MAXIMIZING THE USE OF PATIENT DATA WAREHOUSE VIA VISUAL ANALYTICS

2019 PhUse US Connect

Sandy Lei, Simson Alex
PATIENT DATA WAREHOUSE

ENABLING EXPLORATION OF PATIENT DATA ACROSS JANSSEN TRIALS
Patient Data Warehouse – PDW

- A centralized repository for:
  - Study documents: Protocol | SAP | CSR
  - Clinical trial data: SAS datasets

- A technical capability for, originally for QS:
  - Data driven decision making
  - Search, subset, integrate, export, analyze

Timeline:
- Early-2013: Conceptualized (JEDI Funded)
- Late-2013: Study Browser v1
- 2015: Trial Catalog
- 2016: Study Browser v2
- 2017: AWS Migration
- 2018: Visual Analytics Tool

Stabilize

Implement

Optimize
PDW System Architecture

- Database
- Data/Metadata
- Application Server
- SAS Server
- Trial Catalog
- Study Browser
- Visual Analytics
- Various Web Applications

Data flow:
- Data flow from Database to Application Server
- SQL flow from Application Server to Database
- Response flow from SAS Server to Application Server
- Request flow from Application Server to SAS Server
PATIENT DATA WAREHOUSE CONTENTS

- 480 Trials With Patient Data
- 794 Reporting Efforts
- 1298 Folders
- 41775 SAS Datasets
- ~ 1.1 Billion SAS Data Rows

*All SAS datasets were loaded into database as is*
Data Divided Into Two Broad Sections

- Meta Data
- Is the knowledge about ~33 Million Records
- SAS Value
- ~1.1 Billion SAS Rows

Metadata is 3% of SAS data
Business Needs

• Traditional Data Access and Data Integration (per study data stored under file servers): For large scale cross-study integration/investigation, analysis and reporting, reliance on SAS programming is high and costly when data, programs and outputs are stored under study folders and subfolders.

• Patient Data Warehouse (centralized database and automated data access, data integration and analyses): The Patient Data Warehouse platform with clinical trials data and metadata stored in a centralized relational database, which afforded us with the ability to quickly access data across studies and made way for the use of modern techniques in data extraction, integration and analyses.
PDW Use Cases

1. Signal Detections (9 studies) - PDW POC for Global Medical Safety
2. Placebo Data Extraction (10 studies) - EMIF Metabolic: reuse of Placebo data from completed clinical trials
3. Placebo Response (10 studies) – identify possible factors which resulted in high placebo response rate
4. Hy’s Law by compound (17 studies in a compound)
5. Time-to-event for Major Adverse Events (10 studies)
6. Concomitant Medication & PK response (80 studies) - Impact of a concomitant medication co-administration on PK of approved Janssen mAbs
7. Explore opportunities to increase trial participation (252 studies)
8. Patient disposition rate by reasons of discontinuation (304 studies)
9. Efficacy Analysis by Biological Age (all PDW studies) – Janssen Prevention Center proposal for analysis based on biological age in clinical trials
PATIENT DATA WAREHOUSE
DATA SOURCE (SAS Data Folders)

SDTM: AE SAS Dataset

Browsed from SAS Universal Viewer
PDW Database – Core Concepts

- **Study Info Metadata** – SAS data location file per TA, compound, indication, phase, protocol number, reporting effort, data type (analysis, tabulation, raw), etc.

- **Study Data Metadata** *(created automatically during SAS to database loading)*
  1. Dataset
  2. Variable
  3. Variable Value
  4. Parameter Value

- **Study Data** – a normalized structure stores each data element (one cell in a SAS table) as one row in a patient value database table, no changes to the source data value

- PDW metadata are designed to facilitate the cross-study metadata review and data extraction.
PDW STUDY BROWSER

ENABLING AUTOMATED DATA INTEGRATION & SUBJECT SUBSETTING
PDW Study Browser (SB) – Metadata Review & Data Extraction

• PDW SB is designed for metadata review and data extraction:
  – Cross-study metadata review (any data element).
  – Automated data integration to stack datasets across studies.
  – Automated patient profile creation for subjects meet criteria.

• Users take the extracted data (output datasets/files) to perform analysis using other tools such as Spotfire, SAS, R, etc.

• Many earlier PDW use case support were utilizing PDW SB to output files and load files into Spotfire to create analysis.
Users’ decision for analysis & reporting tools to perform research/analysis based on extracted data from PDW Study Browser.
A visualization of a Hy’s Law subject
AST/ALT > 3xUL and TB > 2xUL in 30 days
PDW VISUAL ANALYTICS

ENABLING AUTOMATED DATA INTEGRATION & ANALYSIS
PDW Visual Analytics
Data Integration & Analysis Tool
(in development)

End-to-End platform by enabling automated analysis from metadata browsing to data extraction to final integrated analysis:

• Set up the connections from a visualization tool to PDW database
• Set up on-demand framework for data query from PDW database to facilitate analysis
• Design and create reusable analysis templates which will populate analysis results based on the data query created by the selected metadata.
Figure 1: 46 SDTM studies were selected from 3 compounds

Figure 2: Select interested datasets (AE, CM, DM)

Figure 3: Select variables to facilitate analysis

Figure 4: refreshed analysis results from an analysis template
Conclusion

• The new capabilities to search, query and perform analysis over hundreds of studies from PDW platform is a big achievement.

• Adherence to clinical data standards is key to address diverse questions accurately, confidently, and in a timely manner.

• As PDW data content increases, we are continuously improving our technological solutions to address performance issues.

• The PDW Visual Analytics platform technologies will shape the clinical trial processes to make analysis at “real-time” possible during any time point of the drug development cycle.
Q&A