	Dr. Chen
Subtask 5: Share and/or publish findings of our usability study.	22-24	Dr. Caban TBD - Software Developer	Dr. Chen TBD - Research Assistant

Research Methodologies

Leveraging existing, approved and non-funded DoD Project for data:

- **Protocol:** DoD IRB #374953
- **PI:** Jesus J Caban, PhD
- **Approval date**: May 30th, 2012 ➔ Dec 2018

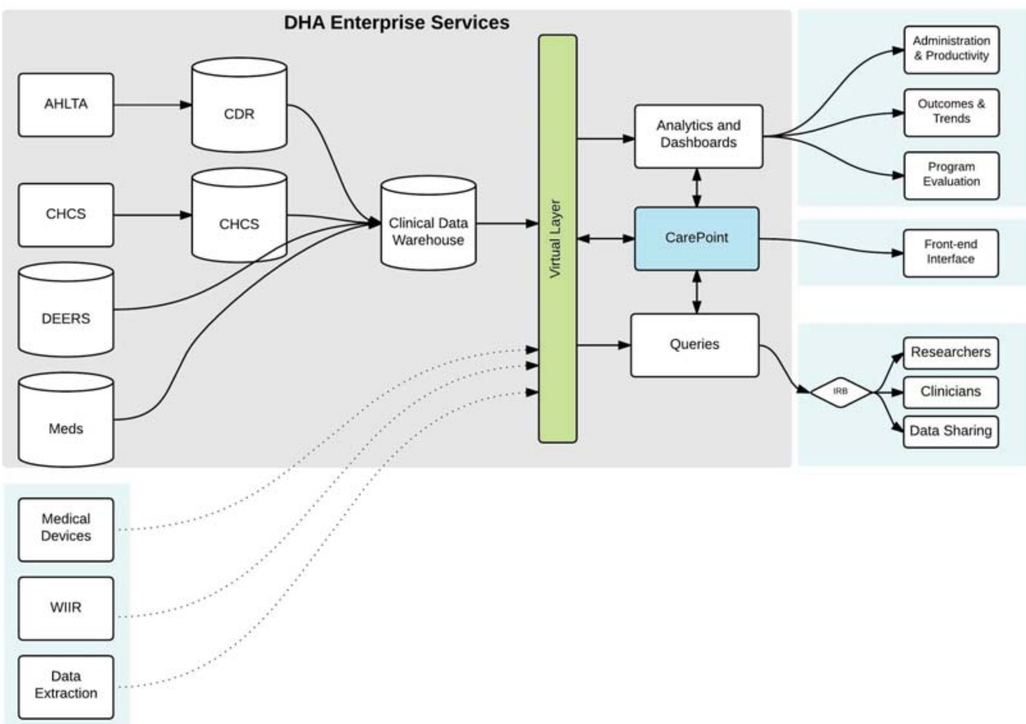
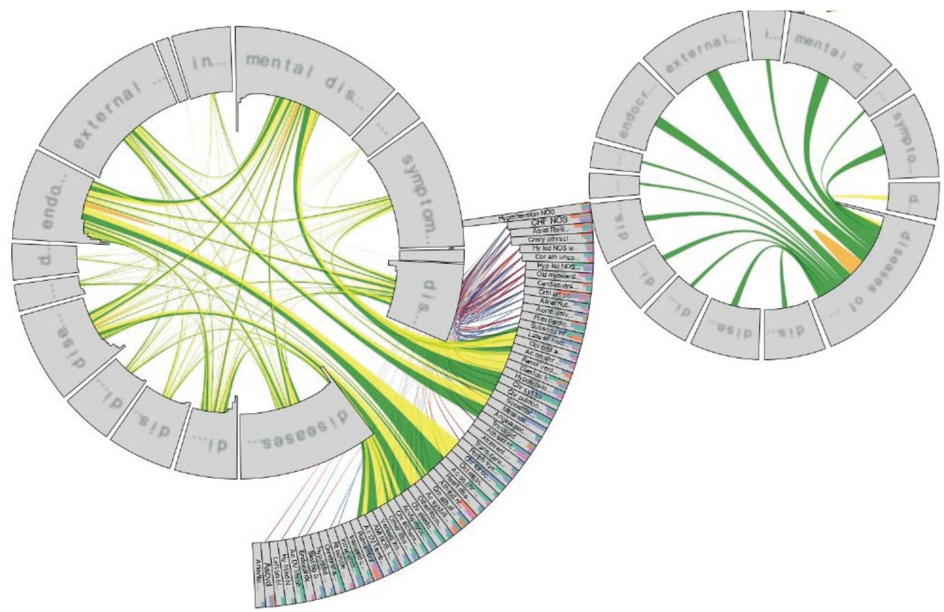


Table #1: Descriptive Statistics of our Dataset

Number of Patients		112,738
Gender	Male	89.6%
	Female	10.4%
Age	18-24	35.8%
	25-34	40.4%
	35-44	18.3%
	45-64	5.4%
	Other	0.1%
Branch of Service	US Army	77,172
	US Marine Corps	15,607
	US Air Force	9,823
	US Navy	9,011
	Other	1,125
SMs with evident War-Related TBIs		44,451
Number of TBI-related Clinical Encounters		17,189,609
Mean number of encounters per patient per year		24
Number of Medication Transactions		9,964,413
Number of Laboratory Results		3,664,461
Number of Radiology Notes		1,123,196

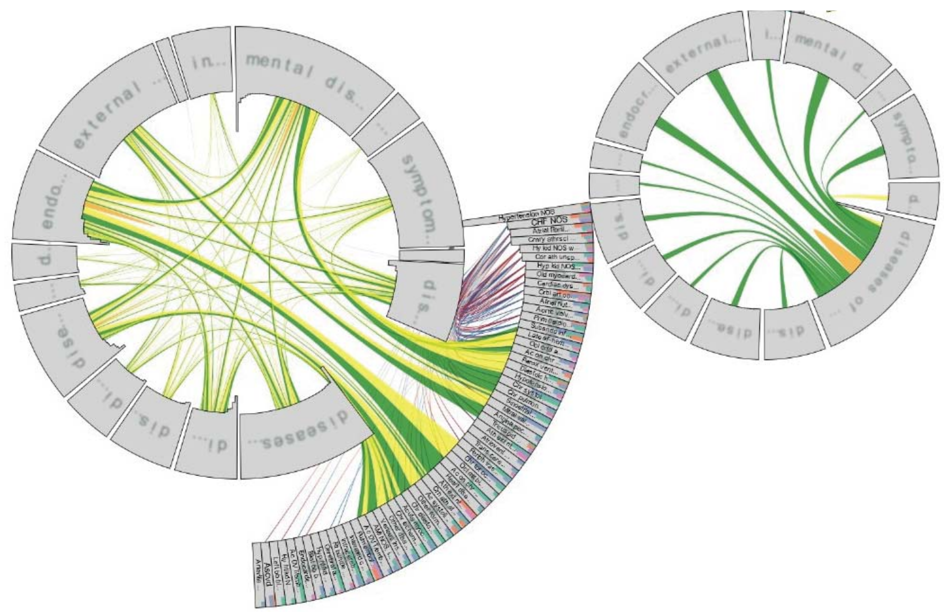
Taxonomy of Clinical Data Types

1. 1D/Linear
   - lists of data items, organized by a single feature (e.g., EKG)
2. 2D/Planar (X-Ray, Geospatial)
   - Geographical distribution of a specific disease
3. 3D/Volumetric
   - medical imaging (CT, MRI, etc…)
4. Temporal
   - T=time series
5. nD/Multidimensional
   - category proportions (e.g. Neurop
6. Tree/Hierarchical
   - ICD codes, many medical tests
7. Network
   - Clinical research

