Better Data, Better Tools, Better Decisions: Introduction to the Office of Computational Science

July 2018

Lilliam Rosario, Ph.D.
Director, Office of Computational Science
Center for Drug Evaluation and Research
Food and Drug Administration

Disclaimer

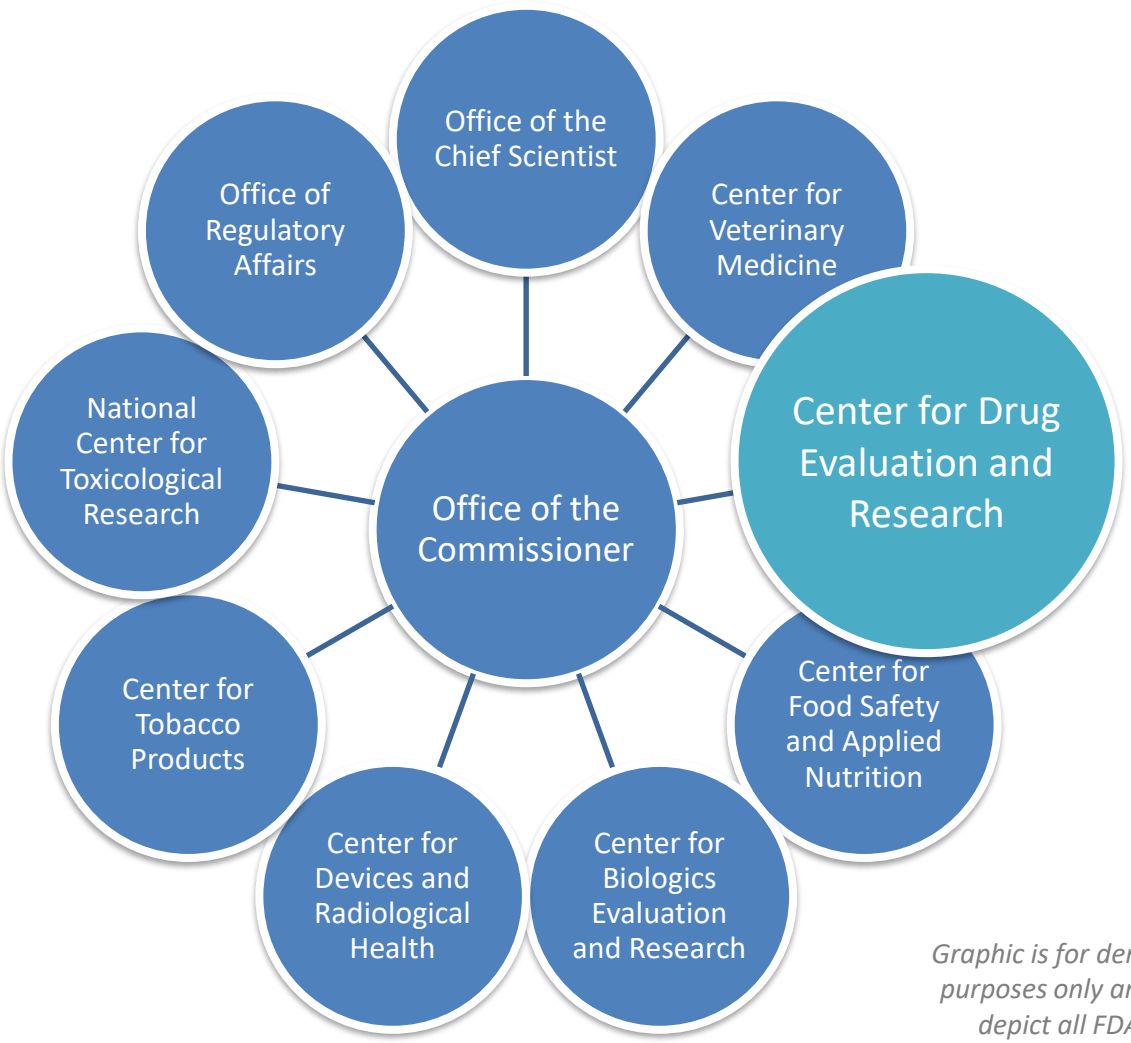
The views and opinions presented here represent those of the speaker and should not be considered to represent advice or guidance on behalf of the U.S. Food and Drug Administration.

Agenda

- Office of Computational Science (OCS) Overview
  - Where We Are
  - Who We Are
  - What We Do
- Current Research Projects

