FDA Experience with the Sentinel Common Data Model: Addressing Data Sufficiency

Michael D. Nguyen, MD

Office of Surveillance and Epidemiology
Center for Drug Evaluation and Research
US Food and Drug Administration

Disclaimer

- The views expressed in this presentation do not reflect the officials views or policies of the FDA

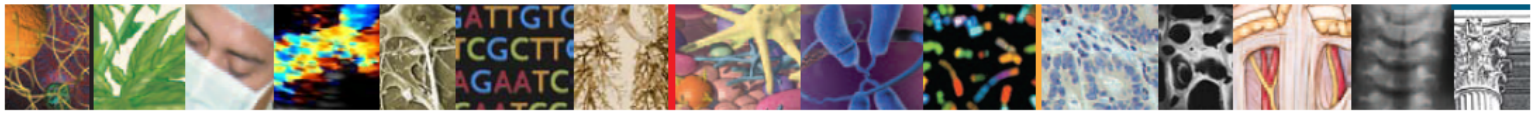
Greatest Strength, Greatest Weakness

"The benefit of the common data model is that applications can work against data without needing to explicitly know where that data is coming from."

https://docs.microsoft.com/en-us/common-data-service/entity-reference/common-data-model

Plan for Talk

- Foundational needs and goals for Sentinel
- Sentinel system is more than the Sentinel CDM
- What the CDM does for FDA and why it meets our needs
  - CDM as data manager and curator
  - CDM as unifier and buffer
  - CDM as enabler of analytic scale and customization
  - CDM as accelerator of public health response
- Example of FDA study design process
- Things we wish the CDM could do or fix
- Towards some guiding principles from an FDA perspective



The NEW ENGLAND JOURNAL of MEDICINE

FDA

Perspective

Developing the Sentinel System — A National Resource for Evidence Development

Rachel E. Behrman, M.D., M.P.H., Joshua S. Benner, Pharm.D., Sc.D., Jeffrey S. Brown, Ph.D., Mark McClellan, M.D., Ph.D., Janet Woodcock, M.D., and Richard Platt, M.D.

The Food and Drug Administration (FDA) now has the capacity to "query" the electronic health

convening an ongoing series of discussions among stakeholders to address the near- and long-term

Legislative Requirement to Consider Sufficiency of Sentinel (ARIA) before PMR

Section 905

Mandates creation of an Active Risk Identification and Analysis System

Section 901

New FDAAA PMR authority

"The Secretary may not require the responsible person to conduct a study under this paragraph, unless the Secretary makes a determination that the reports under subsection (k)(1) and the **active postmarket risk identification and analysis system** as available under subsection (k)(3) will not be **sufficient** to meet the purposes set forth in subparagraph (B)."

Defining ARIA Sufficiency

"When to use the Sentinel System for a particular question"

- **Adequate data**
  - Drug of interest and comparator
  - Health outcome of interest
  - Confounders and covariates
- **Appropriate methods**
- **To answer the question of interest**
  - assess a known serious risk related to the use of the drug
  - assess signals of serious risk related to the use of the drug
  - identify an unexpected serious risk when available data indicate the potential for a serious risk
- **To lead to a satisfactory level of precision**