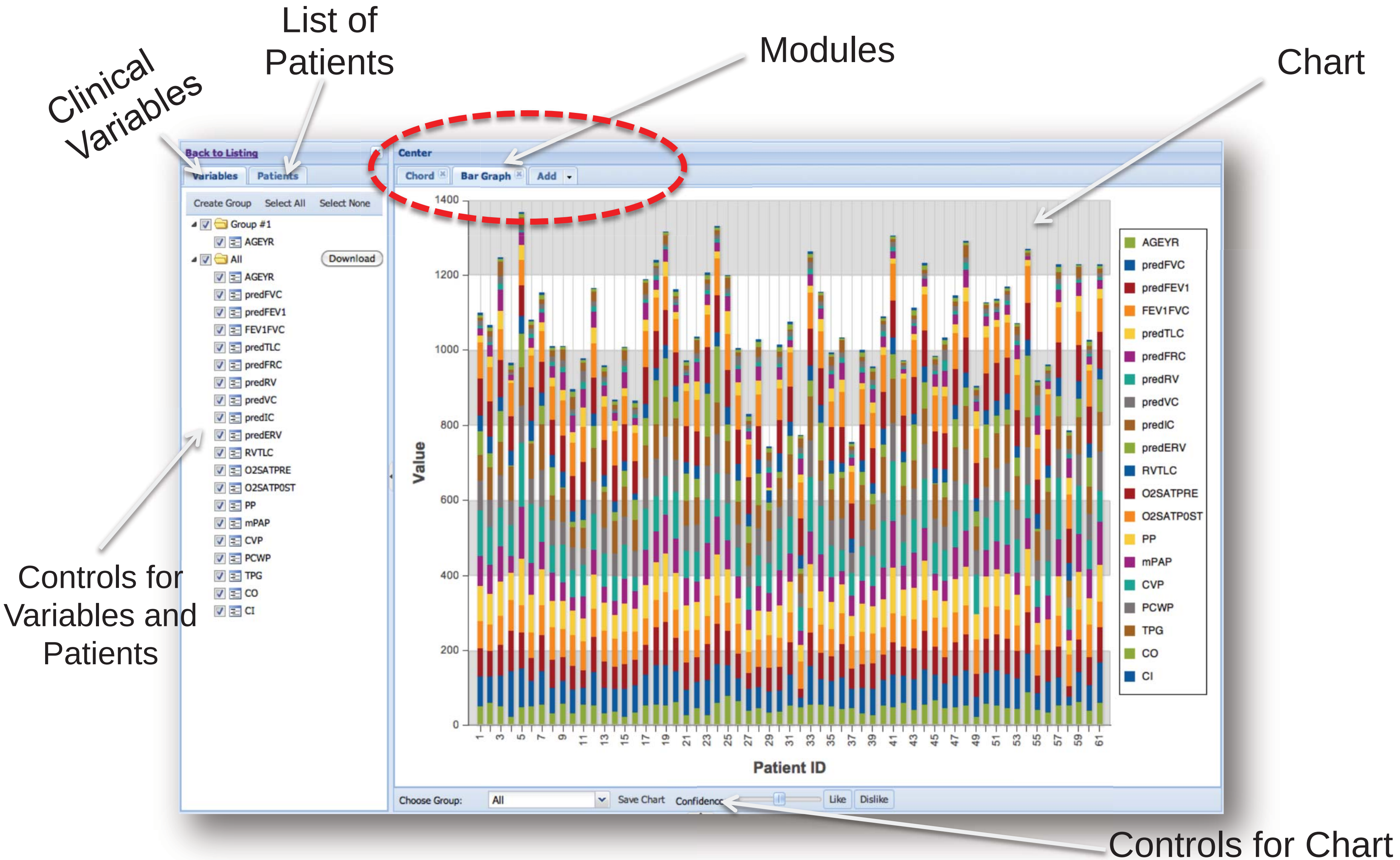


Taxonomy of Clinical Data Types

1. 1D/Linear
    - lists of data items, organized by a single feature (e.g., EKG)
2. 2D/Planar (X-Ray, Geospatial)
    - Geographical distribution of a specific disease
3. 3D/Volumetric
    - medical imaging (CT, MRI, etc…)
4. Temporal
    - T=time series
5. nD/Multidimensional
    - category proportions (e.g. Neurop
6. Tree/Hierarchical
    - ICD codes, many medical tests
7. Network
    - Clinical research

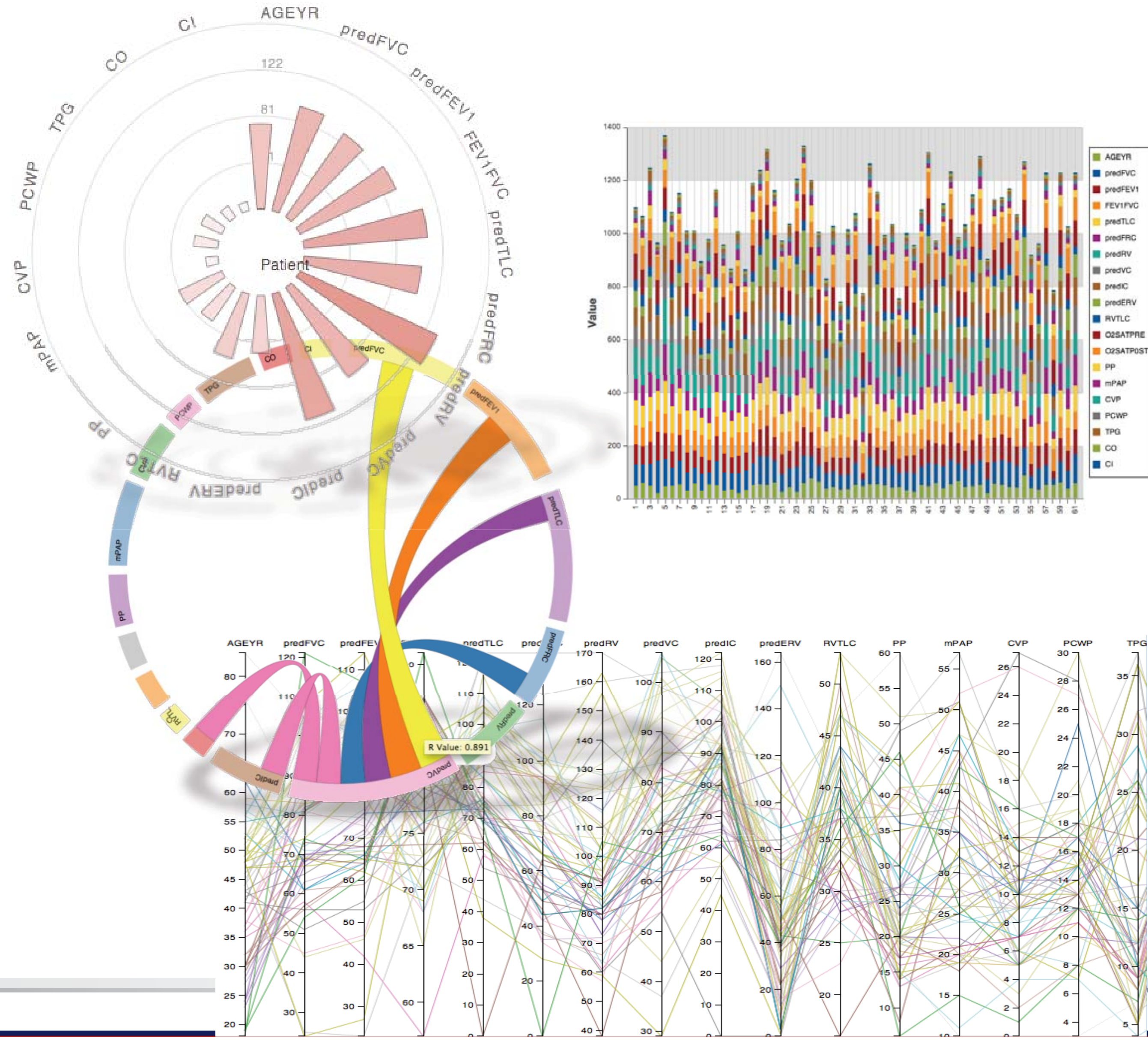


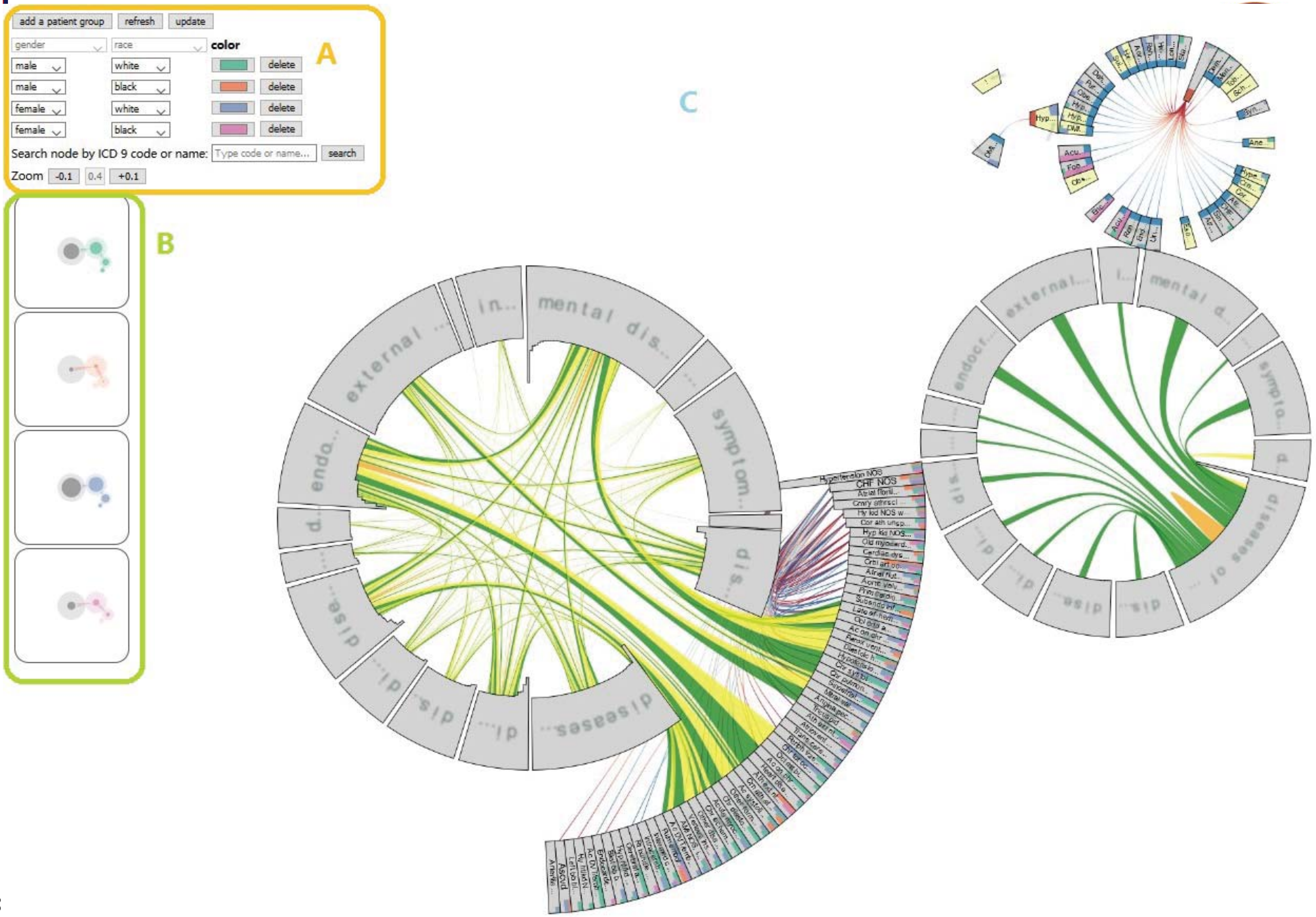
Results - Last Year



Results - Last Year

1. Parallel
2. Chord
3. Windrose
4. Lines
5. Statistics
6. Area Chart
7. Histogram
8. Pie Chart
9. Bar Graph
10. Radar Fill