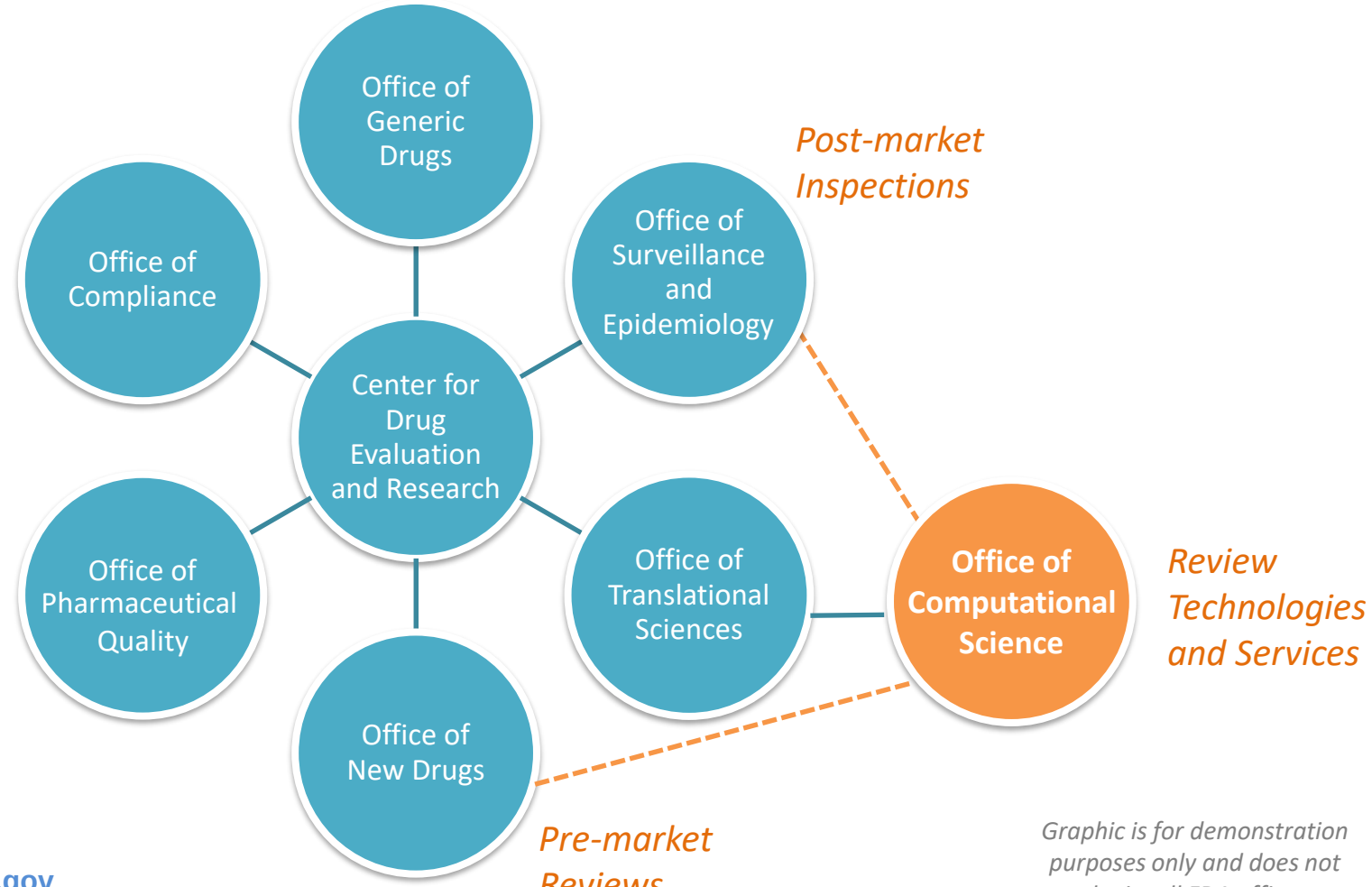
OVERVIEW OF THE OFFICE OF COMPUTATIONAL SCIENCE

Where We Are



Graphic is for demonstration purposes only and does not depict all FDA offices

Where We Are: CDER Review Offices



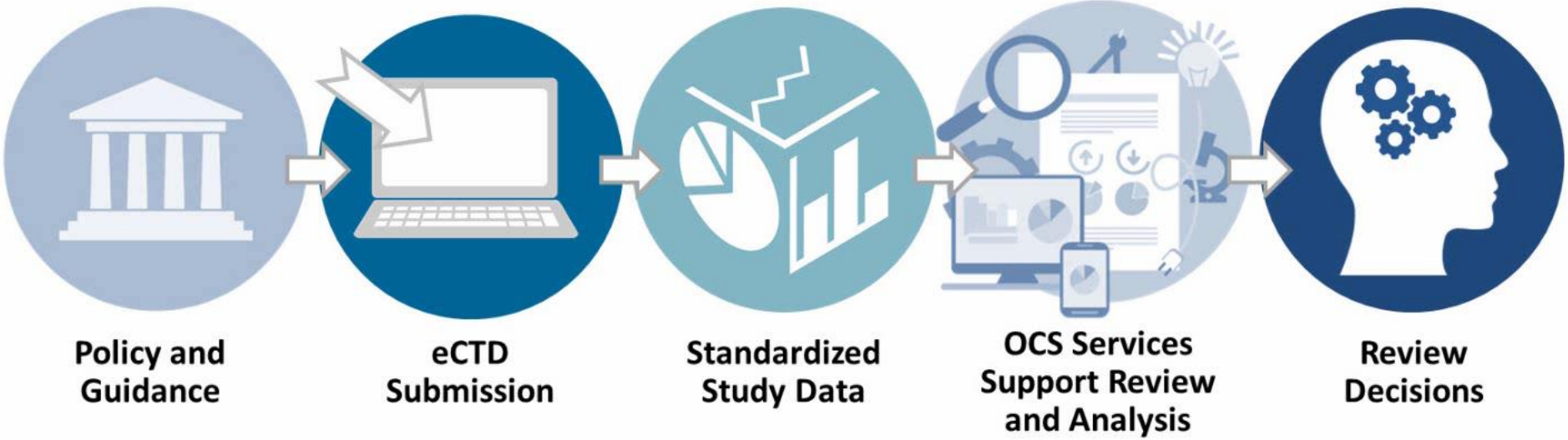
Graphic is for demonstration purposes only and does not depict all FDA offices

Who We Are


Our Vision
Drive the modernization of CDER's scientific review process through the implementation of tools, services, and training to enable reviewers to apply their expertise to information


Our Mission
To provide CDER reviewers innovative and reliable solutions that improve and strengthen the scientific review process by integrating data, tools, and training

