Visual Design for the Visual Analytics for Progressive Patient Profiling with Applications in Cardiovascular Diseases

1. **Design Goals**

With the growing population and increasing demand for healthcare, it is a challenge for higher precision disease diagnosis. On the other hand, the data accumulation of the EMR (electronic medical records), LIS (Laboratory Information Systems), etc. provides a chance for visual analytics for progressive patient profiling. According to the current situation, we propose three tasks:

**T1) P3 with outcome visualization:** Retrieve similar historical CVD patients according to progressive information of query patients with progression of information in hospital, and summarize the outcomes of the retrieved CVD patients in a visual and intuitive way.

**T2) P3 with intelligent alerts of changed outcomes:** Provide intelligent alerts in a visual and interactive way as progression of query information leads to significantly changes of outcomes of retrieved similar patients.

**T3) Data quality integration and visualization:** Integrate data quality indicators into P3 visualization.

So due to the tasks, we design a prototype of visual analytics system.

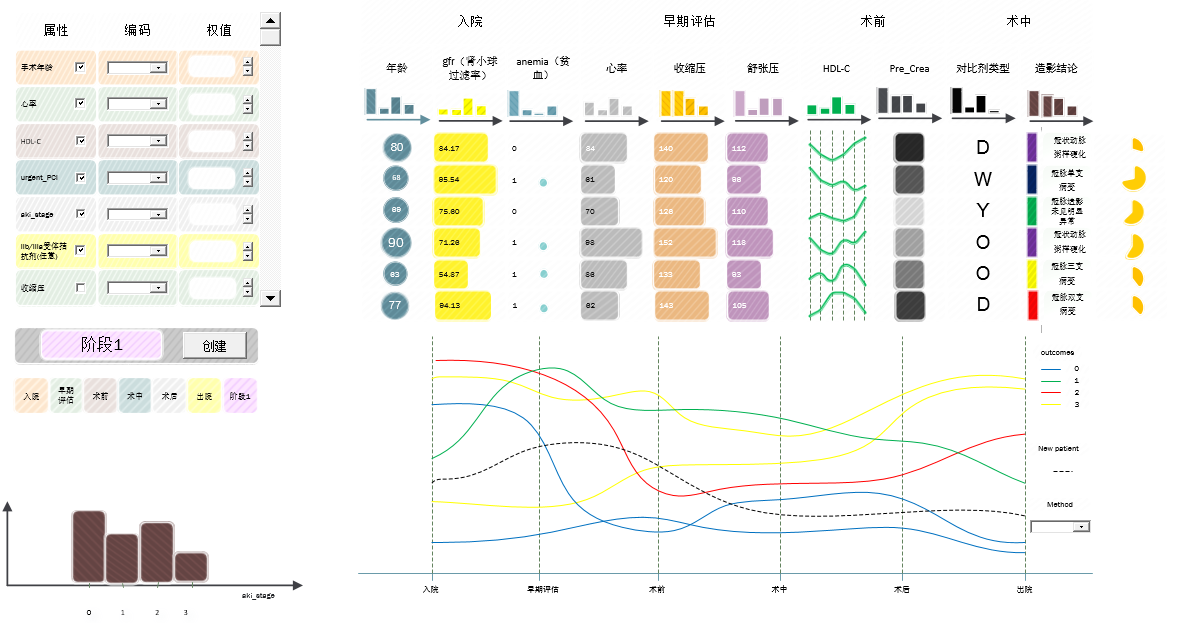


Fig 1. The design overview of the whole system interface.

1. **Prototype of visual design**

**2.1 Attributes view**

Here we provide a chart for attributes selection and process definition.

In this view, we aim to provide three main functions here.

1. **Demonstrate the stages and attributes.**

According to the data, the attributes belong to the different stages, like pre-operation, intra-operative and post-operation. We use a table to show what kinds of attributes each stage has.

Each attribute has different sources of heterogeneity, like numerical, categorical and time-series. We use different visual encoding method to display different types of data.

1. **Customized stages**

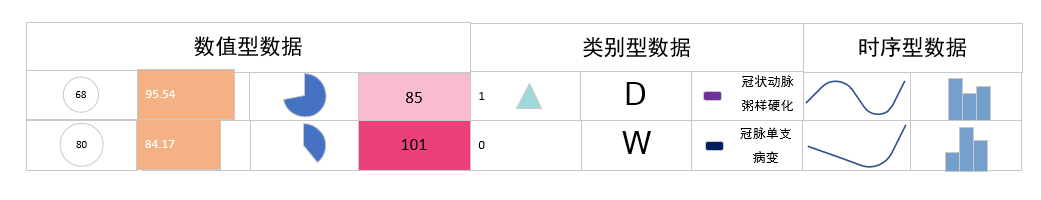
In the data, we can find the attributes has been classified into stages. However, in practical situation, the doctor or the user may focus on certain attributes. Therefore, we provide user-defined stages. Users can use checkbox to select attributes of interest and create a new stage.

1. **Weight adjustment**

We rank the similar patient in the patient view according to the similar patient computing method. Users may pay more attention to certain attributes rather than all attributes. We provide a bar to adjust the attribute’s weight in the similar patient computing method. Based on the customized weight, the rank in the patient view will change immediately.