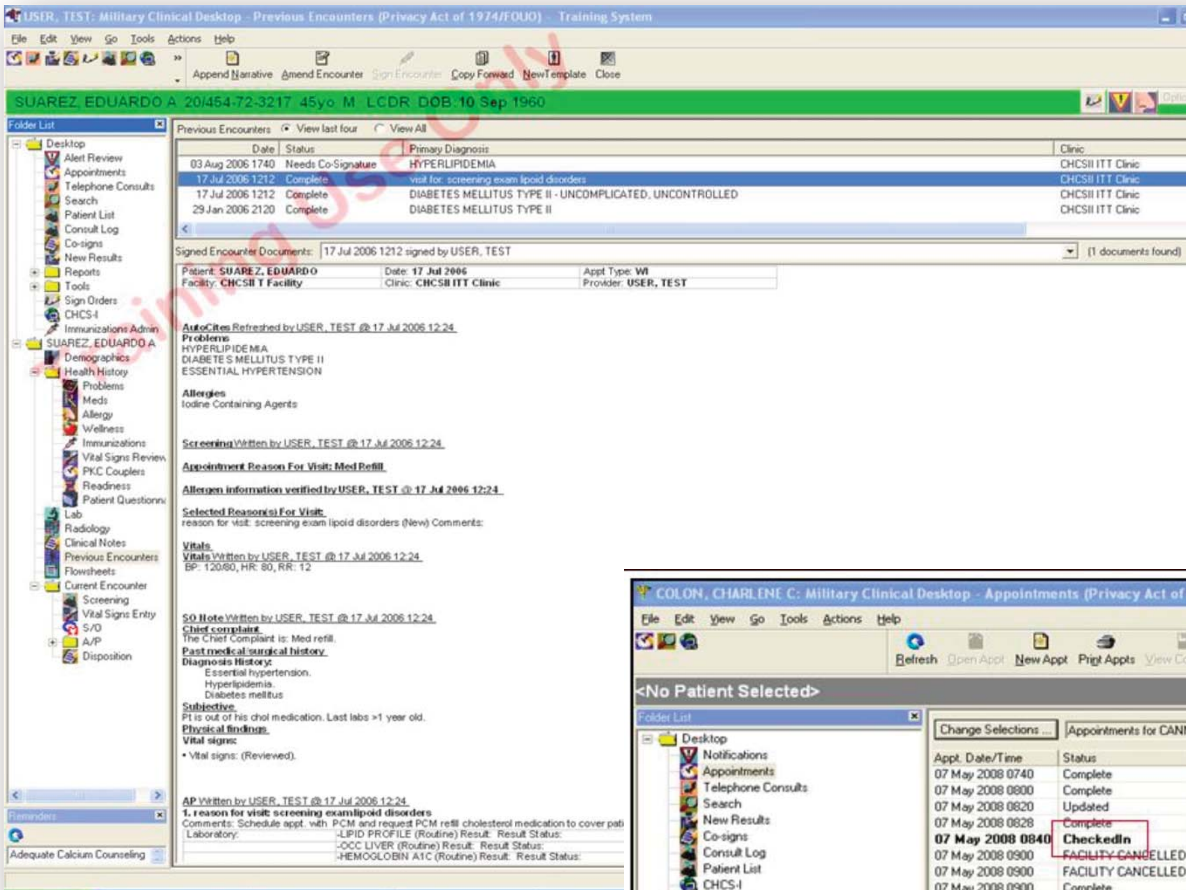


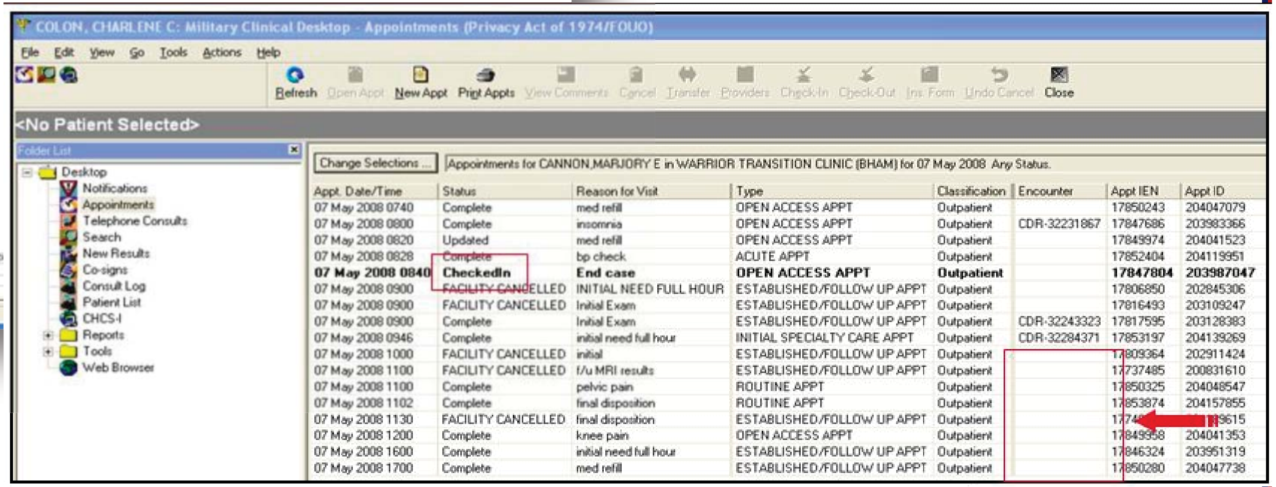

add a patient group
refresh
update
gender
race
color
male
white
delete
male
black
delete
female
white
delete
female
black
delete
A
Search node by ICD 9 code or name:
Type code or name...
search
Zoom
-0.1
0.4
+0.1
B
C
external
mental dis...
symptom...
endo...
diseases
endocr...
mental d...
sympto...
external...

32

Visually Exploring EHR data for a Single Patient

Defense Health Agency





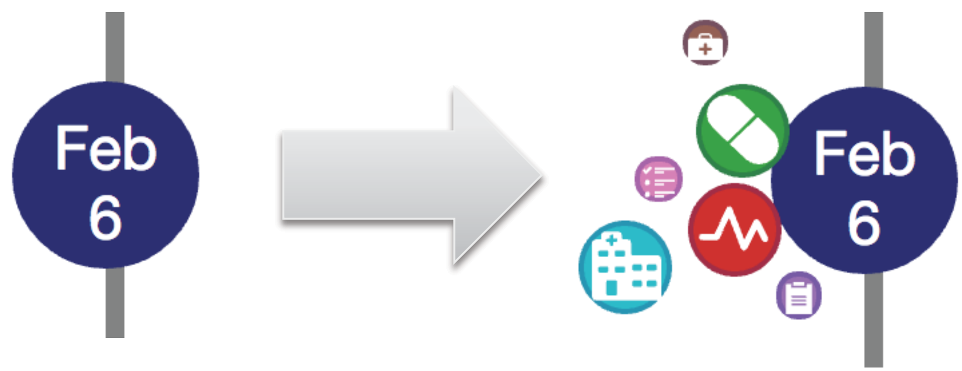
Timeline Construction

- Represent the patient’s longitudinal history as a timeline.
- Encodes multiple attributes to a single encounter

Visually encode clinically relevant information into each node.

Summary Node

- Encode information about the clinical data elements into attribute nodes:
  - Lab Test Results
  - Medications
  - Vitals
  - Clinical Note
  - Diagnosis Codes
  - Procedure Codes
  - Radiology Notes



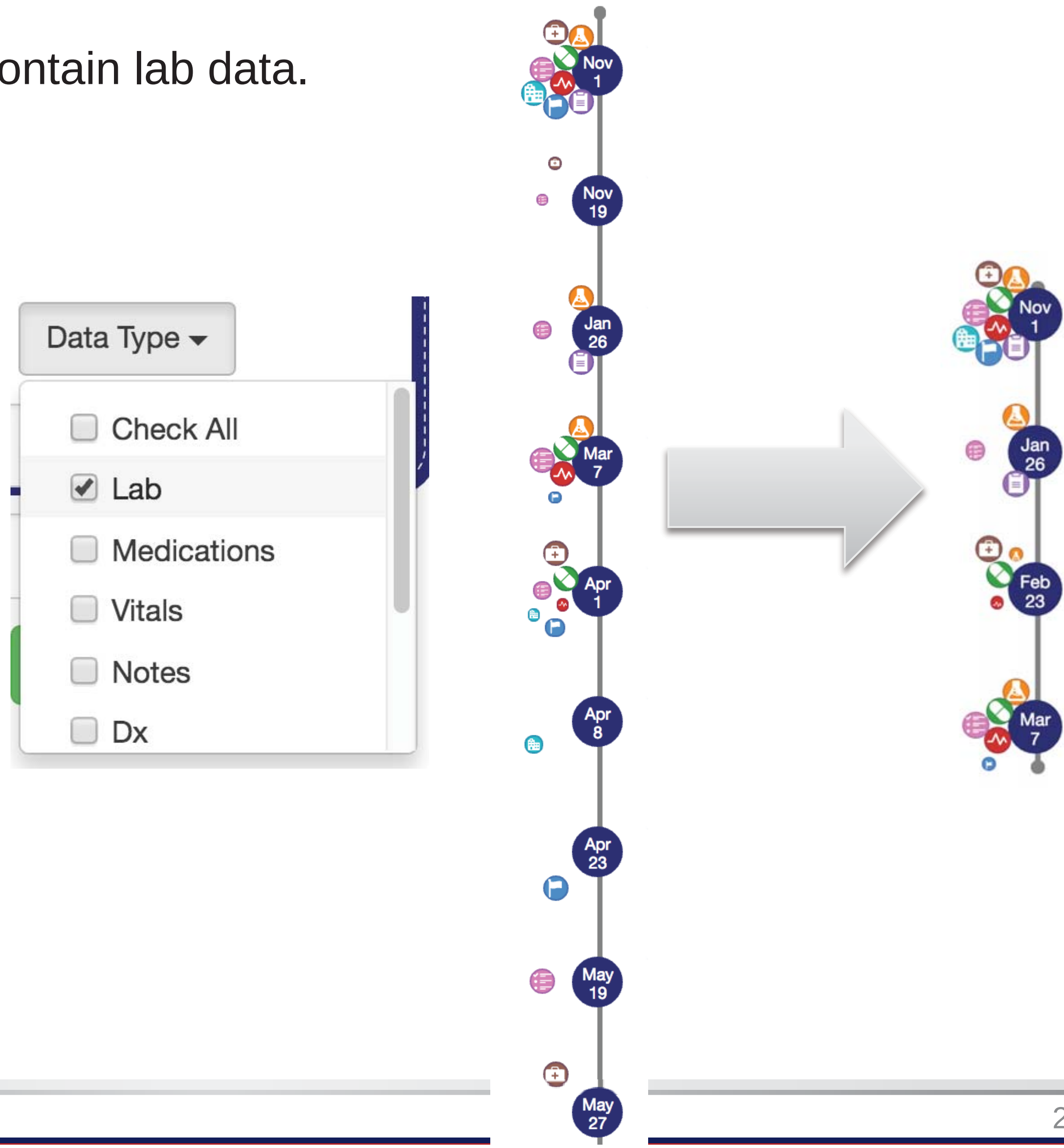
36

Attribute Nodes

- Size of each attribute node is varied based on:
  - » Volume of data collected for the encounter
  - » Clinical significance of the captured data.