From Policy to Review



What We Do

Safety Assessments and Signal Detection

Safety Analyses

Adverse Events Outputs

- AE MedDRA Comparison Analysis
- PT, HLT, HLGT, SOC, SMQ
- Toxicity Grade Summary
- Preferred Term Analysis by Toxicity Grade
- Two-term MedDRA Analysis
- AEs by Arm > 2%
- Serious AEs by Arm
- AEs by Severity
- Serious AEs by Severity
- Risk Assessment (AE and SMQ)
- Graphical Patient Profile

Vital Signs Outputs

Vitals Standard Analysis and Explorations

- Vital Signs results over time (Box and Whisker, Line Summaries, Baseline vs Min/Max)

Demographics Analysis

- Age
- Sex
- Race
- Ethnicity
- Country
- Site ID
- Disposition by Arm

Subject Disposition Analysis

- Disposition Event by Arm for All Subjects
- Disposition Event by Arm for Exposed Subjects

Laboratory Findings

Liver Lab Analysis Panel

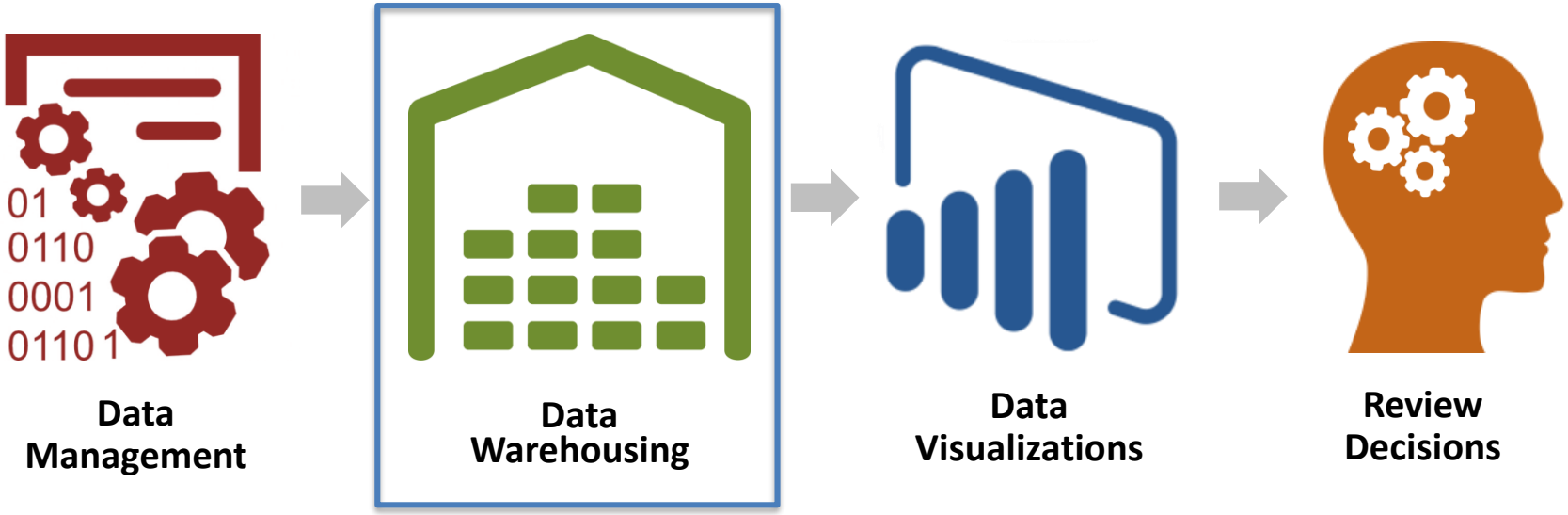
- Labs Greater Than Upper Limit Normal
- Possible Hy's Law Cases
- Max Lab Values Compared to Baseline
- Max AST and ALT vs. Max TB Lab Results per Subject
- Max Lab Results per Subject by Study Day

Standard Analyses of Explorations of Lab Data

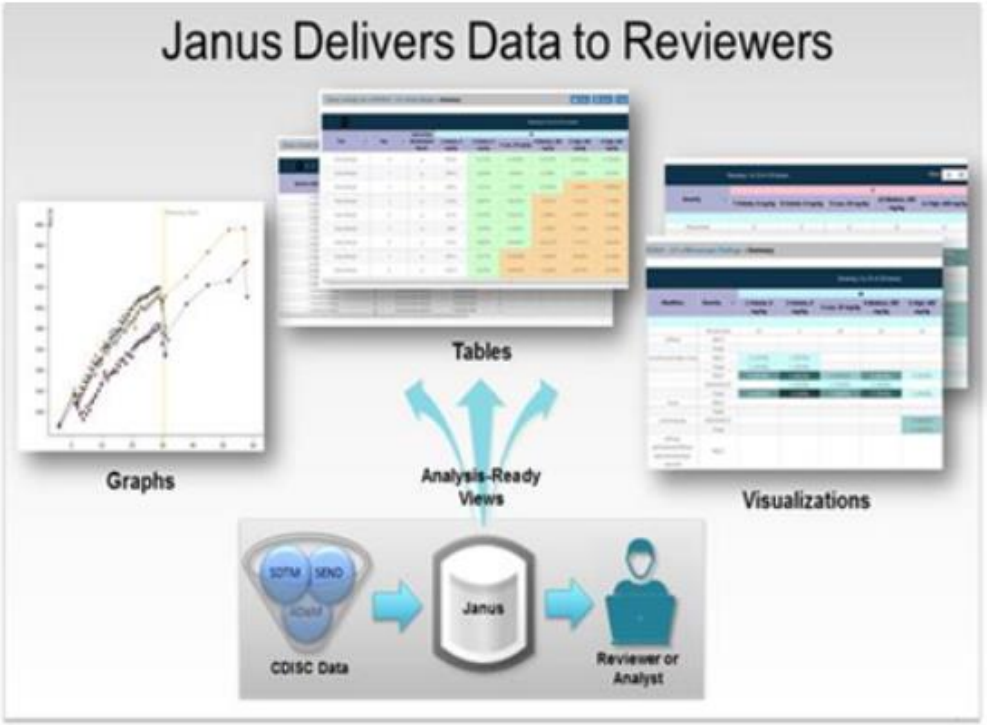
- Organ Class: Lab results over time (Box and Whisker, Line Summaries, Baseline vs Min/Max)