Summary of Foundational Needs and Goals

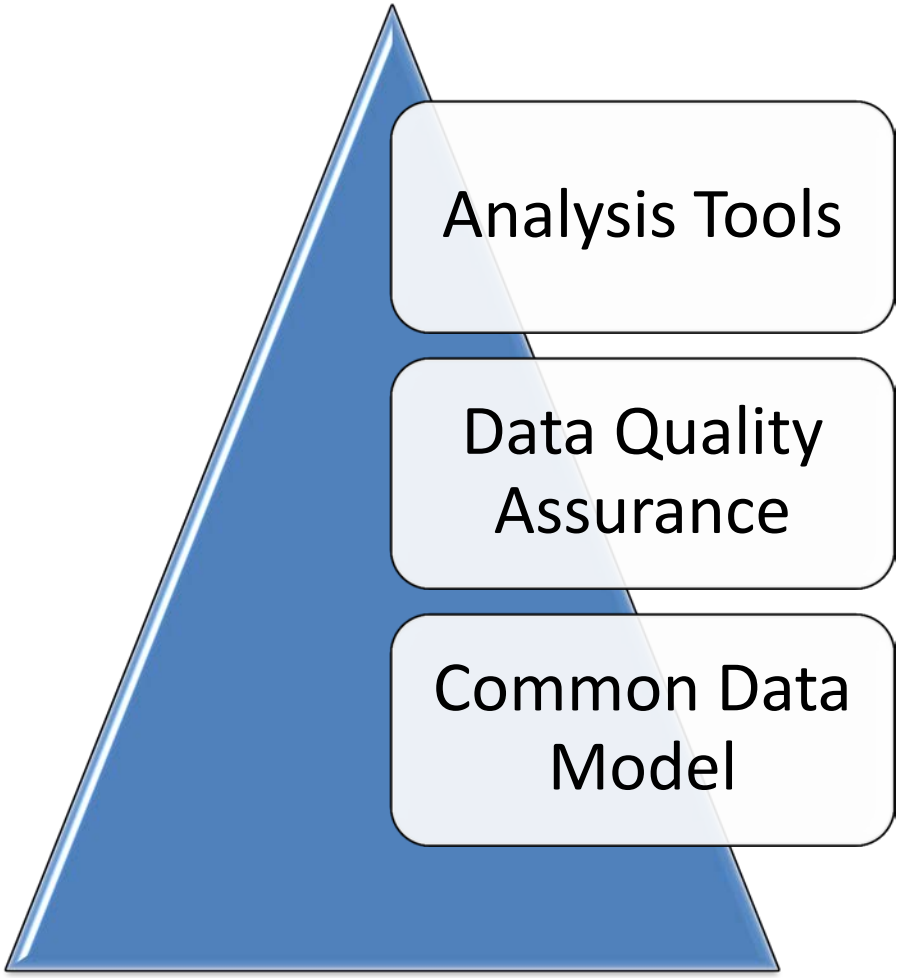
Guiding Ideals

- Address a gap in safety surveillance
- Fit within the existing regulatory paradigm, process, and culture of FDA
- Generate credible scientific evidence to support medical product regulation about risks and benefits
- Serve as a national resource for evidence development
- Meet legislative requirement to create an active postmarket risk identification and analysis system

Operational Translation

- Must be capable of 1st class epidemiologic science and function within a regulatory ecosystem built upon clinical trials
- Must account for an end user comprised of a multidisciplinary team led by an FDA epidemiologist
- Must provide actionable evidence to regulators and policy makers
- Must support numerous use cases such as safety surveillance, medication errors, comparative effectiveness, etc.
- Must be transparent to facilitate consistent decisions about when to use the system and communicate its results to a wide audience

Sentinel is More than the CDM



CDM is Part of an Ecosystem

- Sentinel CDM cannot be isolated from the Sentinel ecosystem
- Sentinel CDM designed to work with elaborate QA process and highly customizable reusable analysis tools

https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-data-quality-assurance-practices

Sentinel

FDA

SEN

PRA

COMP

PRACT

PHARI

HEALT

Prepar

Output tables are designed to evaluate one or more data checks, i.e., pre-defined data quality measures or characterizations. Approximately 1,200 data checks are evaluated during each DP data refresh. Each data check is designated a "level 1," "level 2," "level 3," or "level 4" data quality check depending on the complexity of a data characteristic/issue:

- Level 1 data checks review the completeness and content of each variable in each table to ensure that the required variables contain data and conform to the formats specified by the SCDM specifications (e.g., data types, variable lengths, SAS formats, acceptable values, etc.).
- Level 2 data checks assess the logical relationship and integrity of data values within a variable or between two or more variables within and between tables (e.g., variable ADMITTING_SOURCE in the Encounter table is populated only for inpatient and institutional encounters).
- Level 3 data checks examine data distributions and trends over time, both within a Data Partner's database (by examining output by year and year/month) and across a Data Partner's databases (by comparing updated SCDM tables to previous versions of the tables). For example, a level 3 data check would ensure that there are no large, unexpected increases or decreases in records over time.
- Level 4 data checks examine the occurrence and prevalence of nonsensical diagnoses and examine variations in care practices across Data Partners (e.g., the proportion of prostate cancer diagnoses among women). Level 4 checks are designed to provide more targeted data analyses and profiling of Data Partner data, and are not necessarily designed to detect and correct errors.

https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-data-quality-assurance-practices