Filtering Example #1

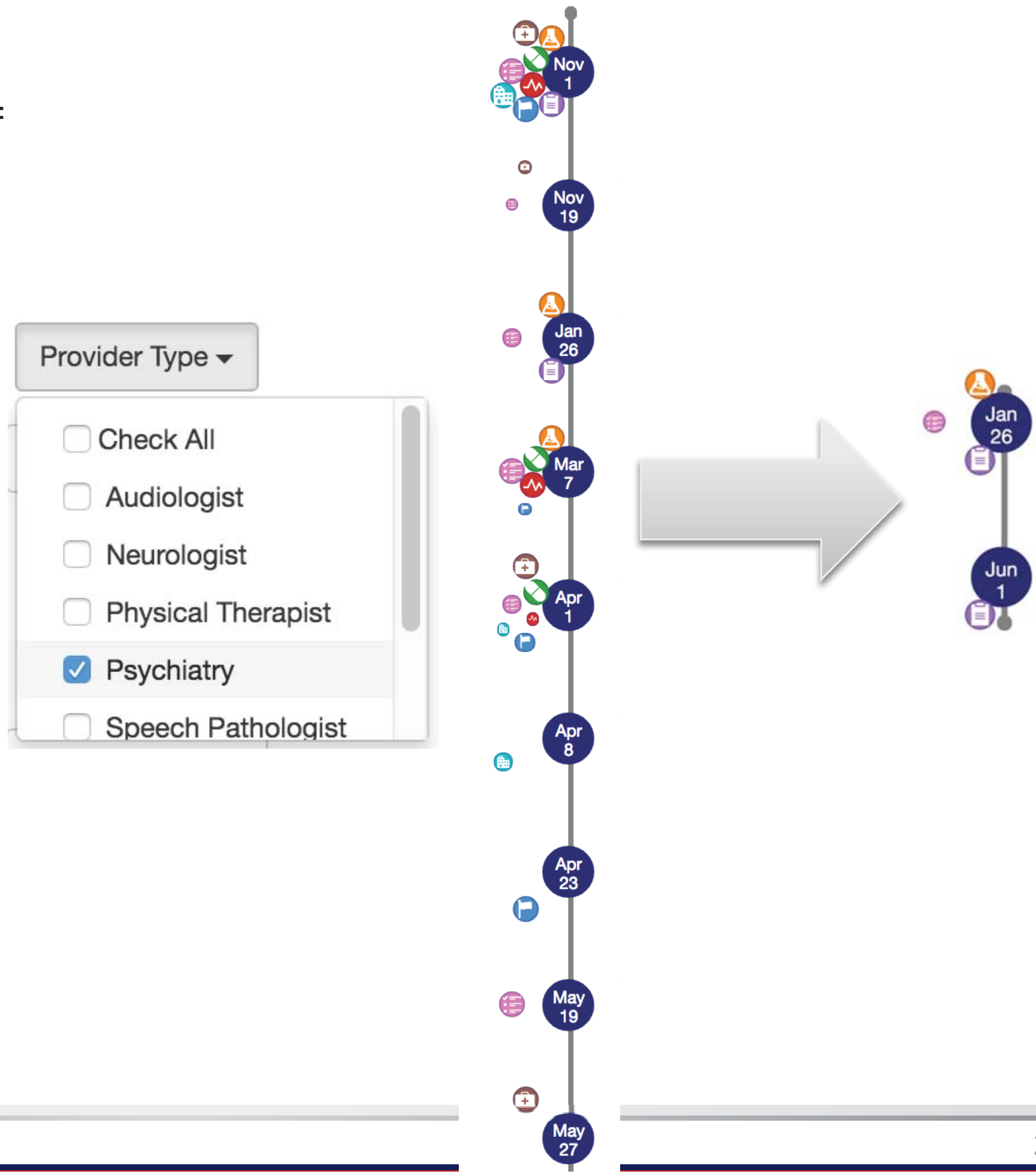
- Only show encounters that contain lab data.



38

Filtering Example #2

- Only show encounters that occurred with a specific type of provider.

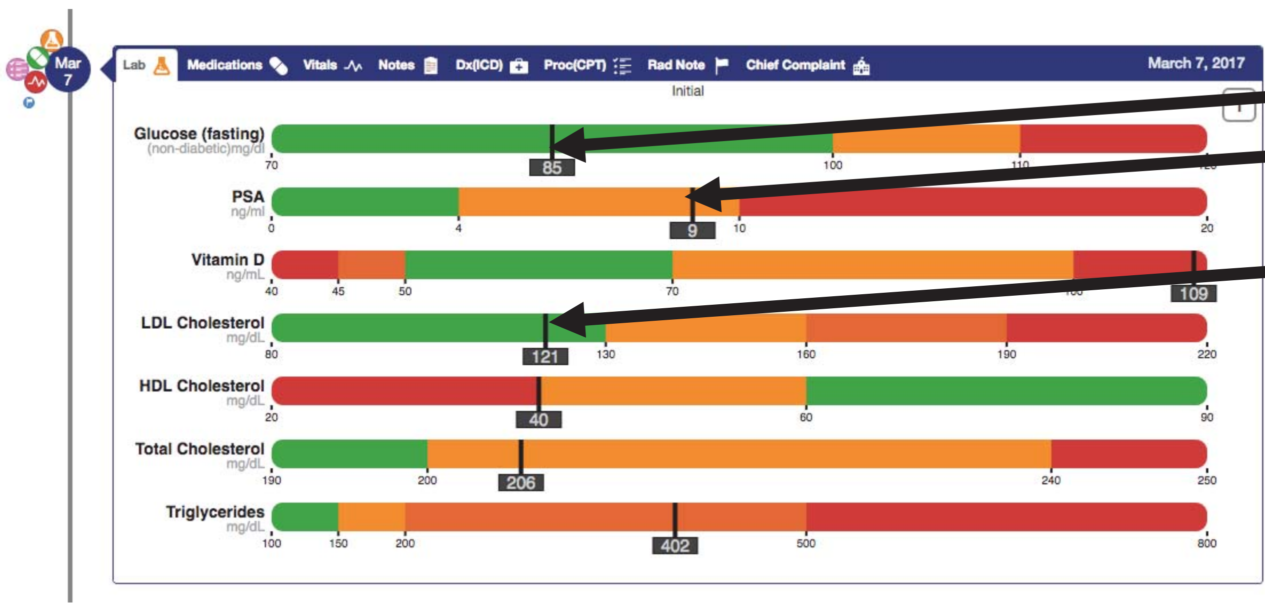


Timeline Interface

- Allows providers to find the encounter that they want to look at.
- Need to:
  - » Allow to drill down further.
  - » Utilize visualization to communicate each attribute.

Data View - Labs

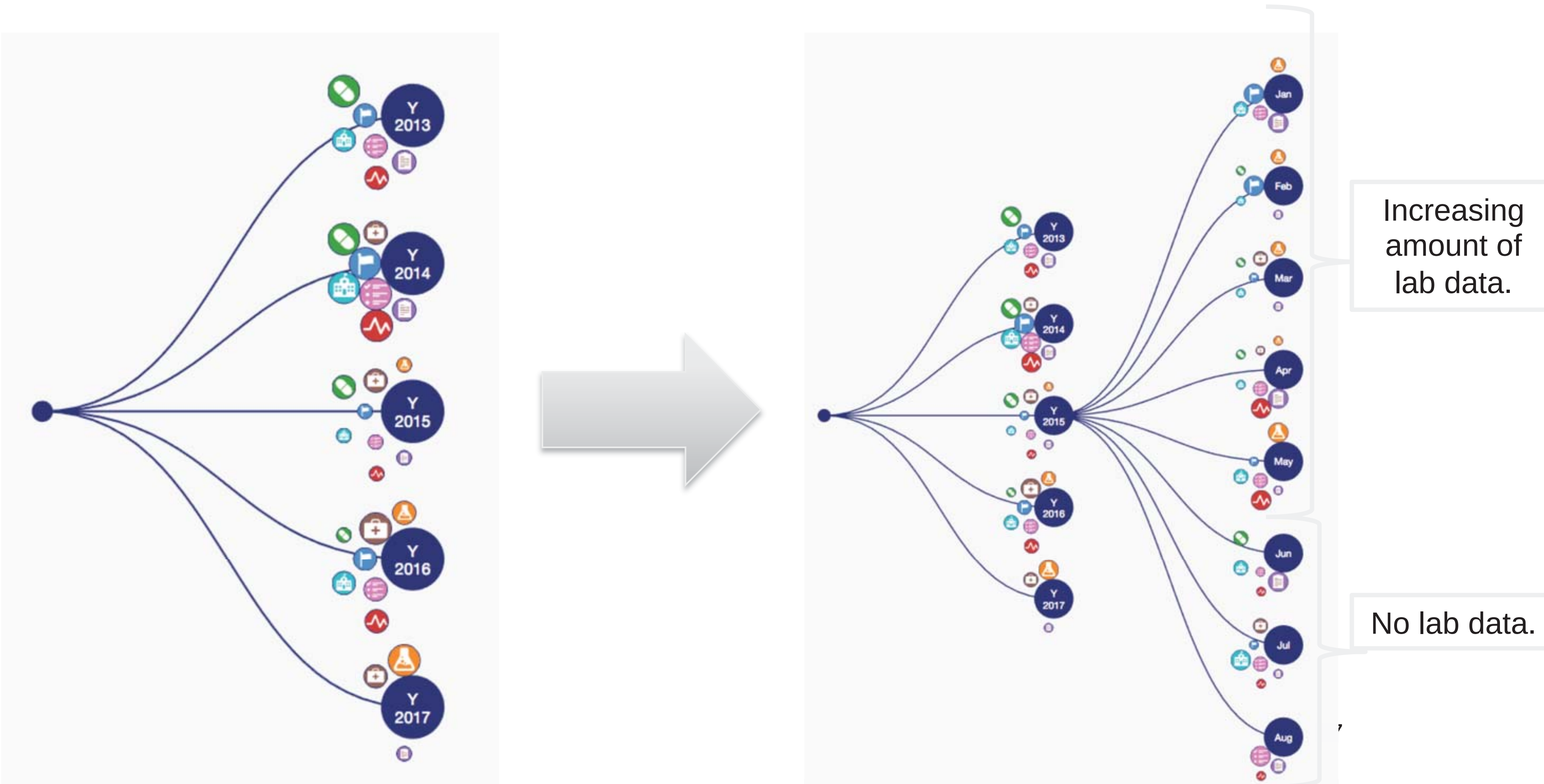
- Visualization of lab tests conducted for a given encounter.
- Color scale indicating whether patient is within normal range.



