Tibco’s Spotfire software was used to analyze Prof. Vaughn-Cooke’s survey data on diabetes treatment [1]. The dataset consist of diabetes blood-glucose levels and monitoring non-adherence data from 100 patients over a span of 60 days. In addition to non-adherence data, there are patient features that could be correlated with non-adherence. For example, race, education, anxiety, insurance, among others. The data is provided in the form of a Microsoft Excel spreadsheet. Additionally, several features were computed and added to the dataset. These features included; average blood glucose level, standard deviation of blood glucose level, average adherence and standard deviation of adherence.

Inaccurate Provider Adherence Prediction

Can insurance providers accurately predict which patients will adhere to diabetes monitoring? In an effort to answer this question, insurance provider predicted adherence was plotted against a calculated actual adherence.

The resulting scatter plot shows that the provider predicted adherence dose not correlate well with actual adherence. Accurate prediction would result in points close to the line Y=X. Points are well off of this line. Supporting this argument, the correlation R-Squared correlation coefficient was calculated be a low, 0.26. Points were colored according insurance provider. One insurance provider does not appear to be better than another.

Patients with Low Adherence are Different

Can patients be grouped based on what unsafe acts they commit? Unsafe acts are the underlying reasons that patients do not adhere to prescribed treatment. If groups of patients can be identified, perhaps a similar correction can be applied to that group.
To answer this question, patients were clustered using a normalized frequency of 11 unsafe acts recorded in the study. To normalize, the number of unsafe acts was divided by the number of test required in the 60 day period. The graph was colored with red indicating a high normalized unsafe act score. The dendogram on the left of the graph shows that patients with many of high scores were less similar than patients with low unsafe acts scores. This may indicate that these patients would require individualized correction plans.

Additionally, unsafe acts were clustered based on patient unsafe act scores. The dendogram at the top of the visualization shows that unsafe act 2 and unsafe act 3 were most similar in terms of patient adherence. The same patients that ‘Don’t believe testing is important’ also ‘Don’t understand how/when to test’

Unsafe Acts Decrease with Diabetes Duration

How does adherence to treatment change as the amount of time since diagnosis? An informal argument could be made that the patient becomes more familiar with the tools and therefore adherence will increase the longer the patient has had the chronic disease.
11 histograms, one for each type of unsafe act, were constructed with diabetes duration on the x-axis and a normalized unsafe act score on the y-axis. All graphs show a general trend that as the diabetes duration increased the number of unsafe acts decreased.

**What factors are responsible for adherence motivation?**

Using HCE tool we study the factors that correlate with *adherence motivation* using the scores overview matrix. We list the factors that are negatively correlated in descending order of their absolute correlation coefficient.
1. Anxiety
2. Treatment Complexity
3. Number of Medications
4. Insulin frequency
5. Diabetes duration

A closer look reveals that all of these are negative factors and therefore negatively correlated with motivation i.e. the motivation to adherence of a patient is high due to these factors being low. On the other hand we now observe the factors that are positively correlated, again listed in descending order
1. Support – with good social support, the chances of adherence to motivation are better.
2. Insurance coverage – with better insurance coverage there is a motivation for adherence.

**Does medical condition of a patient correlate with adherence?**

We observe that the adherence of the patient is somewhat negatively correlated with the level of anxiety of the patient. Leaving out the outliers, we notice a decrease in the average of the overall
adherence over the entire time-period to decrease with the anxiety level of the patient. An additional feature that becomes evident from this scatter-plot is that females (in pink) suffer from higher levels of anxiety than males.

Tool Review

Tibco Spotfire software and HCE were critiqued. Tibco Spotfire was evaluated against Dr. Shneiderman’s 8-Rules of interface design [3] from the perspective of an analyst unfamiliar with the tool. Several pertinent rules were selected for mention:

Spotfire’s Positive Qualities:

1) Easy reversal of actions – Spotfire had a back button that would reverse previous actions one at a time. Additionally, there was a memory of instructions allowing un-do for all preceding actions.
2) Strive for consistency – Spotfire always had filters and details on demand in the same default location for all visualizations.
3) Enable frequent users to use shortcuts – Spotfire offered shortcuts that are similar to those used in other programs. For example; Ctrl+N was used for a new page and Ctrl+C was used copy.

Spotfire’s Areas of Improvement:

1) Design dialog to yield closure – This design principle states that actions should be grouped with a beginning, middle and an end. The interface allowed a wide variety of actions to be taken at each step. After making a graph, it was unclear what to do next. Perhaps there could be an optional window with a recommended next action. For example, a recommendation could be action X was frequently followed by action Y.
2) Offer simple error handling - When training a decision tree classifier an error was generated because the number of features that were being used exceeded a limit. The error was indicated by a small red exclamation mark on the bottom left of the screen and not in a prominent location. Additionally the error message was cluttered and contained excessive amounts of information. Error handing would be improved by making errors more visible and concise.

HCE Positive Comment:

1) It provides an easy interface for comparison of attributes based on different parameters such as correlation coefficient, least square error, quadracity, uniformity, number of outliers etc. which are extremely helpful and provides a quick insight of the data. It also does a color coding of the
attributes in accordance with the values of the parameters and thus helps us to quickly identify hidden relationships between large numbers of attributes.

HCE Negative Comment
1) It seems that HCE does not provide support for non-numeric data and therefore it requires the user to convert the non-numeric data into numeric data.

Michael Sorensen
University of Maryland
Department of Computer Science
mdsoren@umd.edu

Souvik Bhattacherjee
Department of Computer Science
University of Maryland
bsouvik@cs.umd.edu