Special Assessments – Hy's Law

OCS Creates Services and Technologies to Support Regulatory Review Decisions



Janus Clinical



CURRENT RESEARCH AND COLLABORATIONS

Drug Induced Liver Injury (DILI) Research

Challenge:

Liver toxicity is the most common cause for the discontinuation of clinical trials on a drug and the most common reason for an approved drug's withdrawal from the marketplace.

Approach:

Create Liver Toxicity Knowledge Base (LTKB) to develop content-rich resources to improve our basic understanding of liver toxicity, for use by scientific researchers, the pharmaceutical industry, and regulatory bodies. The project involves the collection of diverse data (e.g., DILI mechanisms, drug metabolism, histopathology, therapeutic use, targets, side effects, etc.) associated with individual drugs and the use of systems biology analysis to integrate these data for DILI assessment and prediction.

Goal:

Develop novel biomarkers based on knowledge accumulated from the project.

https://www.fda.gov/ScienceResearch/BioinformaticsTools/LiverToxicityKnowledgeBase/ucm2024036.htm

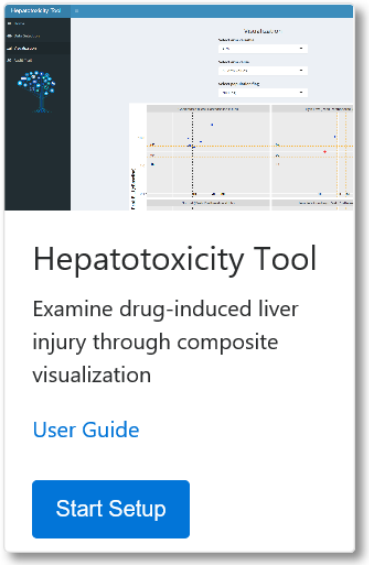
Drug Induced Liver Injury (DILI) Research

Challenge:

A rise in liver test values above normal limits predicts fatal DILI when accompanied by liver dysfunction (Hy's law). In subjects with liver disease, baseline pre treatment test values exceed normal limits. A rise in liver test values over baseline while on treatment can represent liver disease progression or DILI. No tools are available to identify DILI in these subjects.

Approach:

OCS ORISE research fellows compared the variability in liver test markers in clinical trials of healthy volunteers to patients with liver disease and developed a tool to visualize the change in liver tests from baseline to complement current DILI screening with Hy's Law analyses.



Results:

The Hepatotoxicity Tool complements Hy's law analysis with a visualization of the change over baseline test values and provides FDA reviewers a screening tool for DILI in treatment trials for liver disease.

Bereket Tesfaldet, et al. Variability in Baseline Liver Tests in Clinical Trials: Challenges in DILI Assessment *In*: Springer Protocols "Drug-Induced Liver Toxicity" Chen M, Will Y (eds) 2017.

Drug Induced Liver Injury (DILI) Research

Challenges:

- Defining DILI +/- is challenging – consider causality, incidence, and severity of liver injury events caused by each drug.
- Biomarkers and methodologies are being developed to assess hepatotoxicity but require a list of drugs with well-annotated DILI potential
- A drug classification scheme is essential to evaluate the performance of existing DILI biomarkers and discover novel DILI biomarkers but no adopted practice can classify a drug's DILI potential in humans.
- Drug labels used to develop a systematic and objective classification scheme[Rule-of-two (RO2)]. However highly context specific, rarity of DILI in the premarket experience, the complex phenotypes of DILI, drugs are often used in combination with other medications.

Approach:

- Integrate the post-marketing data into the drug-label based approach: the FDA FAERS database to improve the DILI classification.
- Develop a statistical prediction models for better predicting DILI: the structured & unstructured data (premarket and post market DILI narrative reports).

Results:

- Model Comparison and Improvement
- Visualization of results in interactive reporting tool
- Application to other adverse event scenarios

Assessing Cardiovascular Risk in Diabetes Trials

Challenges:

Cardiovascular (CV) safety in clinical trials relies on investigators' adverse event reports using standardized MedDRA queries (SMQ). To asses the CV safety of diabetes drugs in large CV outcome trials (CVOTs), FDA requires expert adjudication in addition to investigator SMQ reports. CVOTS provide a unique opportunity to compare SMQ report performance to expert adjudication.

Approach:

OCS and CDER reviewers compared the sensitivity and specificity of SMQ hazard ratio estimates with expert outcomes as the gold standard.

Results:

In adequately designed clinical trials, SMQ derived endpoints were concordant with expert adjudication. Narrow queries were more specific but less sensitive than broad queries.