Fig 2. This view shows the overview of the data attributes. There are three columns, which shows the attribute name, encoding method and weight in the similarity-computing algorithm.

  
Fig 3. The visual encoding of each kind of data. There three columns represents each kind of data. The numerical data can encode with circle, bar and pie chart. The categorical data use some glyph to show the meaning. Therefore, as to the time-series data, we use a small chart to represent the data changing.

In the Fig3. We investigate some visual encoding method in different categories of data. In the numerical data, the radial of circle, the length of bar, the percentage of pie and the saturation of color encode the number. In categorical data, we use glyphs like triangles, letters and color to represent the classification. We use small charts to show the time-serious changes.

**2.2 Patient view**

In this view, we provide an overview of the patients.

1. **Attributes presentation**

Each patient is in a row of the table, and each column represents the attributes. Due to the various sources of heterogeneity of data, we provide the customized encoding of data. In the attributes view, users can choose visual encoding of every attributes and it will change in the patient view.

1. **Patients Ranking**

The up-to-bottom order in the patient view represents the similarity of the query patients. The similar patients computing method is chosen by the user in the attributes view. Moreover, the weight of attributes can adjust by the users in order to select the attributes of interest or ignore the attributes.

1. **Data quality**

For the T3, we integrate the data quality visualization in this view. The header of the table shows the statistical distribution of each attributes using histograms. For the patient data’s confidence, we consider using pie chart to show the proportion of missing values.

Here we show the kinds of visual encoding of attributes.

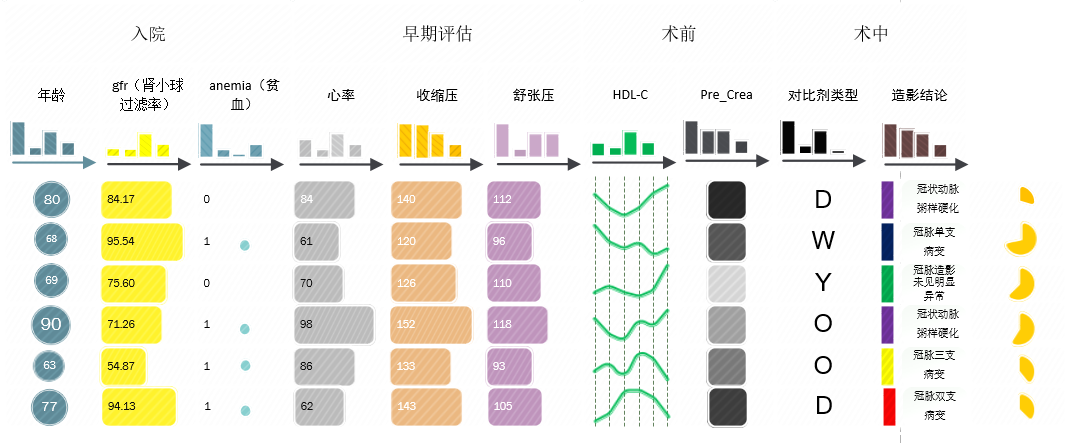


Fig 4. We use different visual encoding methods to show the attributes. Users can choose method through the attributes view.

**2.3 Outcome view**

We present the outcomes of the selected patients in the patients view and the whole data set (background) both in this view. This view shows the statistics information in different stages of the data set.

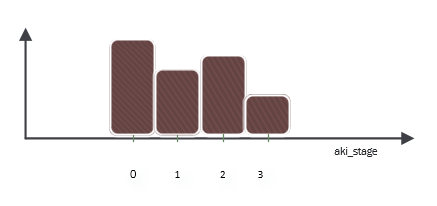


Fig 5. The statistic information of dataset. For example, the figure shows the outcome result of the whole dataset.

**2.4 Storyline view**

In this case, the horizontal axis encodes to the sequence of the acquired time-series information (like stages), and each line represents a retrieved similar patient, which converges and diverges with other patients’ lines, depicting their similarities.

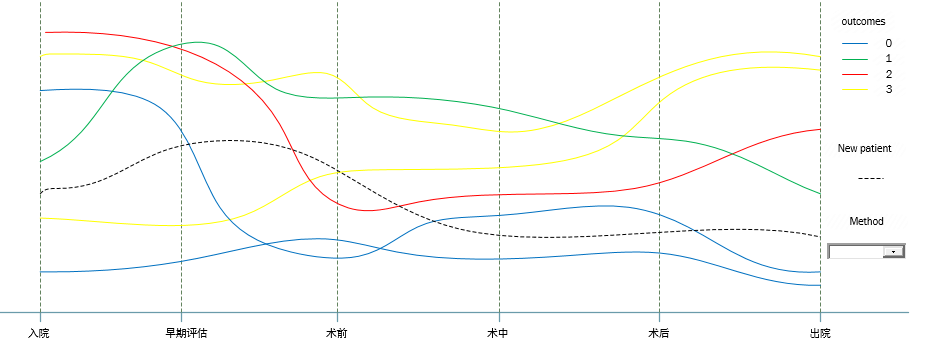


Fig 6. The similar patient storyline view. The color represents the outcomes and the distance between lines encodes the similarity of patients.