Markers indicating the patient's measured value against the possible range.

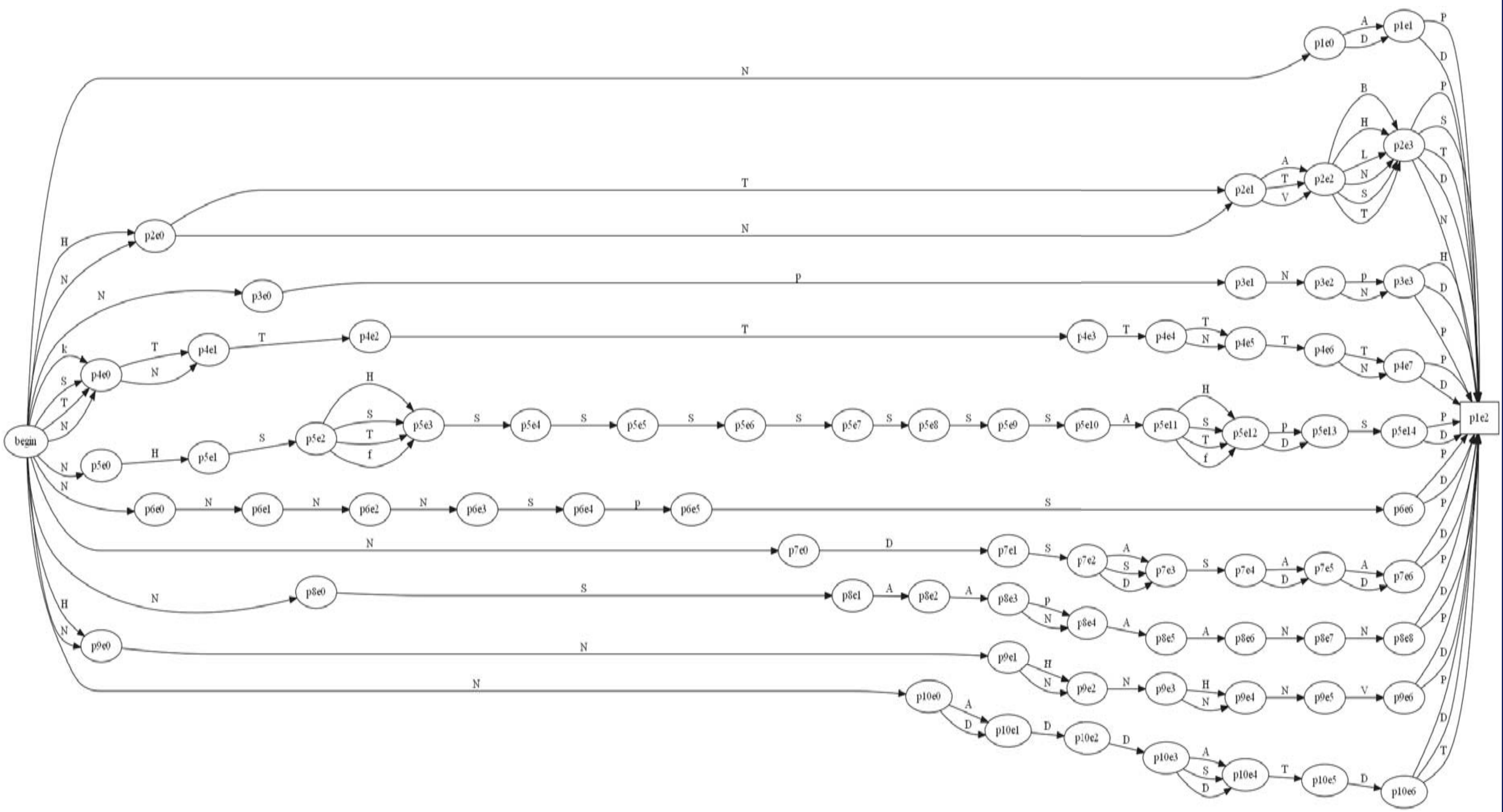
41

Patient Summary



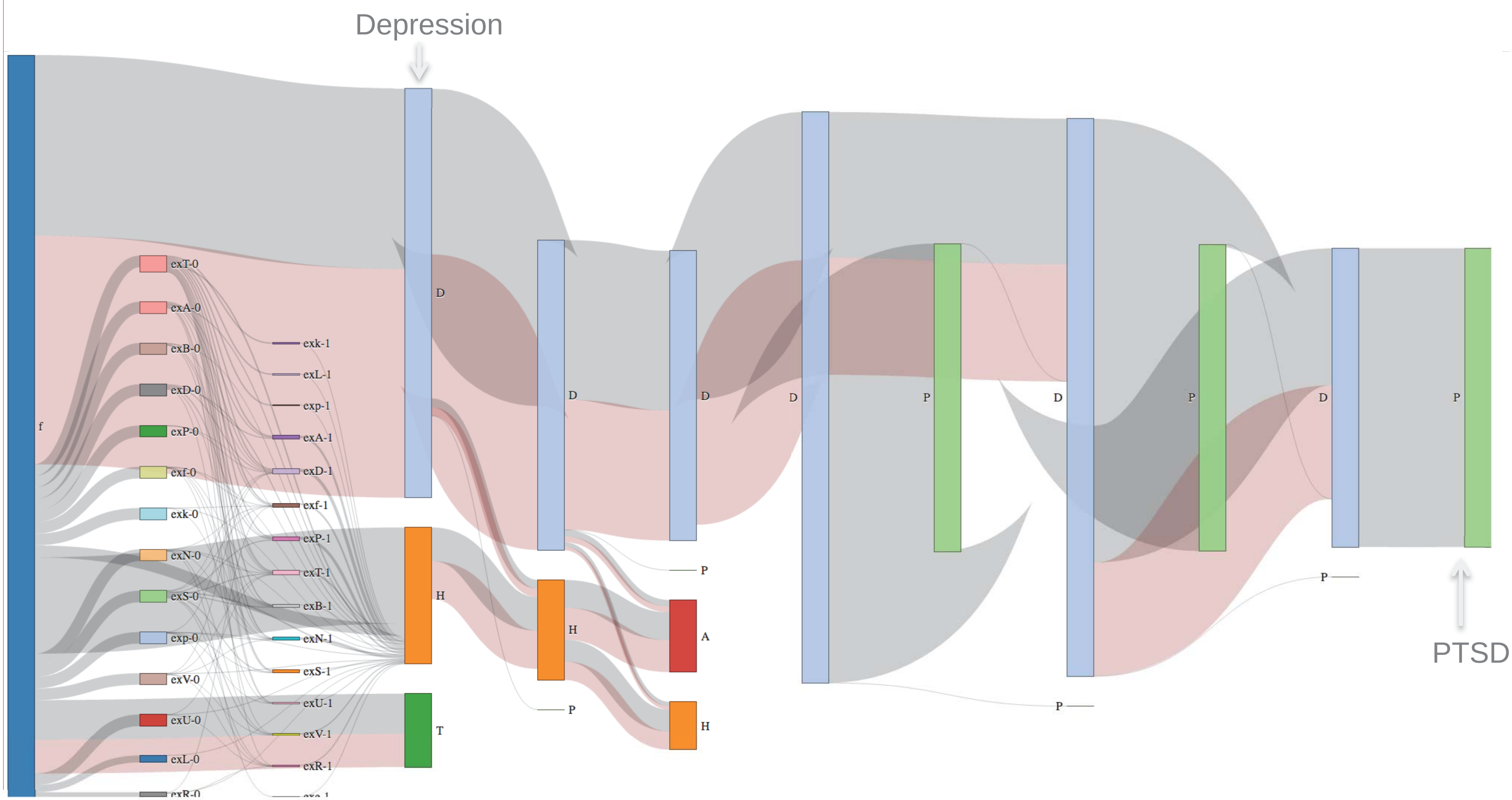
Visually Exploring EHR data for a Population or a Cohort of Patients