Table 1. Adjudicated versus SMQ-derived MACE in CVOTs

	Adjudicated MACE	SMQ-derived MACE	
		Custom MACE	SMQ MACE
CVOTs	HR (95% CI)	HR (95% CI)	
EMPA-REG	0.86 (0.74, 0.99)	0.90 (0.77, 1.06)	0.87 (0.77, 0.98)
EXAMINE	0.96 (0.80, 1.16)	0.99 (0.83, 1.17)	0.98 (0.86, 1.11)
DEVOTE	0.91 (0.78, 1.06)	0.85 (0.72, 0.99)	0.85 (0.74, 0.98)
ELIXA	1.02 (0.89, 1.17)	1.02 (0.87, 1.19)	1.03 (0.89, 1.19)
LEADER	0.87 (0.78, 0.97)	0.82 (0.72, 0.93)	0.88 (0.79, 0.98)
SAVOR	1.00 (0.89, 1.12)	1.08 (0.96, 1.23)	1.04 (0.93, 1.15)
SUSTAIN-6	0.74 (0.58, 0.95)	0.72 (0.55, 0.93)	0.74 (0.59, 0.93)

Note: Cox model only included treatment arm as a covariate.
Study designs and patient populations differed among trials; therefore, the event rates across CVOTs should not be compared.
Abbreviations: CI, confidence interval; HR, hazard ratio; MACE, major adverse cardiac events; SMQ, standardized MedDRA query.

Patel T, Tesfaldet B, Chowdhury I, Kettermann A, Smith JP, Pucino F, Navarro Almario E
Endpoints in diabetes cardiovascular outcome trials. Lancet. 2018 Jun 16;391(10138):2412.

Assessing Cardiovascular Risk in Diabetes Trials

Challenge:

Application of innovative computational analytics to large datasets could uncover patterns of differential CV risk for patient subgroups or individuals. To improve public health outcomes, OCS partnered with the National Heart, Lung, and Blood Institute and academic investigators through the Meta-Analysis InterAgency Group (MATIG) to share resources and expertise in exploratory analyses of patient-level data from public access databases.

Approach:

Through MATIG, OCS applies systematic evidence-based approaches and machine learning techniques to identify prognostic factors for CV outcomes from patient-level data in publicly available CV therapy trials. OCS developed a research compendium, mapped data to a standard data model and used standard definitions to enable analysis of harmonized trial data.

Results:

Novel analysis tools applied to harmonized data uncovers new insight from existing publicly funded trial data, magnifying the returns on public investment in these trials. Data standards facilitate this reproducible, transparent research and fellowship participation in these activities fosters data science research careers.

Patel, T, et al. on behalf of MATIG. Pooled patient level data are better suited to investigate the link between dipeptidyl peptidase-4 inhibitors and the risk of heart failure in type 2 diabetes. *BMJ*. 2016 May 24; 353:i2920.

Describing Cardiovascular Injury

Challenge:

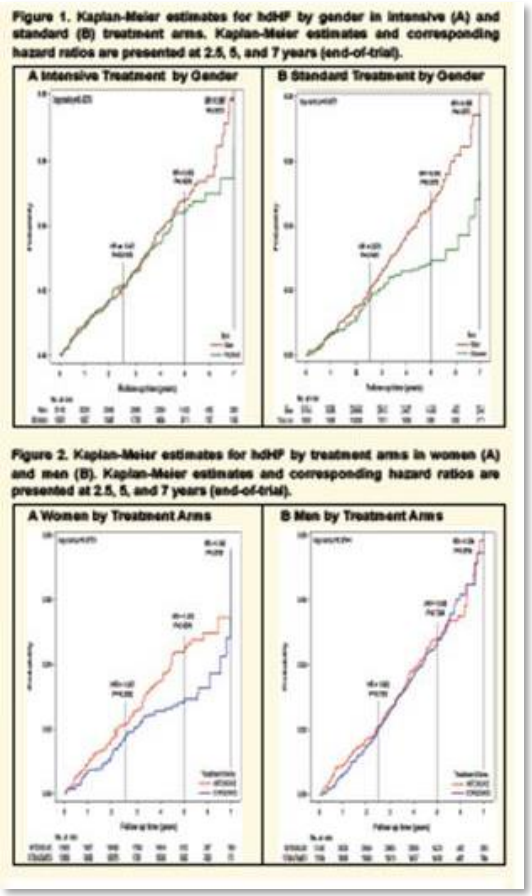
There is a need to assess the influence of sex (and the biological basis) on treatment outcomes.

Approach:

Reanalyzed publicly available data using a new analytic method to learn whether these findings need to influence the way diabetic female patients are treated.

Results:

- Women with type 2 diabetes tended to have an increased risk of hdHF events with intensive vs. standard glucose-lowering treatment in the ACCORD trial. No such difference was observed among men.
- This hypothesis-generating secondary analysis, without adjustments for multiple comparisons, warrants confirmatory studies.
- The findings call attention to the importance of outlining gender differences in treatment responses in clinical trials.



Patel T, Tesfaldet B, Navarro Almario E, et al. Risk of Hospitalization or Death due to Heart Failure with Intensive Glucose-Lowering Therapy in Diabetic Women. American College of Cardiology (ACC) 66th Annual Scientific Sessions & Expo. 2017 March 17; Washington, DC.

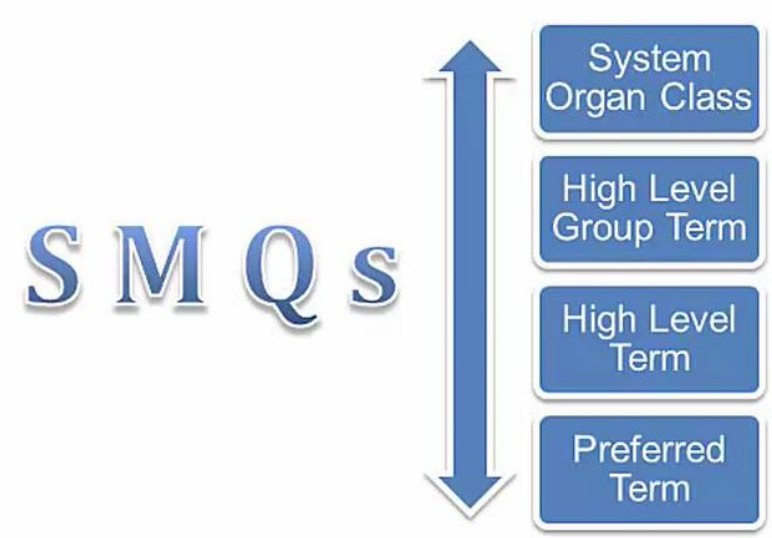
Adverse Event (AE) Signal Detection

Challenge:

- There is a need to identify AEs defined by analysis parameters and use risk difference, relative risk, and Standard MedDRA Queries.