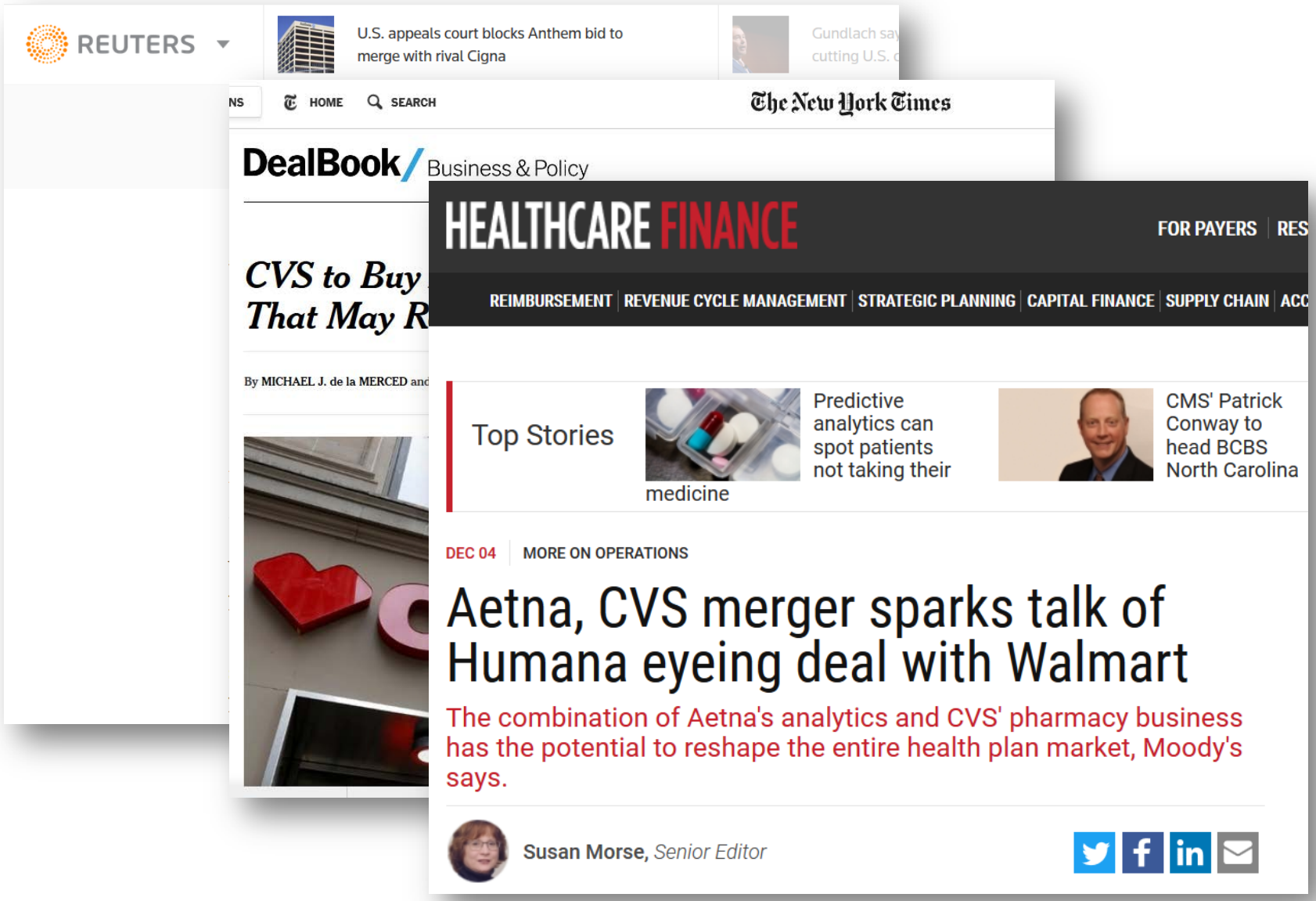
Plan for Talk

- Foundational needs and goals for Sentinel
- Sentinel system is more than the Sentinel CDM
- What the CDM does for FDA and why it meets our needs
  - CDM as data manager and curator
  - CDM as unifier and buffer
  - CDM as enabler of analytic scale and customization
  - CDM as accelerator of public health response
- Example of FDA study design process
- Things we wish the CDM could do or fix
- Towards some guiding principles from an FDA perspective

CDM as Data Manager and Curator

- **Originating legislative mandate set forth 2 key ideals**
  - Substantial sample size
  - Use of both private and public data sources
- **CDM facilitates data management**
  - Routinely extracts and transforms data across multiple sites
  - Core data elements well-defined with consistent and known clinical meaning and understanding of data provenance
- **CDM facilitates data curation**
  - Enables robust quality assurance testing across sites
  - Allows analytic tools to run on a trusted and curated dataset

