Spotfire for Clinical Review: Design for Scale

Ramesh Rajamani

03 Nov 2016
Context

This presentation highlights key design areas for consideration for developing a Clinical Data Review solution.

- Clinical Data Review solution using Spotfire visualizations
- Solution that scales up to several hundred studies
- Cater for studies with large subject population (＞10,000 subjects)
- High end user performance
- Easy to develop and maintain
Request from Clinical Team: Study 1

Study 1 Data

Data Extract and Derivations

Derivations
- Compute change from baseline
- Normalize lab data
- Compute study day

Visualizations
- Adverse Events
- ECG and Vitals
- Labs – Change from Baseline
- Patient Profile
Request from Clinical Team: Study 2

Derivations
- Compute change from baseline
- Normalize lab data
- Address partial dates
- Compute study day (new)

Visualizations
- Adverse Events
- ECG and Vitals
- Labs – Change from Baseline
- Patient Profile
- AE – Lab Correlation (new)
Request from multiple Clinical Teams

Study 1
Data

Data Extract and Processing

Spotfire Reports

Study 2
Data

Data Extract and Processing

Spotfire Reports

Study 3
Data

Data Extract and Processing

Spotfire Reports

Study Data

Data Extract and Processing

Spotfire Reports
Constraints in this approach

- Multiple copies of code and reports for each study
- Resource intensive: Allocated to each study needs
  - Development, Testing, Validation and related support
- Challenges in scaling the solution to multiple studies
## Solution Considerations

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
</table>
| **Common solution approach for all studies** | - Common reporting design across studies  
- Handle variations in Study Protocol and Design  
- Address study specific reporting needs |
| **Derivations and Computations needed for visualization** | - Often derived information required for data review (study day, change from baseline, lab result normalization)  
- Such variables are computed much later in the downstream process  
- Solution needs to provide for these computed variables in visuals. |
| **End to end data refresh cycles**           | - Optimize data refresh time across all the layers (Data Source -> Warehouse / Mart -> Spotfire) |
| **End User Performance**                    | - Handle large data volumes for mega studies ( > 10,000 subjects)  
- Domains like lab can have > 5 Mil. Records  
- Reports need to load up to 30 data domains  
- Provide robust end user performance, while handle above loads |
| **Ability to scale up for few hundred studies** | - Solution that is robust to handle few hundred studies  
- Simple process to board a study and provision for clinical teams |
| **Testing and Validation**                  | - GxP Solution  
- Simplified testing and validation cycle |
## Design Areas

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High Level End to End Flow</td>
</tr>
<tr>
<td>2</td>
<td>Merging Study Data by Domain</td>
</tr>
<tr>
<td>3</td>
<td>Study level Configuration</td>
</tr>
<tr>
<td>4</td>
<td>Reusable Spotfire Template</td>
</tr>
<tr>
<td>5</td>
<td>Derivations – Database or Spotfire</td>
</tr>
<tr>
<td>6</td>
<td>Study Template Design in Spotfire</td>
</tr>
<tr>
<td>7</td>
<td>Spotfire Automation Services</td>
</tr>
<tr>
<td>8</td>
<td>Spotfire Scheduled Updates</td>
</tr>
</tbody>
</table>
Design Areas: High Level End to End Flow

1. High Level End to End Flow
2. Merging Study Data
3. Study level Configuration
4. Reusable Spotfire Template
5. Derivations – Database or Spotfire
6. Study Template Design in Spotfire
7. Spotfire Automation Services
8. Spotfire Scheduled Updates

Data Collection
- Clinical Data Management
- Clinical Trial Management
- Safety
- Third Party

Manage and Process Data
- Data Warehouse
  - Integrate data sources
  - Merge study data
  - Exclude blinded data

Clinical Reporting Mart Layer
- Data Mart
  - Study level configuration
  - Derivations and computations
  - Exclude columns not for reporting

Display Data
- Spotfire Reporting
  - Parameterized Global Template
  - Automation Services to embed data in report and publish to Study folder
  - Scheduled update for mega studies

Spotfire Clinical Data Review – Design for Scale
11/7/2016
Source system typically store data (AE, LAB, ECG..) separately for each study.

This creates challenges in applying a uniform approach to process and report on data.

Merging of study data by domain (AE, LAB, ECG..) addresses two key areas:

- Common approach to derivations / computations & Reusable Spotfire template.
Design Areas: Study Level Configuration

- Derived data is often needed to make the visualization more effective
  - Example: Study Day, Change from Baseline, Lab Result Normalization
- Study configuration file to address study specific variations and processing needs.
  - Example: Randomization date, Base line visit, Blinded domains definition.
- Study configuration is a one time setup during study onboarding
Design Areas: Reusable Spotfire Template

- Not practical to develop all reports each time a new study is on-boarded
- Create a common Spotfire template that can be reused across studies
- Most safety reviews can be achieved by this reusable template
- Study specific reports will be fewer and complement this global template
## Design Areas: Reusable Spotfire Template

- High Level End to End Flow
- Merging Study Data
- Study level Configuration
- Reusable Spotfire Template
- Derivations – Database or Spotfire
- Study Template Design in Spotfire
- Spotfire Automation Services
- Spotfire Scheduled Updates