Approach:

- Provide a quick and comprehensive look at the safety data by performing a series of exploratory adverse event analyses on data from clinical trials and non-denominator databases.



Results:

- Analyzes AEs at all levels of MedDRA hierarchy and Standard MedDRA Queries (SMQs)
- Provides AE counts at subject or event level by treatment group
- Compares AEs between study arms using a collection of risk estimators (odds ratio, risk difference, relative risk)
- Performs custom queries as user-defined preferred term groupings

AE Signal Detection: MAED

4	PT	Events	Number of subjects	Proportion (%)	Events	Number of subject	Proportion (%)	RD (per	RD C.I. (lower bound	RD C.I. (upper bound	RR	RR C.I. (lower bound	RR C.I. (upper bound
5	Dysgeusia	86	72	44.44	9	9	11.25	33.19	22.87	43.51	3.951	2.085	7.486
6	Anaemia	130	75	46.3	20	16	20	26.3	14.64	37.95	2.315	1.449	3.698
7	Dry skin	46	38	23.46	12	7	8.75	14.71		23.7	2.681	1.253	5.734
8	Decreased appetite	56	46	28.4	15	13	16.25	12.15			1.747	1.004	3.042
9	Irritability	42	37	22.84	13	10	12.5	10.34			827	0.958	3.483
10	Arthralgia	61	45	27.78	17	14	17.5	10.28			587	0.928	2.715
11	Rash	42	25	15.43	6	5	6.25	9.18			469	0.982	6.209
12	Cough	62	42	25.93	23	14	17.5	8.43			481	0.861	2.548
13	Diarrhoea	61	41	25.31	22	14	17.5	7.81			446	0.839	2.493
14	Dyspnoea	49	41	25.31	15	14	17.5	7.81			446	0.839	2.493
15	Vomiting	33	24	14.81	6	6	7.5	7.31			975	0.841	4.638
16	Dry mouth	32	26	16.05	8	7	8.75	7.3	-1.08	15.68	1.834	0.832	4.043
17	Thrombocytopenia	12	10	6.17	0	0	0	6.17	2.47	9.88	5.5	0.722	41.872
18	Abdominal pain upper	18	14	8.64	2	2	2.5	6.14	0.63	11.66	3.457	0.805	14.843
19	Dizziness	36	26	16.05	8	8	10	6.05	-2.62	14.72	1.605	0.761	383
20	Asthenia	59	38	23.46	22	14	17.5	5.96	-4.62	16.54	1.34	0.772	2.336
21	Leukopenia	18	11	6.79	3	1	1.25	5.54	0.96	10.12	432	0.714	41.1
22	Fatigue	119	92	56.79	64	41	51.25	5.54	-7.81	18.89	108	0.86	1.4
23	Pain	15	15	9.26	5	3	3.75	5.51	-0.59	11.61	2.469	73	8.28
24	Abdominal discomfort	8	8	4.94	0	0	0	4.94	1.6	8.27	4.5	58	34.9
25	Gastrooesophageal reflux disease	10	10	6.17	1	1	1.25	4.92	0.49	9.36	4.938	543	37.9
26	Urinary tract infection	16	10	6.17	1	1	1.25	4.92	0.49	9.36	1.938	0.64	37.9
27	Anxiety	23	20	12.35	10	6	7.5	4.85	-2.83	12.53	646	0.688	3.9
28	Nausea	100	68	41.98	37	30	37.5	4.48	-8.57	17.53	19	0.8	1.6
29	Respiratory tract congestion	10	9	5.56	1	1	1.25	4.31	0.02	8.59	4.444	573	34.474
30	Productive cough	12	11	6.79	2	2	2.5	4.29	-0.88	9.46	2.716	617	11.963
31	Constipation	23	19	11.73	6			4.23	-3.38	11.84	1.564	5	3.762
32	Neutropenia	43	23	14.2	12			4.2	-4.29	12.69	1.42	5	3.032
33	Affect lability	6	6	3.7	0			3.7	0.8	6.61	3.5		27.972
34	Anger	7	6	3.7	0			3.7	0.8	6.61	3.5		27.972

Introduction | SOC | HLGT | HLT | PT | Broad | Narrow

Indicates significant difference between treatment arms

MedDRA Hierarchy

SMQs

MAED
MedDRA-Based Adverse Event Diagnostics

MAED

Office of Computational Science
Better Data | Better Tools | Better Decisions

FDA

Natural Language Processing

Challenge:

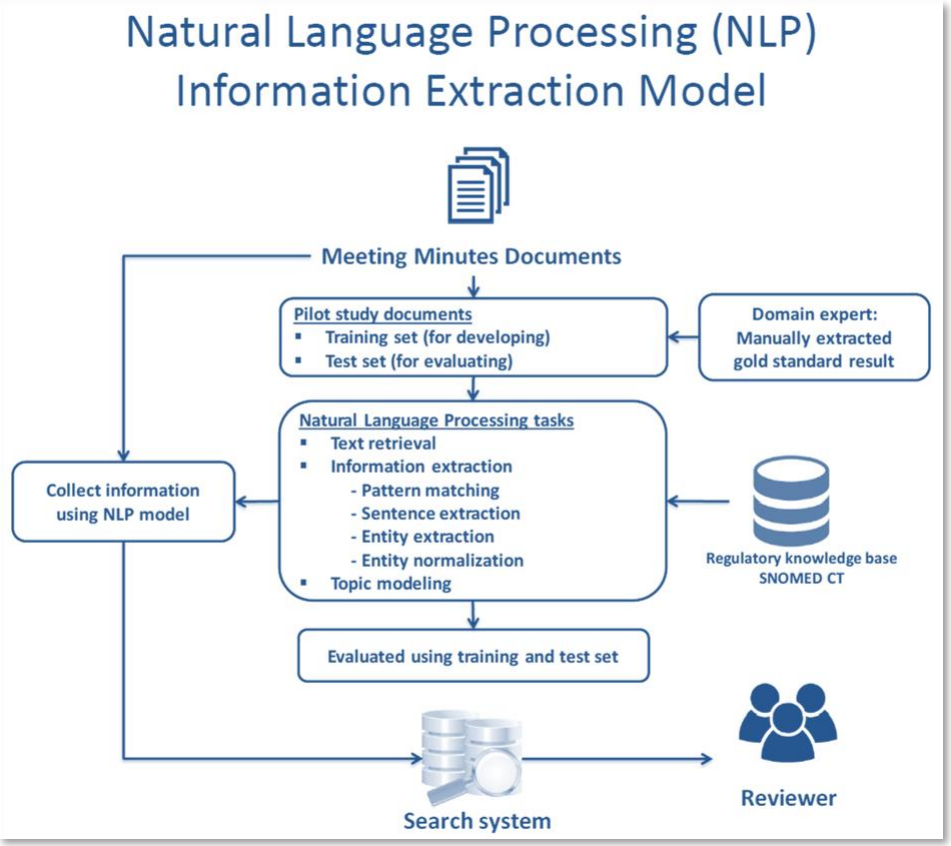
Unable to search and retrieve past meeting minutes for past regulatory decisions and other complex information.

Approach:

Use natural language processing (NLP) to extract semi-structured and unstructured information, which combined with established ontologies, will allow for document retrieval through improved search capabilities including hierarchical search

Results:

- Designed, developed, and evaluated an automated text mining tool that uses NLP to help with CDER's knowledge management efforts
- Extracted meeting minutes metadata with high precision and recall
- Developed proof of concept for extracting Q&A sections of meeting minutes using rule-based pattern matching



More information: https://www.lexjansen.com/phuse-us/2018/tt/TT07_ppt.pdf

Office of Computational Science
Better Data | Better Tools | Better Decisions

FDA

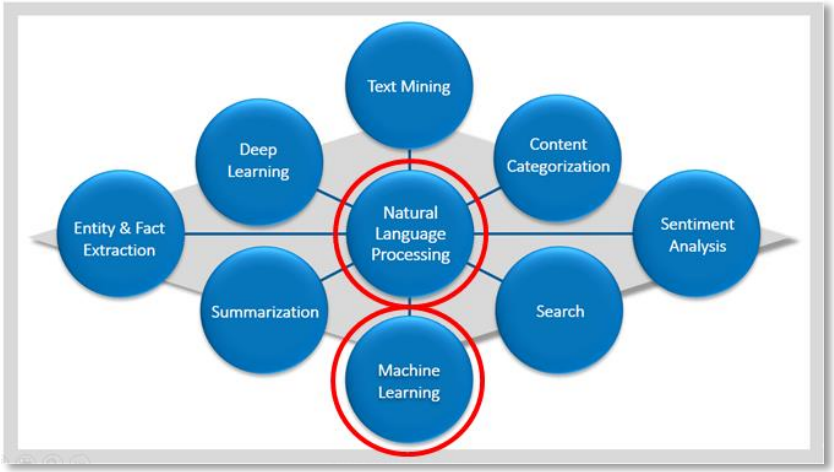
Machine Learning

Challenge:

- Determining conclusions from adverse event data in a timely manner is challenging for manual analysis alone

Approach:

- Utilize text analysis to employ a wide range of statistical, machine learning, and linguistic techniques for automated/semi-automated processing of text data
- Generate structured data from unstructured, then apply modeling techniques



Results:

- Uncovered **relationships** between these drugs and hepatic failure
- Offered information on **drug combinations** related to hepatic failure
- Clarified factors that can **predict** a greater degree of hepatic failure from serious events (death) versus less serious (treatable) events
- Assisted with **rating the impact** of using these drugs on patients