CDM as Change Buffer and Unifier



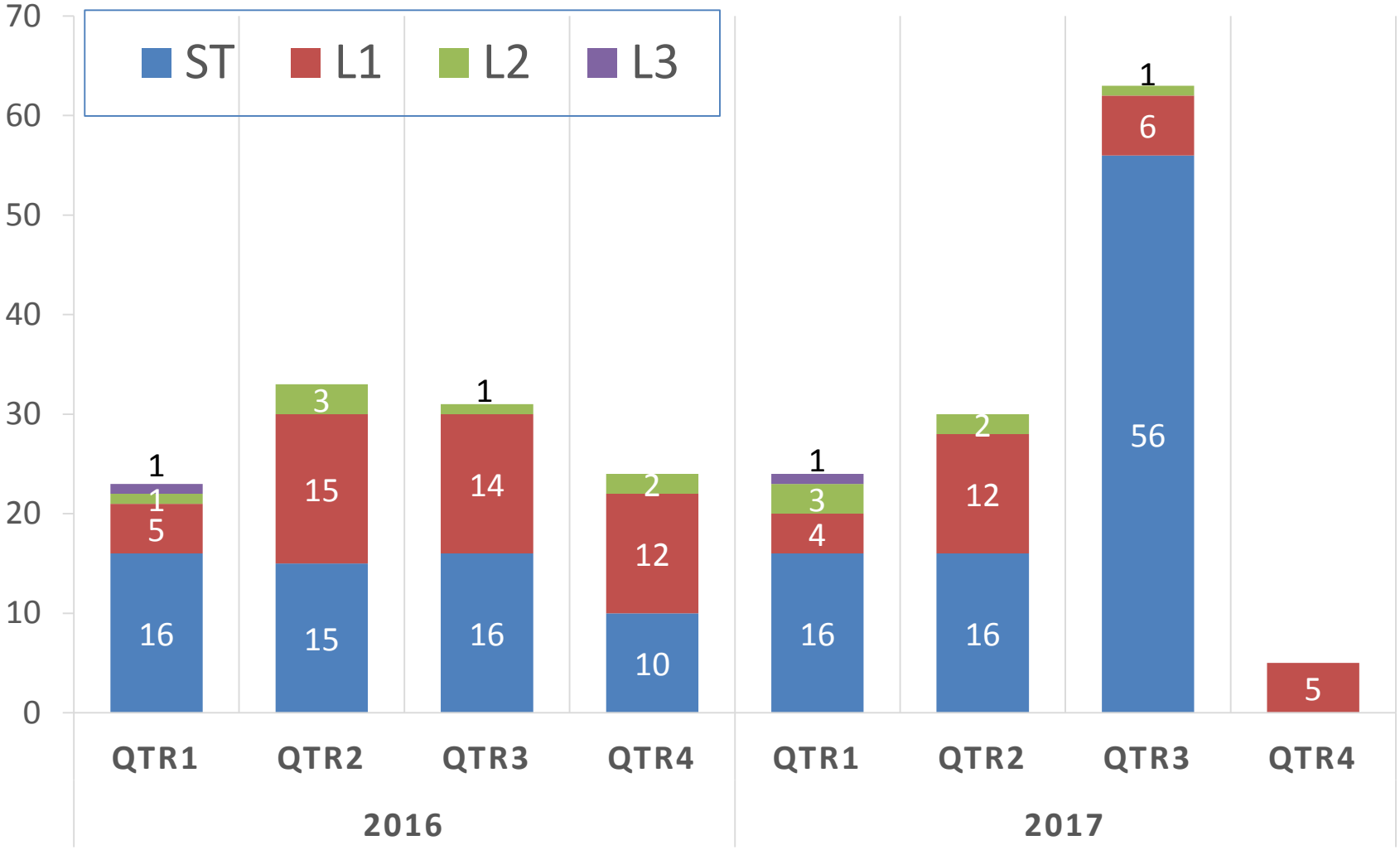
CDM as Change Buffer and Unifier

- **CDM as Buffer**
  - Market and regulatory forces will result in a constantly changing healthcare system
  - Buffers against changes in IT platforms and data infrastructure that results from mergers, acquisitions, routine business needs
  - **FDA exercises version-control** over CDM to ensure regulatory needs are always met
- **CDM as unifier**
  - CDM allows diversity of data sources to participate (e.g., national health insurer, integrated delivery system, registry, eHR)
  - CDM unifies different sources with values with known meaning
  - CDM does not mix data from different data sources (e.g., eHR has prescribing data, while claims have dispensing data; each characterize exposure differently)
  - CDM achieves requisite sample size

CDM as Enabler of Analytic Scale and Customization

- **CDM supports analytic scale**
  - Data within the CDM is quality checked routinely before any analysis is run, instead of as-needed basis
  - Once curated, any number of analyses can be run
- **CDM supports highly tailored analyses**
  - FDA needs dozens of finely customized analyses, not thousands of standard analyses that permute design choices
  - CDM and tools must have minimal mapping and allow FDA to precisely tailor parameters to the specific question
    - Exposure and outcome definitions, stockpiling, covariate adjustment
    - CDM and tools do not automate/build-in study design choices or algorithms
  - Allows FDA to explain analyses to other regulators, and allows others to reproduce analyses

Sentinel ARIA Analyses (N=233)



From all FDA Centers, by date of distribution to Data Partners

CDM as Accelerator of Public Health Response

- Sentinel is an "opt-in" public health program
  - Data partners participate because they believe in the public health mission
  - But they always clear analyses and results
- Simple but rich CDM, combined with well described analytic tools, and standard output formats facilitate:
  - Data Partner decision to participate in queries
  - Operational speed for clearance of urgent requests (<1 week across all 16 claims based data partners)

Summary: What the CDM does for FDA and why it meets our needs

Data management & curation + Change Buffer & unifier = Data Quality

Data Quality + Analytic customization = Validity

Validity + Analytic scale = Speed (public health response)

Sufficiency to Address the Regulatory Question

Data management & curation + Change Buffer & unifier = Data Quality

+

Data Quality + Analytic customization = Validity

+

Validity + Analytic scale = Speed (public health response)

Plan for Talk

- Foundational needs and goals for Sentinel
- Sentinel system is more than the Sentinel CDM
- What the CDM does for FDA and why it meets our needs
  - CDM as data manager and curator
  - CDM as unifier and buffer
  - CDM as enabler of analytic scale and customization
  - CDM as accelerator of public health response
- Example of FDA study design process
- Things we wish the CDM could do or fix
- Towards some guiding principles from an FDA perspective

