Implementation of Personalized Longitudinal Health Records to Improve Patient and Clinicial Engagement

Nicole Brye, Aven Huang, Kenneth Nguyen, Rohith Pillai, Kamen Redfield, Anjana Sriram, Qiaoxuan Wang

March 15, 2023

Abstract

Major insufficiencies in healthcare systems today leave patients and clinicians alike wanting for a more effective way of receiving and giving care. Several major deficiencies are 1) the lack of an integrated Electronic Health Record (EHR), lifelog, and personal omics data, and 2) the lack of standardization across healthcare systems, data standards, and terminologies. This incompatibility creates inefficiencies in operating personalized medicine, leading to problems with interoperability and introducing ambiguity into the healthcare environment, especially amongst patients, their providers, and organizations. Our work focuses on developing a data-integrated patient timeline dashboard. While many specialty-specific and administrative clinical dashboards may already exist, there is not much –if any– literature available that describes general clinical dashboards that are oriented around patient care, can be used across various medical specialties, and explain how computations on the dashboard are conducted. Our work aims to consolidate the overwhelming amount of data and graphics in a patient's medical history into a single accessible platform. Key aspects of this dashboard include options for what performance indicators to see, variety in the data sources being used for visualization, and interactive capabilities for both patients and clinicians. To accomplish the creation of this dashboard, we worked in two sub-teams, with one focused on understanding the types of available data and standards to be followed, researching relevant UI/UX practices to be followed when working with medical data, and the other discovering potential software tools that can be used to store this data and create dashboards.

1 Introduction

The healthcare industry is one that produces large quantities of data coming from a wide variety of sources; however, the data is typically not accessible to or presented in a way that is useful for a patient or their healthcare provider. Patient data can come from an assortment of sources - prior physicians, wearable devices, apps, etc - which are typically not standardized with one another. This creates a lot of issues for clinicians. The lack of standardization discourages coherent patient health records and the inconsistency of data standards produces further problems with integration and interoperability of patients' electronic records. The data inconsistency at the clinician also reflects patients' difficulties with accessing their health records. The current healthcare infrastructure can often be a barrier for patients, leaving them unaware and uninformed of their medical data and history.

The goal of our work is to eliminate the ambiguity that currently exists among clinicians, patients, and healthcare systems. In order to accomplish this, we are attempting to integrate various sources of patient data into a single patient timeline dashboard. Apart from the specialty-specific dashboards, our purpose is to integrate a variety of data into a generalized stream of information with processed data adhered to the medical industry standards. This would be an individualized resource with consolidated patient data available to both an individual and their healthcare provider. The problems existing in the healthcare space that our dashboard will attempt to address are 1) medical dashboards not being designed with the patient in mind 2) the non-existence of timeline visualizations integrating EHR with personal omics' data 3) the inability of many existing dashboards to integrate data from different wearable device sources and 4) the lack of personal trauma data in medical histories (something that is often requested by patients)

2 Methods

Our dashboard is aimed at integrating data from a variety of sources - specifically, wearable devices and electronic health records. Therefore, the initial phase of our work focused on standardizing data from different sources. This was quite an extensive process, as it required thorough data exploration, identification of useful metrics, elimination of duplicate metrics, and standardization of units. We looked at data taken from two specific wearable devices - Oura rings and Apple watches - as well as data from patient electronic health records. We calculated important statistic functions - such as Pearson correlation - from the data so that they could also be integrated into the dashboard.

Once we had an initial set of usable data, the next step of the process was making the data into a format that was compatible with our visualization tool, Tableau. We selected Tableau as our primary visualization tool because of its ability to integrate smoothly into web pages and because of its useful interactivity features. This was an iterative process, in which the sub-team focused on data processing and worked with the visualization team to debug any data-cleaning or standardizing issues and ensure that the data format could be smoothly integrated into Tableau to create the visualizations we needed.