Office of Computational Science
Better Data | Better Tools | Better Decisions

FDA

Predicting Postmarket Adverse Events

Challenge

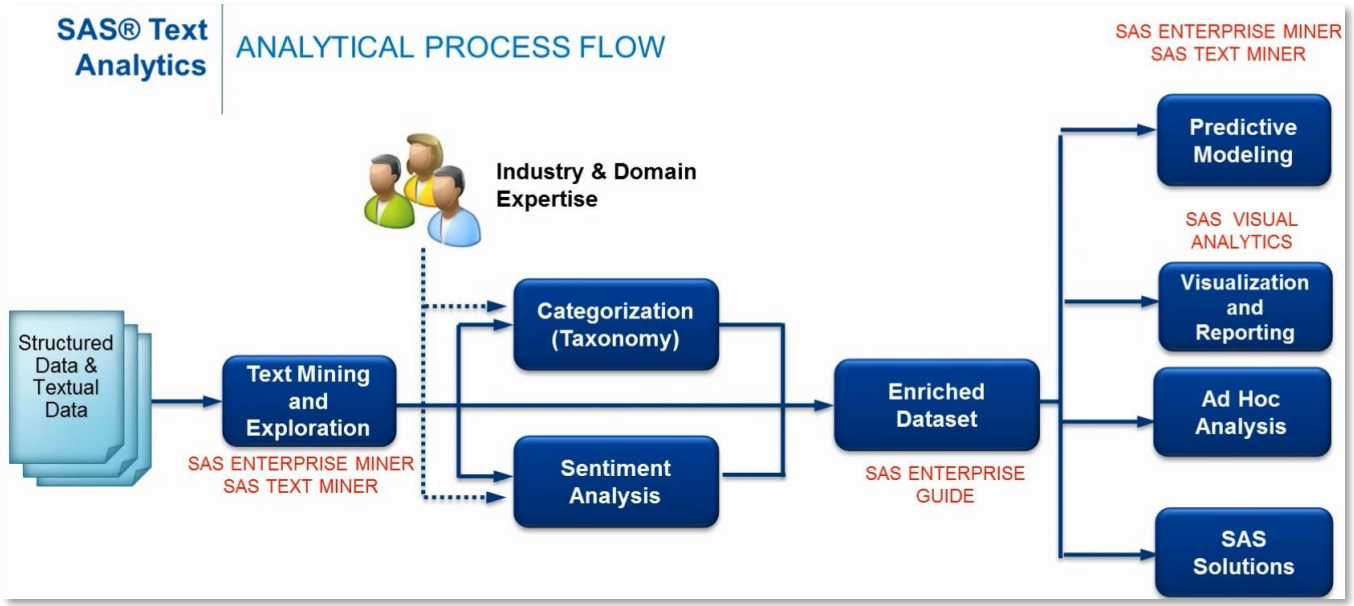
- Given volume of vaccine AE data, particularly unstructured text, it is **difficult to identify trends and serious reactions** by manually parsing through the data

Approach

- Combine SAS Enterprise Guide, Text Analytics and Visual Analytics to closely monitor the safety of vaccines and provide **analytics approach to discover AEs**
- **Model and predict serious AEs** to get sense about the primary characteristics of these events

Results

- Advanced understanding of **relationship** between these vaccines and their AEs
- Identified **common side effects**, including those that hamper course of treatment
- Identified **signs leading to AEs** through use of these vaccines
- Enhanced **ranking of degree of severity** of AEs
- Rated serious **impact** of using these vaccines



Current Collaborations: PhUSE

OCS collaborates with **Pharmaceutical Users Software Exchange (PhUSE)** in the Computational Science Symposium (CSS) and associated working groups

Current Collaborations: PhUSE

Standard Analyses and Code Sharing:
Improve the content and implementation of analyses for medical research, leading to *better data interpretations and increased efficiency* in the clinical drug development and review processes.

- Code Sharing (Repository)
- Communication, Promotion and Education
- Analysis and Display White papers
- Test Dataset Factory

Current Collaborations: PhUSE

Emerging Trends and Technologies: Share means of applying new technologies to create collaborative projects that will describe, prioritize, assess, and assist advancement of these opportunities.

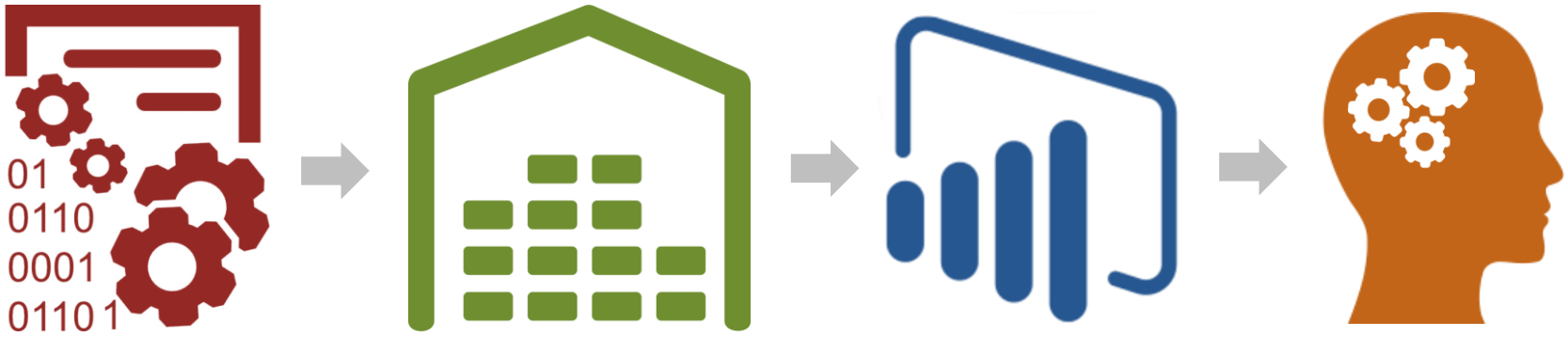
- Cloud Adoption in the Life Sciences Industry
- Data Visualizations for Clinical Data
- Investigating the use of FHIR in Clinical Research
- Clinical Trials Data as RDF
- Introduction to Clinical Development Design (CDD) Framework
- Blockchain Technology
- ODM4 Submissions
- Future Forum Interoperability & Technology
- Key Performance Indicators & Metrics (KPI)

Current Collaborations: PhUSE

Educating for the Future: Develop frameworks by which to educate the PhUSE community on technology advancements and how they can be used to drive innovation in the industry.

- Machine Learning / Artificial Intelligence
- Design Thinking
- Data Engineering

OCS Creates Services and Technologies to Support Regulatory Review Decisions




FDA
U.S. FOOD & DRUG
ADMINISTRATION