ISPE 2017 Conference Symposium

Propensity Score Matching L2 tool

Venous thromboembolism after oral contraceptives
By David Moeny

Stroke after antipsychotics medications
By Lockwood Taylor

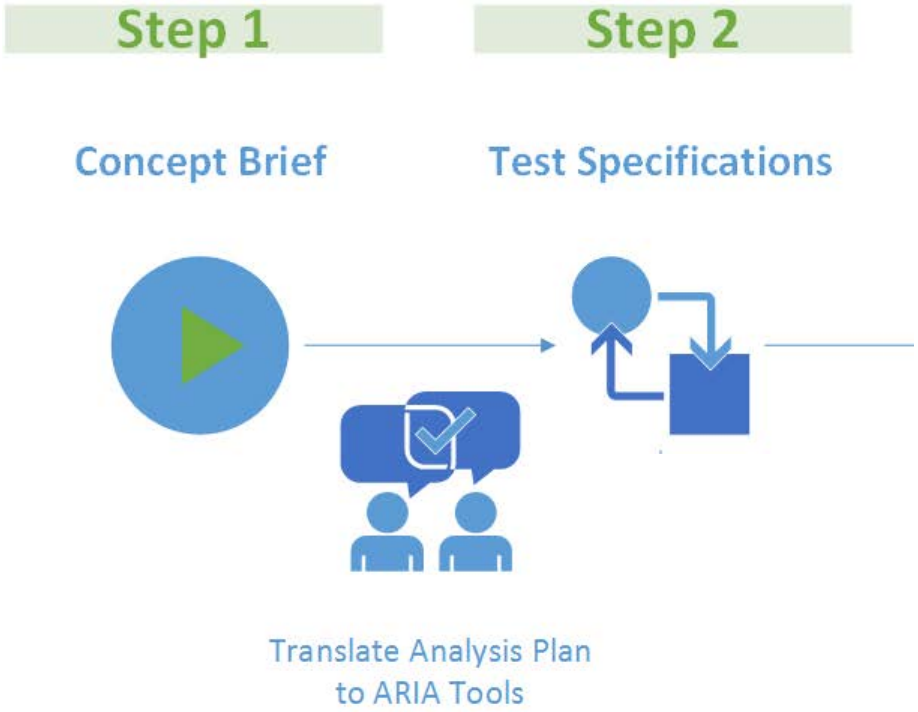
Self-controlled L2 tool

Seizures after ranolazine
By Efe Eworuke

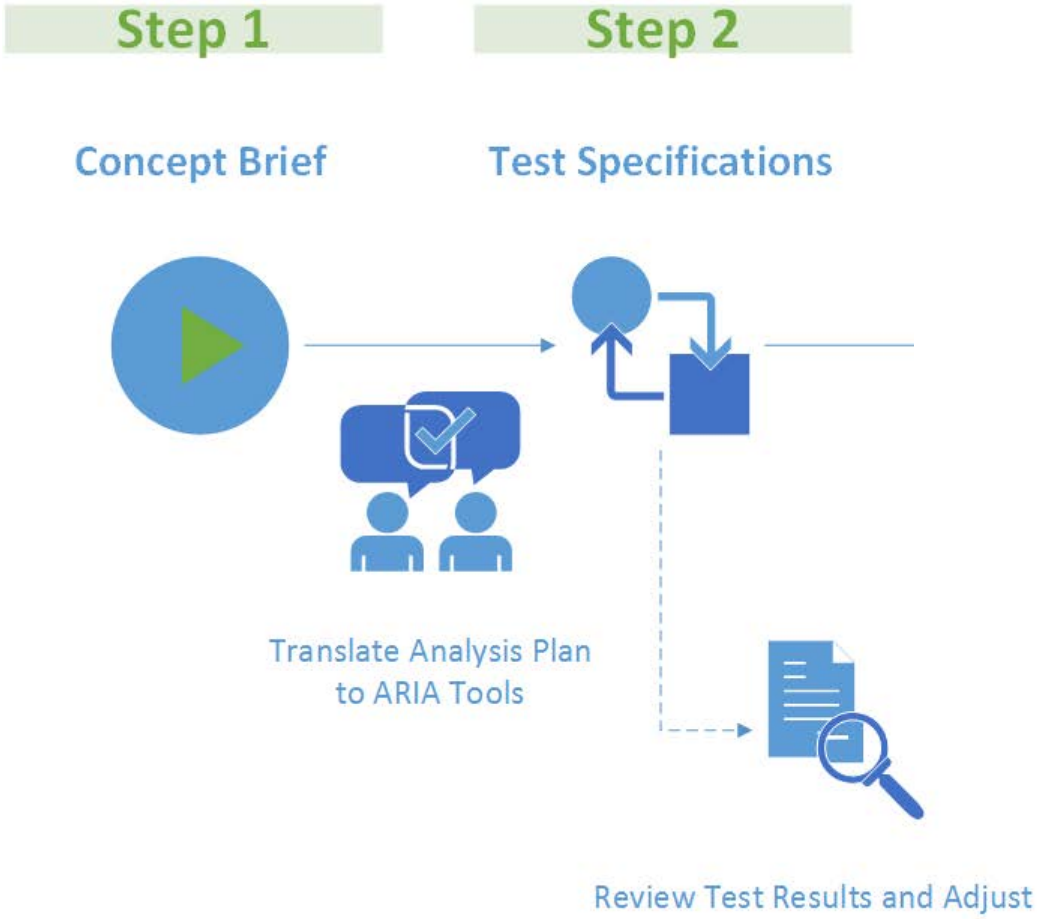
Seizures after gadolinium based contrast agents
By Steve Bird

https://www.sentinelinitiative.org/communications/publications/2017-icpe-symposium-integrating-sentinel-routine-regulatory-drug-review

