The use of recursive partitioning via temporal visual toolsets to efficiently isolate patient cohorts

Gigi Lipori
pflugg@shands.ufl.edu
UF Health
Concept

• Total parenteral nutrition (TPN) was introduced in the 1960’s and has become vital in the prevention and reversal of malnutrition for individuals with various diseases and conditions
• TPN has been linked to mucosal atrophy, reduced GI hormone secretion and liver dysfunction
• To date, the majority of the information we have on parenteral nutrition-associated liver disease is based on data from the pediatric sector and animal models
• The frequency of parenteral nutrition in adults is on the rise, both for acute and chronic conditions
• Our aim is to retrospectively study adults who received TPN in our facility to determine the frequency of parenteral nutrition-associated liver disease, the face the disease, and the severity with which it presents
Approach

• We began by completing a short look-back to gain experience with a small dataset containing relatively few subjects and dimensions
• By using conventional query techniques against administrative and clinical datasets, we easily isolated a set of 4500 adults who received TPN
• We likewise easily isolated a subset of 2600 adults who received TPN and had some indication of liver disease
• Now for the hard part – what were the temporal considerations within this subset
  – How many patients had liver disease occur after TPN was introduced?
  – What was the progression of their disease?
Demonstration

• Let’s load synthetic data for 1K patients (20K facts) into EventFlow, explore the data, determine if we can answer our preliminary research questions, and if we can isolate a meaningful cohort for the research team

• We’ll use simplification techniques, filter techniques and transformation techniques, including category aggregation, interval merging and marker events
<table>
<thead>
<tr>
<th>Event</th>
<th>TPN</th>
<th>TPN gte 2x normal</th>
<th>Liver BX</th>
<th>Liver DX</th>
<th>TPN Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>649</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>656</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>657</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Back to Our Research Questions

• How many patients had liver disease occur after TPN was introduced? 180

• What was the progression of their disease?

- 0. first tpn
  Mean: 7 months 29 days 10:41 hrs
  Median: 2 months 21 days 1:00 hrs
  SD: 11 months 13 days 19:56 hrs

- 1. first liver bx after tpn

- 0. first tpn
  Mean: 9 months 1 days 19:11 hrs
  Median: 5 months 1:00 hrs
  SD: 9 months 23 days 19:35 hrs

- 1. first liver dx after tpn

- 0. first tpn
  Mean: 9 months 17 days 8:08 hrs
  Median: 5 months 21 days 1:00 hrs
  SD: 10 months 17 days 1:15 hrs

- 1. first alt gte 2x normal after tpn

• What else can we see?
Conclusion

• EventFlow provided a flexible and robust platform for us to recursively partition our data and explore various temporal attributes.

• Alternative traditional techniques would have required the creation of a priori hypotheses that would have been difficult to code and would have stifled the loose exploration of concepts.

• Through this exploratory exercise, we developed domain specific concepts that will be extensible to our larger study.
Demo Initial Load
Demo Search and Replace
Demo Combine Events
Demo Merge Events