Clinical Trajectories



44

Exploring Longitudinal Data for a Cohort of Patients

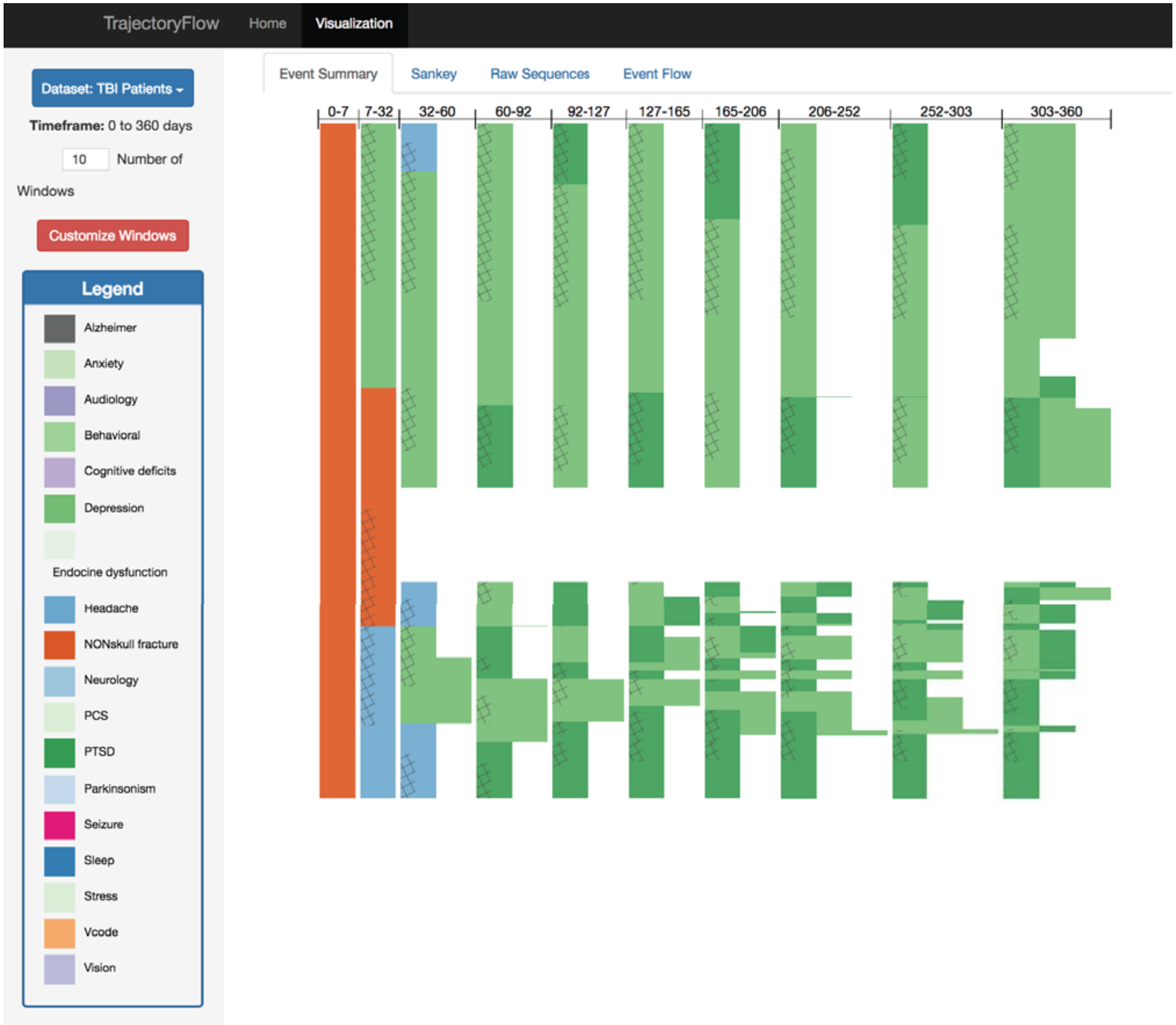


45

mTBI

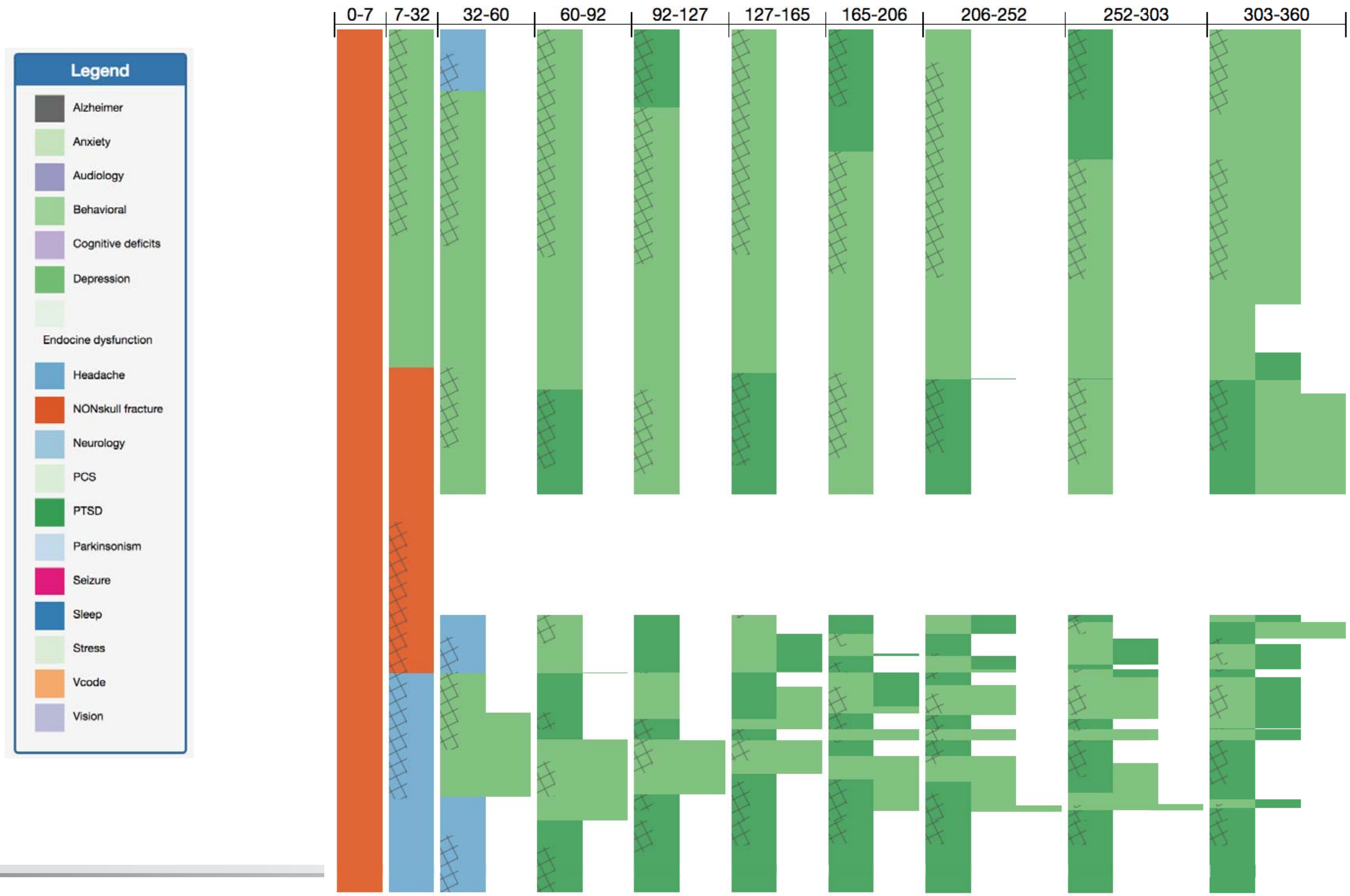
Tool to analyze clinical trajectories
From 1st TBI to 1st diagnosis of PTSD for 5,000 service members.