### Example Template

**Clinical Study:** STUDY WXXK  
**Data Source:** Clinical Data Mart  
**Spotfire Instream Release:** 5  
**Last Refresh Date:** 02-Nov-2016  
**Recommended training:** WebPlayer Jumpstart Video  
**For FAQs, tips and tricks, general or study specific settings:** Sharepoint Documents  
**Please mail any questions regarding the Spotfire Instream Review application to:** instreamreview-spotfire@abc.com  

**ACTION LINKS:**  
- Maximize Zoom Bars  

**NOTE:** Only subjects for which Demography data have been entered will be available in the visualizations.

### DEMOGRAPHICS
- Demographics

### AVERSE EVENTS
- AE Summary  
- AE by Start Date  
- AE Treemap

### MEASUREMENTS
- Vital Signs  
- ECG  
- Lung Function Test (1)  
- Lung Function Test (2)

### LABS
- Lab Summary  
- Lab Change from Baseline  
- Lab Results by Subject  
- Post Baseline vs Baseline  
- Std ± DUSH Plot  
- Bivariate Analysis  
- Lab Subject Selection  
- Lab -> AE Profile  
- AE -> Lab Profile

### MEDICAL HISTORY
- Medical History  
- MS History

### CONCOMITANT MEDICATIONS
- Concomitant Medications  
- CM Subcategories

### PATIENT PROFILE
- Patient Profile  
- Patient Profile (Details)

### EFFICACY
- PASI Scores  
- Tender and Swollen Joint Count  
- GQA  
- ACES  
- Assessment_of_Rebound_&_IGA

### OTHERS
- Protocol Deviations  
- Disposition Events  
- Death
Design Areas: Derivations-Database or Spotfire?

- Derivations / transformations can be programmed in the database or in Spotfire (via calculated columns, insert columns etc.)

- Recommendation is to use one layer i.e. database layer for all computations and use Spotfire only for visualizations.

- Database provides more capability for complex computations.

- Computations done once in database – reused by different Spotfire visuals.

- Better performance as data is precomputed and ready for display.

- Testing and Validation process is simpler as all computations at one place.
Design Areas : Spotfire Template Design

- Template with parameterized information links (study as input parameter)
- Centralize creation of information model
- Use layout standards and guidelines for consistent look and feel
- Provide access to quick links on pages (Reset Filters, Clear Markings, Reset Zoom etc.)
- Design visualizations for max reusability. Provide dynamic user selection.
  - Select X and Y axis
  - Trellis By
  - Color By
  - Cross Tab definition

1. High Level End to End Flow
2. Merging Study Data
3. Study level Configuration
4. Reusable Spotfire Template
5. Derivations – Database or Spotfire
6. Study Template Design in Spotfire
7. Spotfire Automation Services
8. Spotfire Scheduled Updates

11/7/2016
Spotfire Clinical Data Review – Design for Scale
### Design Areas: Linked to Source or Embed Data?

#### When to select a particular approach?

<table>
<thead>
<tr>
<th>Linked to Source</th>
<th>Embed</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Need for real time reporting</td>
<td></td>
</tr>
<tr>
<td>• Data volumes are small.</td>
<td>• No need for real time reporting. Source data for reports refreshes periodically. (once daily)</td>
</tr>
<tr>
<td></td>
<td>• Data volumes are very large.</td>
</tr>
<tr>
<td></td>
<td>• Slow performance of reports, fetch from database is taking a long time to load for user</td>
</tr>
<tr>
<td></td>
<td><em>Automation Services are needed to refresh the data.</em></td>
</tr>
</tbody>
</table>
Automation services is used to embed data in report. Reduces report loading time.

Manual setup is needed for automation job + schedule for each new study.

This manual setup can be automated using the approach below.

1. **Which Studies to automate?**
   - Review Study Configuration
   - File to determine which studies are on-boarded and active

2. **Batch program create XMLs files for Automation**
   - Batch program generates the XMLs files for each study.
   - Batch program inputs
     - Spotfire Global Template
     - Study Code
     - File Output location

3. **Configure Batch Program Execution**
   - Execute automation jobs in parallel with a lag of (n seconds) after each study job trigger
   - Error logging / Email

4. **Trigger Batch program execution**
   - Batch program is triggered on notification of data mart refresh completion

---

Note: The table on the right side of the document seems to be an outline or a summary of the steps mentioned above, but it is not clear from the image how it relates to the text. The table contains placeholders for step numbers and descriptions, but the actual content is not visible.
Design Areas: Spotfire Scheduled Updates

- For mega studies, Spotfire loading time can be > 2 mins even after embedding data.
- Improve performance with Spotfire Scheduled Updates to preload study report in memory and reduce loading time.
- The manual setup of Schedules Updates can be automated using approach below.

1. Which Studies to load with Scheduled Updates?
   - Review Study Configuration File for mega studies
   - Review Spotfire audit logs for studies accessed in the last n(7) days.

2. Batch program generates Schedule Update.XML
   - Batch program generates the XMLs files using:
     - Study configuration
     - Audit logs
     - Existing manual schedules

3. Configure Batch Program Execution
   - Schedule update parameters (loading window)

4. Trigger Batch program execution
   - Batch program is triggered at a specific time to load the reports into memory.

The approach has to be revisited as schedules updates has undergone a change from version 7.5
Thank You