In order for the data to be more compatible with Tableau, we created a pipeline specifically designed to read in, clean, add features to, and subset the data. Our pipeline currently utilizes Oura ring csv data and exports the cleaned version of these files to one smaller csv that is easier for the front-end team to incorporate into Tableau visualizations. Our pipeline removes null values using a threshold, meaning that rows with null values are only removed if they contain entirely null values or more than 10 null values. Our pipeline also cleans timestamps for ease of integration with other datasets, typecasts columns, and subsets the data to a desired date range.

Using the data we had, we created and tested three different predictive models using the Oura Ring sleep data that predicts whether the data point comes from a day that was a weekend or weekday. Weekends can signify when patients were relaxing or performing less strenuous tasks than on a weekday, so looked to see if we could predict this based on their health data. The best performing model was the sklearn Nearest Neighbor classifer(N = 2) which had .77 accuracy on the training set and .85 accuracy on the test set. This prediction model shows that in the future we'll be able to predict a number of things using mobile health tracking devices. Cleaning and feature additions such as these allow us to produce more accurate, compelling visualizations. Overall, our data is saved at multiple stages in the pipeline, and the final, cleaned version is exported to a directory for the convenience of the front-end team.

3 Results

After gathering all of the data, processing it, visualizing it using Tableau, and publishing it to the website, our team established a product for the data-integrated dashboard presented by Tableau. The dashboard provides a comprehensive overview of a user's health status, consolidating multiple sources of health data into one convenient location.

At this development stage, the dashboard includes a variety of metrics, such as calories burned, total sleep hours, lowest heart rate, and Pearson Correlation graphs between different pairs of variables (minutes awake, total calories burned, steps, total sleep hours, lowest heart rate) all displayed on a per day scale. This allows users to easily monitor their daily health status and track changes over time. Additionally, the dashboard features user-selected years and months for dashboard display, which further enhances the ability to track progress.

However, we did encounter some obstacles in regard to incorporating EHRs into the dashboards. This posed a challenge due to HIPPA regulations and the need to protect sensitive patient information. While we were unable to include this information in the public-facing dashboard, we recognize the potential benefits of incorporating it in future iterations of the project.

Furthermore, we also faced difficulties with incorporating direct user input of significant life events into the annotation feature of Tableau. While this feature could potentially add valuable insights to the dashboard, it would require further development and testing to ensure its effectiveness.

4 Conclusion

The end result of our work was a single, integrated dashboard that provided patients with a way to easily view, filter, and engage with their health data. This dashboard is unique because it integrates data from multiple wearable technologies into a single platform; it synthesizes the information from various useful sources and presents it in a unified manner. The key strengths of our dashboard are the interactivity and the customization. Users - including both patients and providers - are able to tailor their viewing experience to best fit their needs, thanks to the filters and toggles on the dashboard.

There were several challenges and obstacles we encountered when developing this dashboard. For instance, there are many ethical considerations in involved with manipulating biometric data is very sensitive and confidential information. Therefore, because of privacy related concerns, we did not have much data to work with; we were limited to using the data of three patients. There is a lot of potential for expanding upon this dashboard. The immediate next step would be to incorporate data from EHRs (specifically, EPIC), which would allow us to incorporate major medical procedures and chronic conditions into patient timelines on the dashboard. Additionally, we would like to automatize the data uploading and processing pipeline, as that would allow us to scale the platform better to more patients. Finally, we are interested in getting more detailed feedback from users in order to determine what features are helpful and how we could possibly expand the functionality of the dashboard.

5 Appendix

We used pandas as the main tool to perform data cleaning and standardization for this project. In order to clean and standardize the data, we removed null values using a threshold strategy. Any row that contained 10 or more null values were removed, and the remaining rows were left in the DataFrame.

In addition, we converted dates so that they were pandas datetime objects, and so that they were easier to work with. In order to consolidate the data into one convenient location, we needed to merge everything based on the date of collection. This is a particularly interesting challenge in Oura, as each metric is recorded at different time points. However, for the purposes of this project, we limited our data to metrics that were recorded daily so that all of the metrics corresponded to one another.