Hierarchical Exploration of Clinical Trajectories

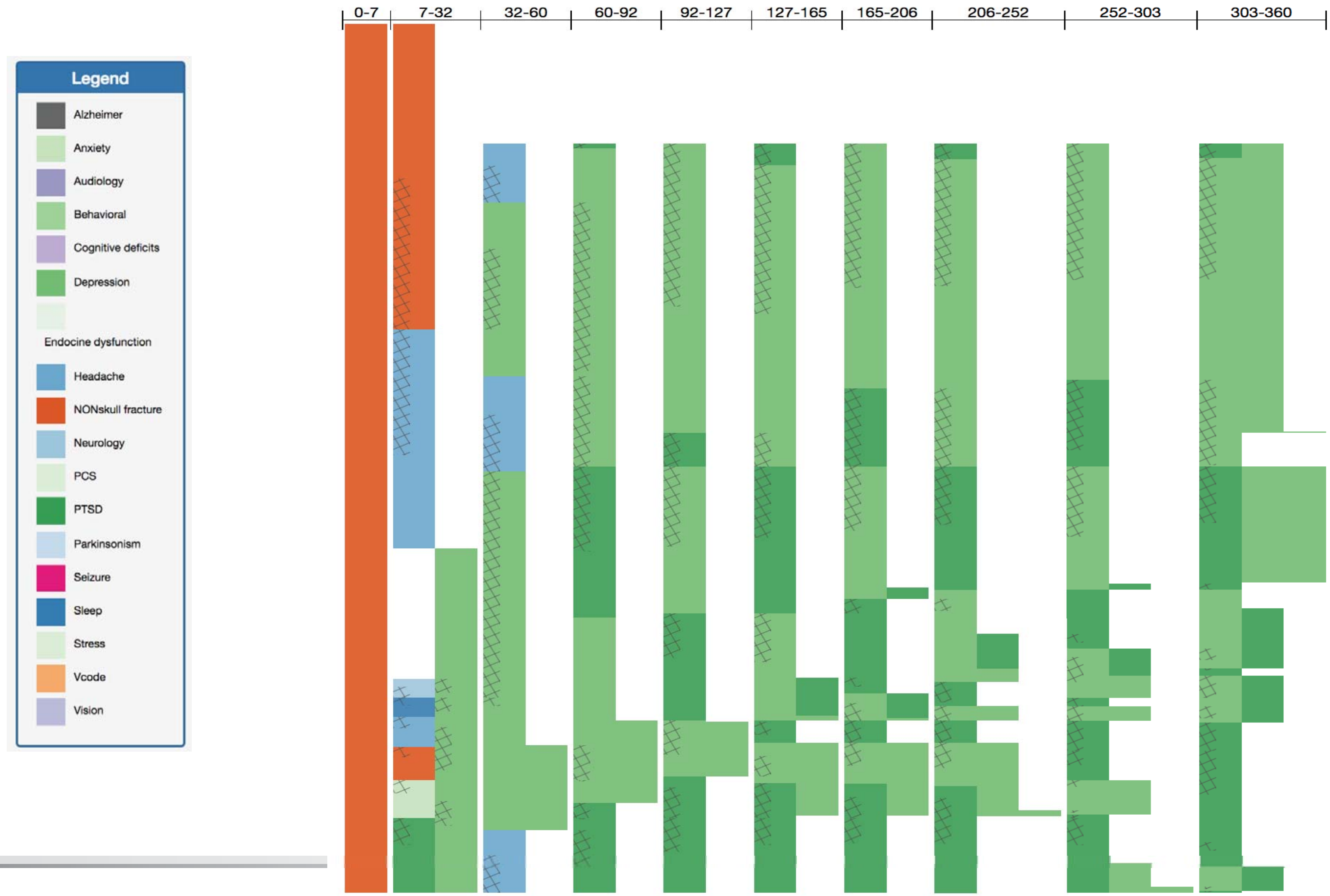


46

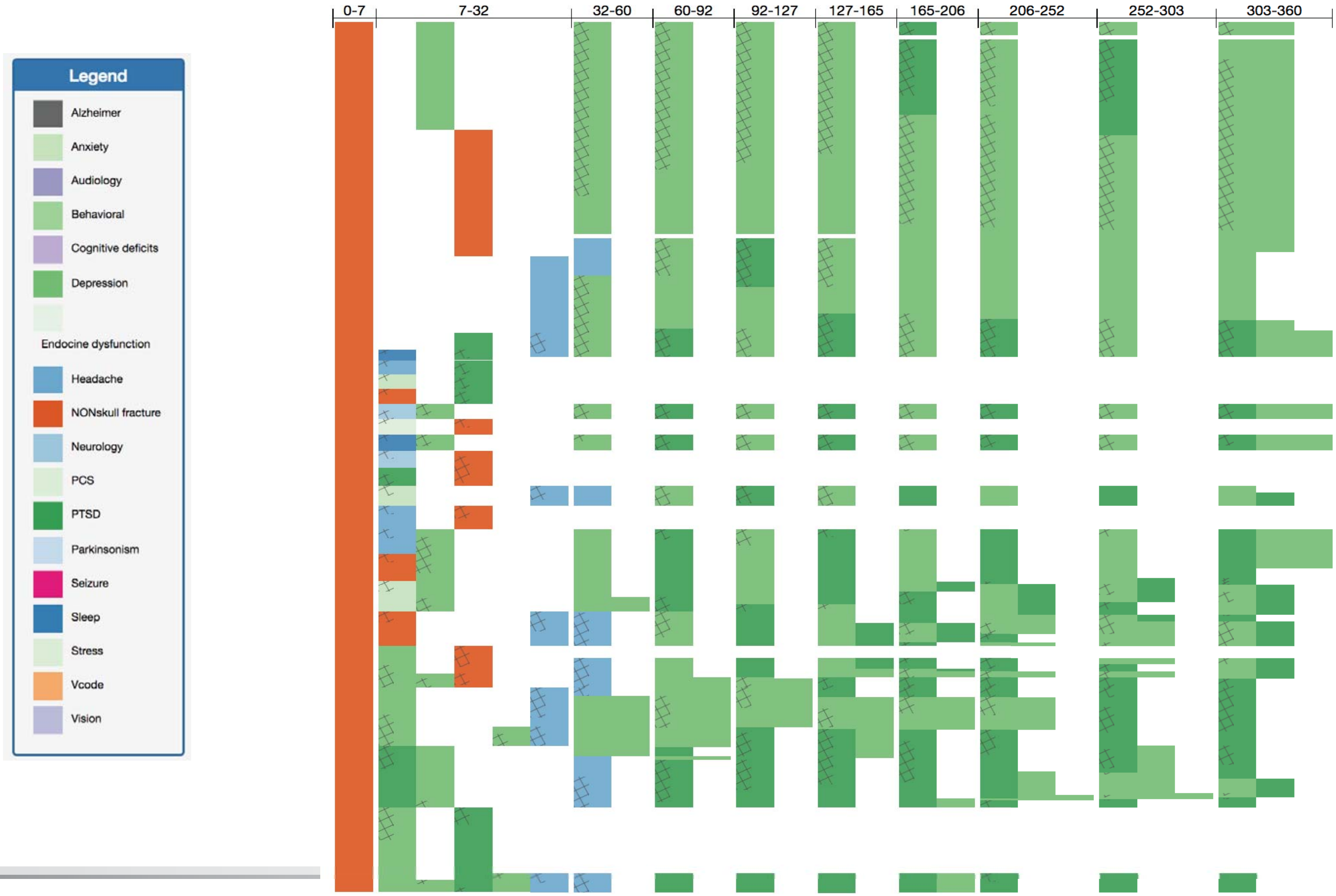
Hierarchical Exploration of Clinical Trajectories

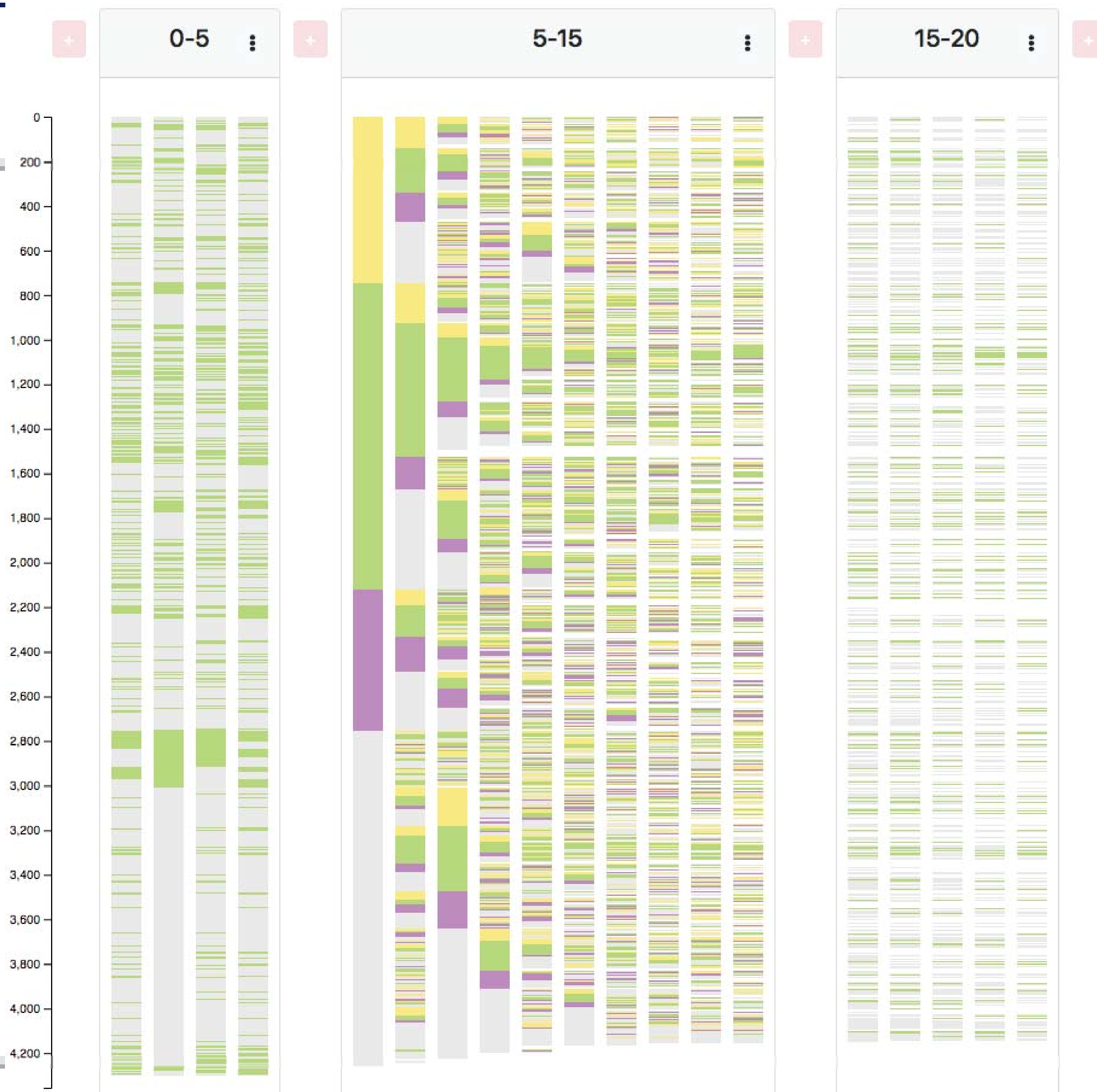


Hierarchical Exploration of Clinical Trajectories



Hierarchical Exploration of Clinical Trajectories




0-5
5-15
15-20
0
200
400
600
800
1,000
1,200
1,400
1,600
1,800
2,000
2,200
2,400
2,600
2,800
3,000
3,200
3,400
3,600
3,800
4,000
4,200

Barriers / Issues

- POP started on April 15th, 2015
  - Delayed on starting the project. It didn't start until August 2015
  - Challenges hiring qualified developer to focus on the project due to budget modifications

Risks & Risk Mitigation Plan

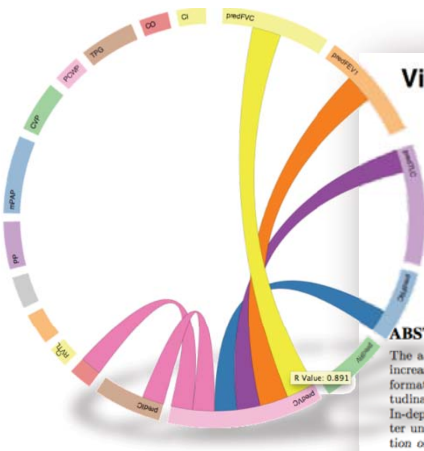
- **Risk:** Delays due to complexity of brining qualified candidates.
- **Mitigation:** Spent significant amount of time training candidate.

DHA Defense Health Agency

Metrics

PROTECT • PROJECT • SUSTAIN

- Number of research publications, presentations, and abstract
- Interest and the engagement shown by clinical providers



Visualization of Longitudinal Clinical Trajectories using a Graph-based Approach

Filip Dabek
NICoE
Walter Reed, Bethesda
filip.j.dabek.ctr@mail.mil

Jian Chen
Dept. of Computer Science and Electrical Engineering
University of Maryland, UMBC
jichen@umbc.edu

Alexander Garbarino
Dept. of Computer Science and Electrical Engineering
University of Maryland, UMBC
garba1@umbc.edu

Jesus J Caban
NICoE
Walter Reed, Bethesda
jesus.j.caban.civ@mail.mil

ABSTRACT

The adoption of electronic health records (EHRs) and the increased participation of hospitals and clinics in health information exchange systems have resulted in unique longitudinal data that describes a patient's clinical trajectory. In-depth analysis of that information is important to better understand the general course to recovery or the evolution of a particular disease. Unfortunately, modelling and understanding the progression of a disease is still a challenging task given that most often patients take vastly different paths after being diagnosed with a specific disease or condition. The different trajectories patients follow, the individual temporal events that each patient goes through, the uncertainties associated with clinical diagnoses, and the irregular time intervals between clinical diagnoses present challenges to researchers trying to analyze the common trajectories of a set of N patients. This paper presents a graph-based visualization method to interactively analyze the longitudinal clinical trajectory of a group of patients. The system allows users to select a specific set of events or conditions, filter the data based on different thresholds, and compare different cohorts while using an interactive virtual space that expands as the user continues to analyze and explore the data. The system has been tested with a dataset of over 89,000 patients and 8.7 million clinical events.

Keywords

Data visualization, visual analytics, longitudinal events, health-care data

1. INTRODUCTION

The rapid advancements in medical sciences and the increased adoption of electronic health records (EHRs) have

VAHC '15 Chicago, Illinois USA
© 2015 ACM. ISBN 978-1-4503-3671-0.
DOI: 10.1145/1235

... associated to a particular cohort of patients, estimates the different events, and generates a single graph illustrating all of the possible transitions. The system then allows users to select a specific set of events or conditions, filter the data based on different thresholds, and compare different cohorts

Understanding the Effects of Co[...] *Big Data*

A Grammar-based Approac[...] Clinical Trajectory after a[...] Inju[...]

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Procedia Computer Science 00 (2015) 1–
Procedia Computer Science
www.elsevier.com/locate/procedia

Leveraging *Big Data* to Model the Likelihood of Developing Psychological Conditions After a mild TBI

Filip Dabek and Jesus J. Caban*
National Intrepid Center of Excellence (NICoE), Walter Reed National Military Medical Center
Bethesda, MD

Abstract

A concussion is an invisible and poorly understood mild traumatic brain injury (mTBI) that can alter the way the brain functions. Patients who have been screened positive with mTBI are at increased risk of depression, post-traumatic stress disorder (PTSD), headaches, sleep disorders, and other neurological and psychological problems. During the last few years there have been an increased interest in understanding the intersection and comorbidities between traumatic brain injuries and psychological conditions. The interest has been driven in part by the effects that psychological disorders can have in recovery time and provider's burden [1]. Early detection of psychological conditions such as PTSD following a concussion might improve the overall outcome of a patient and could potentially reduce the cost associated with intense interventions often required when conditions go untreated for a long time.

Statistical and predictive models that leverage large-scale clinical repositories and use pre-existing conditions to determine the probability of a patient developing psychological conditions following a concussion have not been widely studied. This paper presents an SVM-based model that has been trained with a longitudinal dataset of over 5.3 million clinical encounters of 89,840 service members that have sustained a concussion. The model has been tested and validated with over 16,045 patients that developed PTSD and it has shown an accuracy of over 85% at predicting the condition within the first year following the injury.

Keywords: Big Data, Machine Learning, Concussion, Informatics, mild Traumatic Brain Injury

1. Introduction

A concussion is a poorly understood mild traumatic brain injury (mTBI) that can alter the way the brain functions. During the last decade a significant amount of attention has been given to the acquisition of clinical data from patients suffering mTBI and psychological health (PH) problems after a concussion. The increased awareness has been in part driven by the Department of Defense (DoD), the National Football League (NFL), and many other government and private organizations that have been leading different efforts to raise awareness about the short- and long-term effects of concussions.

A traumatic brain injury (TBI) is defined and indicated by "Any period of loss of or a decreased level of consciousness, Any loss of memory for events immediately before or after the injury, Any alteration in

Anticipated Impact
as an Outcome of Research

- Visually exploring the effects of a particular treatment plan and graphically monitoring changes over time will have a significant impact on the way clinicians assess the sequelae of TBI.

- To guarantee the success of our software application, two clinicians will guide, review, and validate the visualization system so it can be easily used and adapted by other researchers.

- We believe that the design and development of the software will create a community within the NICoE, Walter Reed, DHA, Intrepid Spirits, Uniformed Services University (USU), and other organizations that will use state-of-the-art visual analytic tools to better understand and explore TBI research data.

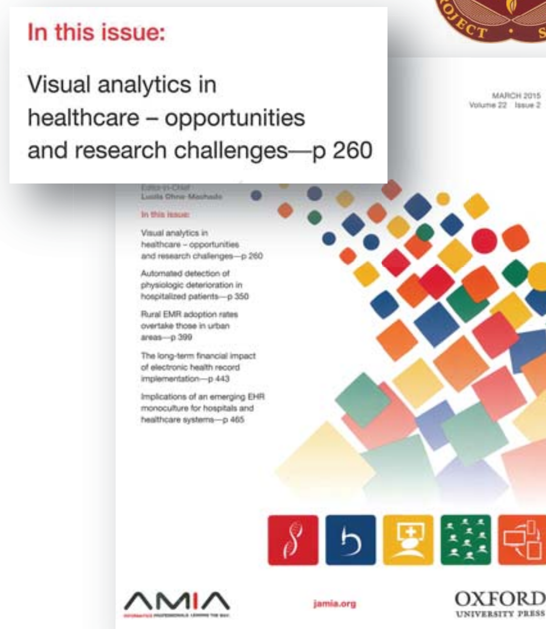
- Components of our visual analytical techniques being incorporated into CarePoint.

Transition Plan

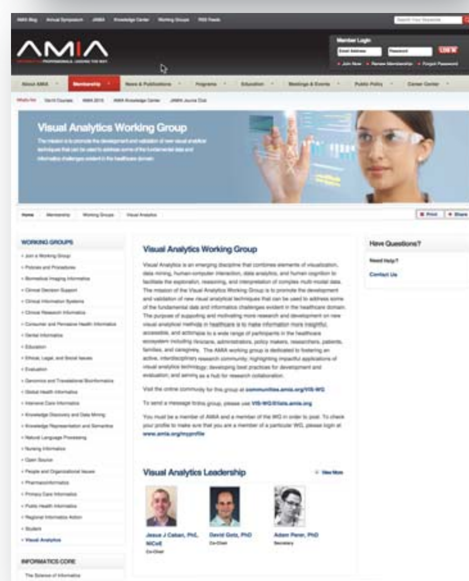
- Research Community: the software application will be primarily developed to target clinicians and researchers who need to fuse physiological, cognitive, behavioral, and imaging measurements. To accomplish that we plan to:
  - Include clinicians in regular meetings so they can guide the research and development of specific charts, correlations, and comparisons most clinicians are interested in analyzing.
  - Validate the framework with other clinicians within the NICoE.
  - Freely distribute the software application to members of the community.

Transition Plan

- 8 Workshops, 2 tutorials, 1 JAMIA special issue, 1 panel, 1 WG on Visual Analytics in Healthcare
  - 2010: IEEE Visualization Conference, Salt Lake City, Utah
  - 2011: IEEE Visualization Conference, Providence, RI
  - 2012: IEEE Visualization Conference, Seattle, WA
  - 2013: AMIA, Washington DC
  - 2014: AMIA, Washington DC
  - 2015, IEEE Visualization Conference, Chicago, IL
    - 2015: AMIA Tutorial: Introduction to Visual Analytics in Healthcare
  - 2016, AMIA Chicago, IL
    - 2016: IEEE Visualization Conference, Baltimore, MD
  - 2017, IEEE Visualization Conference, Phoenix AZ
    - 2017, AMIA Panel, Visualization in Healthcare, Washington DC



www.visualanalyticshealthcare.org/



AMIA Working Group

Conclusions

- Visualization systems can have a significant impact in how providers look at clinical data
- This project has the potential to impact different providers that need assistant with the data analysis and exploration
- Project on track

Contact Info:

Jesus J Caban, PhD
Chief, Clinical & Research Informatics
NICoE, Walter Reed Bethesda
E: jesus.j.caban.civ@mail.mil

Quad Chart

An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

PI: Dr. Jesus J. Caban **Org:** National Intrepid Center of Excellence (NICoE) / Geneva

DMRDP

Problem, Hypothesis and Military Relevance

- **Problem**: The large number of evaluation techniques to assess TBI patients poses challenges to clinicians who must integrate disparate measurements to understand the patient's condition.
- **Hypothesis**: In order to successfully analyze a large number of multi-modal clinical variables and integrate different evaluation protocols, new visual analytics techniques and applications are needed.
- **Military Relevance**: The computational models and visual analytical interfaces will use TBI data of military personnel.

Proposed Solution

- **Objective 1**: The design and development of an intuitive framework to visually analyze and explore a large number of clinical variables.
- **Objective 2:** Perform a usability study to validate application with clinical and research staff treating service members diagnosed with TBI/PTSD
- **Objective 3:** Establish a research community. The software application will be shared with researchers and clinicians at different DoD and federal organizations as a validated software application to visually integrate and analyze large number of clinical variables.
- **Budget Expenditure to Date:**
  - Projected Expenditure: $101K
  - Actual Expenditure: $303K **Updated:** 15 JAN 2018

Timeline and Cost

Activities	FY15	FY16	FY17	FY18
R&D visual analytics application				
Compare application with existing techniques				
Perform a usability study				
Validate software application with TBI Data				
Software Integration with other tools and libraries				
Share application				
Estimated Total Budget ($452k)	$45K	$103K	$113K	$94