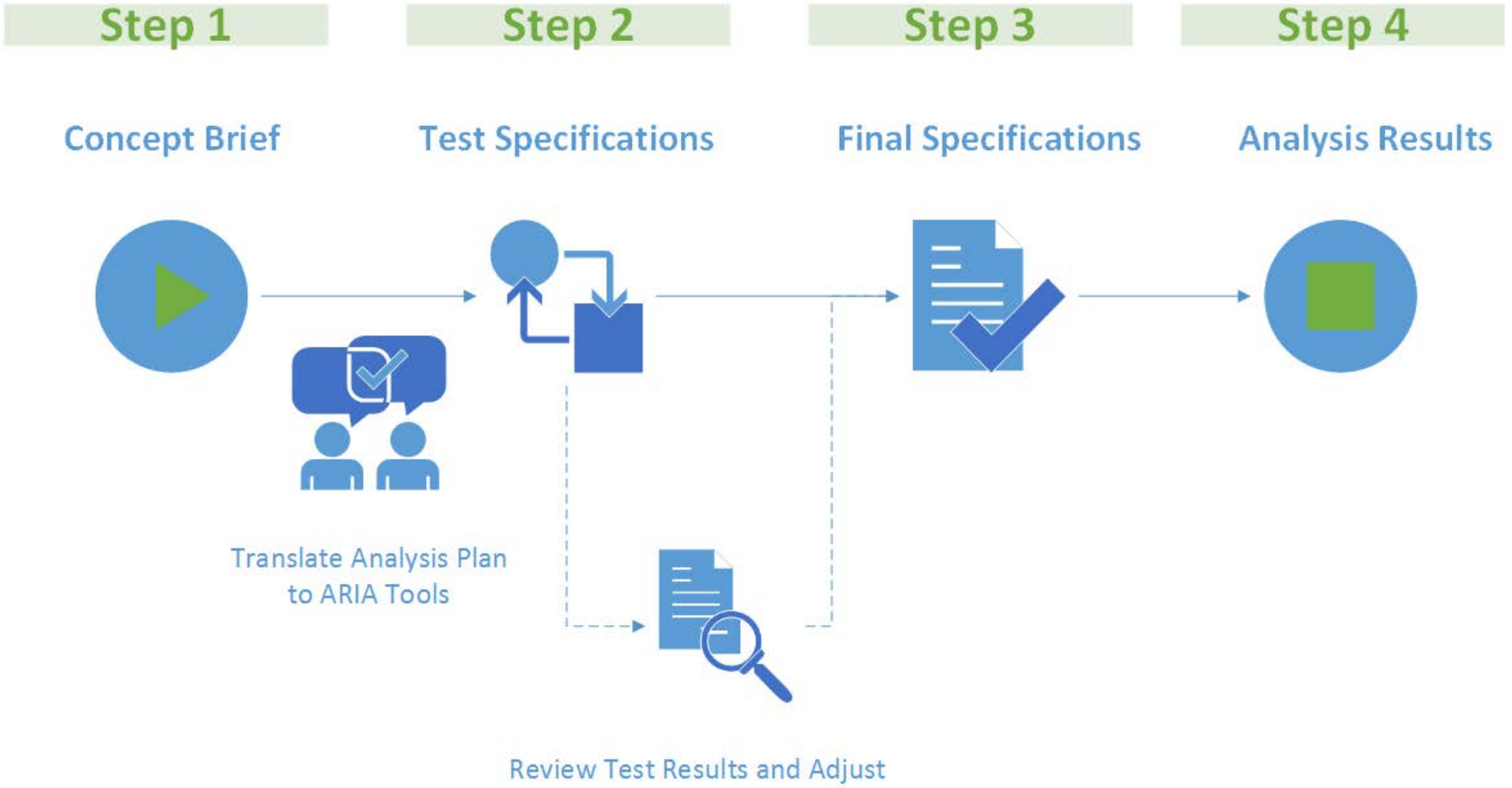
Query Development Process



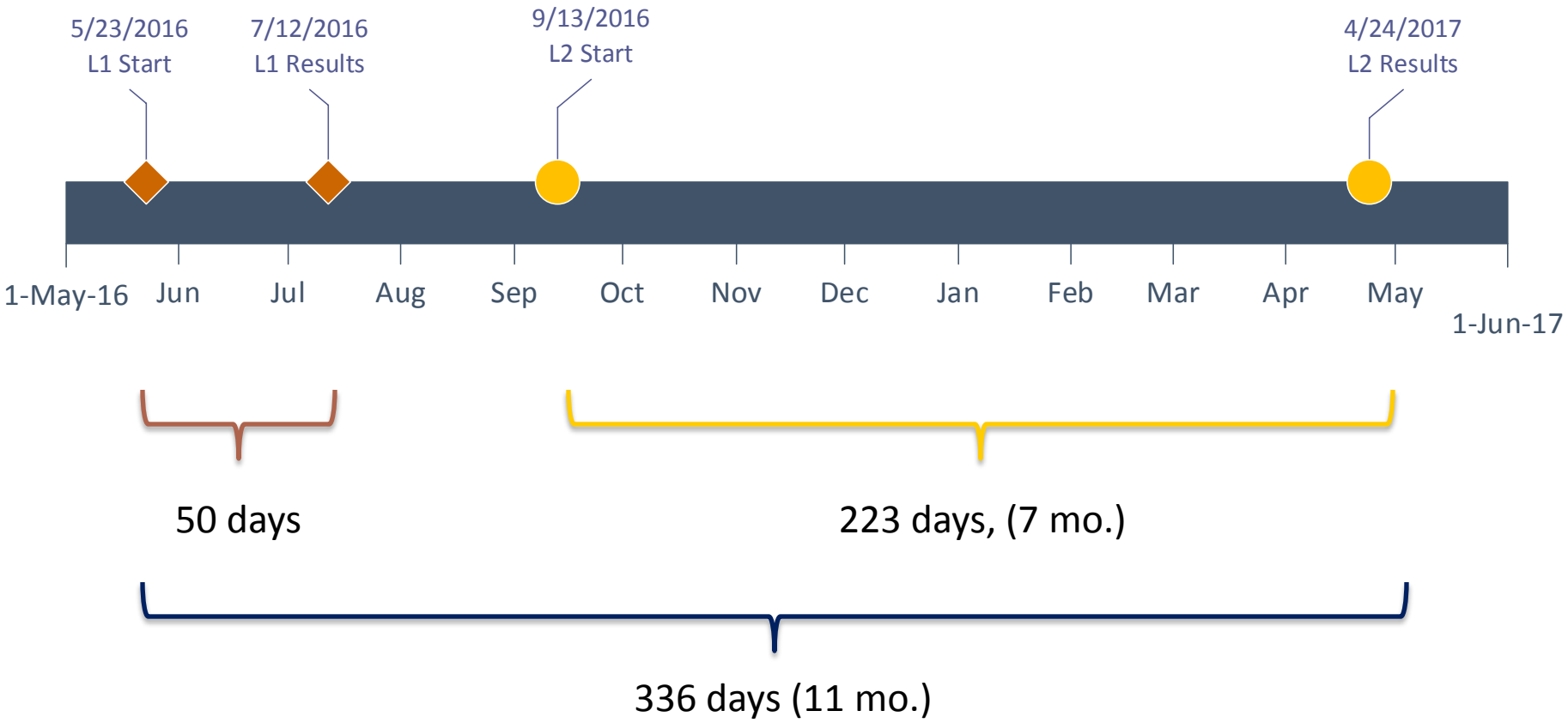
Query Development Process



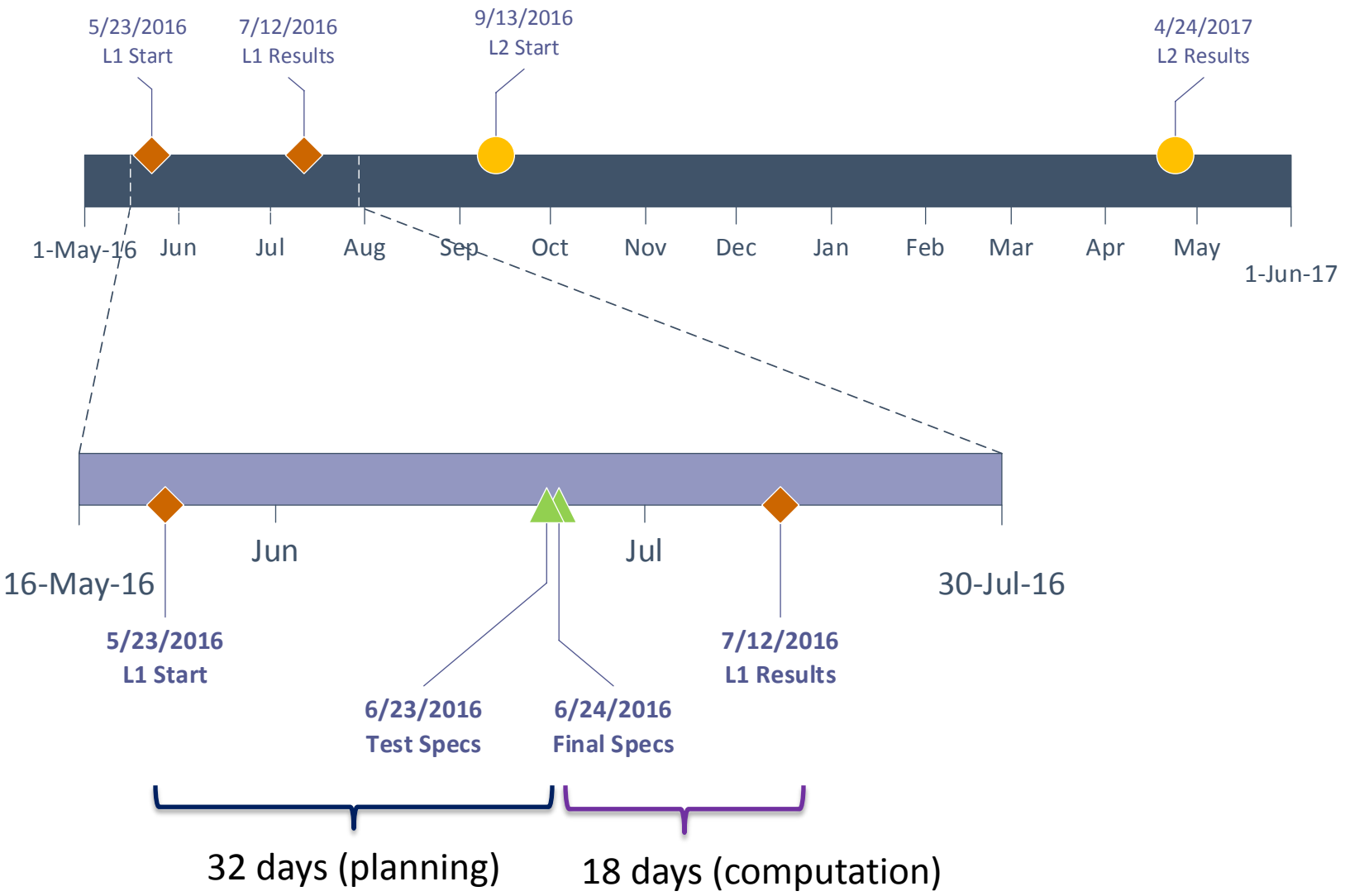
Query Development Process



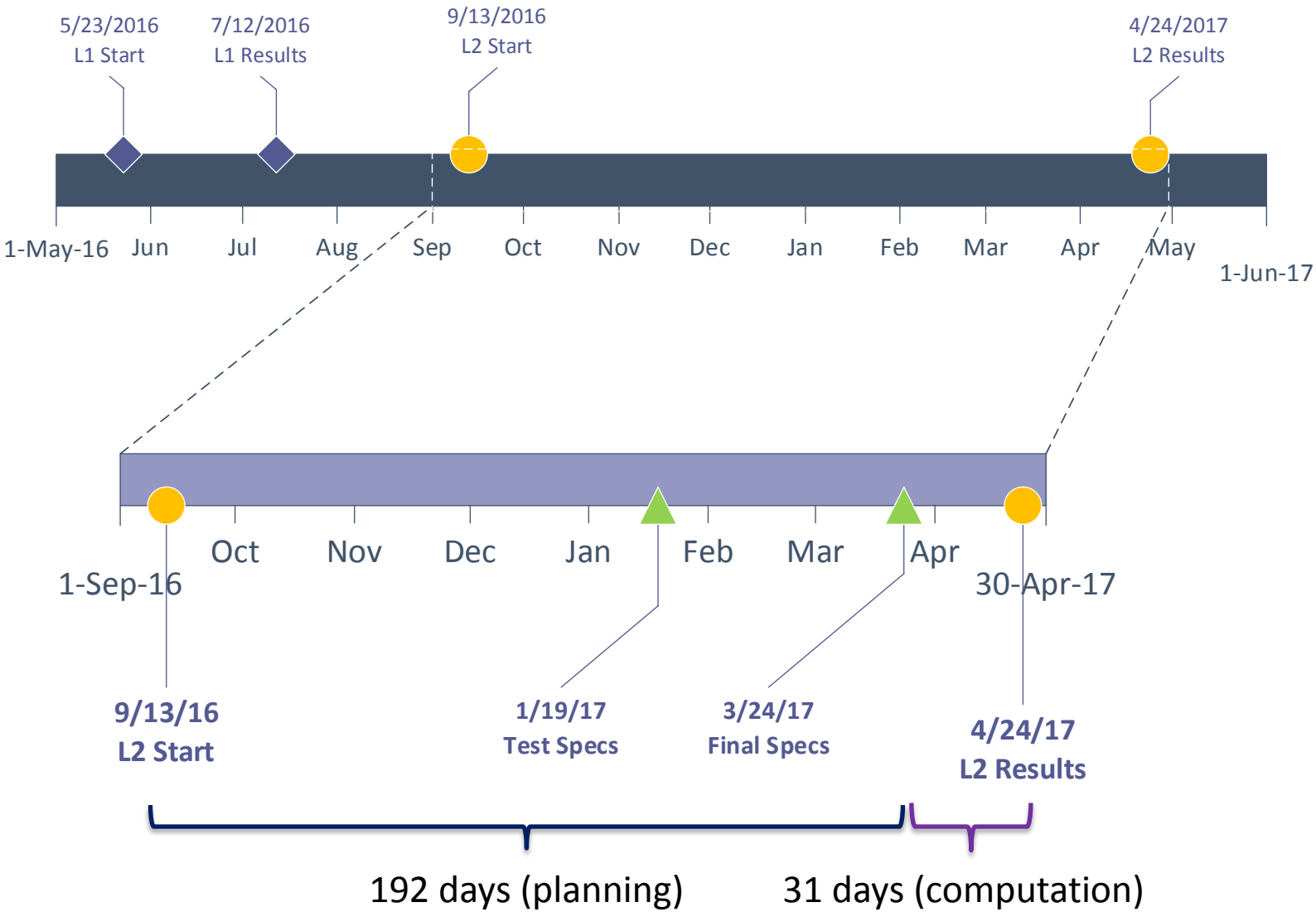
Timeline: Oral Contraceptive Analysis



Detailed Timeline: Level 1 Descriptive Analysis



Detailed Timeline Level 2 Inferential Analysis



Total Time to Assess Safety Issue

Safety Issue	Type and No. Analyses	Total Time, days
Ranexa (ranolazine)	ST, L1, L2	302 (10 mo.)
Gadolinium contrast agents	L1, L2	741 (24 mo.)
Antipsychotics	L1, L2	277 (9 mo.)
Oral contraceptives	L1, L2	336 (11 mo.)

- ST = Summary table, simple counts
- L1 = Level 1, complex descriptive analysis
- L2 = Level 2, inferential analysis

Things We Wish the Sentinel CDM Could Do or Fix

- Improve the accuracy of claims data for health outcome identification
- Overcome fragmentation of the U.S. Healthcare System through an encrypted universal patient identifier
- Provide information about disease stage and progression (e.g., Child-Pugh liver disease prognostic score, tumor stage, diabetes disease progression)
- Provide access to insurance formularies to differentiate physician from formulary prescribing decisions

Drug Safety
August 2015, Volume 38, Issue 8, pp 683–686 | Cite as

Common Models, Different Approaches

Authors | Authors and affiliations

Joshua J. Gagne

Editorial
First Online: 19 June 2015

In recent years, a number of initiatives have established database networks for studying drug safety, including the Mini-Sentinel [1] and Observational Medical Outcomes Partnership (OMOP) [2] programs in the US, the Canadian Network for Observational Drug Effect Studies (CNODES) [3], the Asian Pharmacoepidemiology Network (AsPEN) [4], and the Exploring and Understanding Adverse Drug Reactions (EU-ADR) in Europe [5]. These networks, each comprising data for up to hundreds of millions of individuals, facilitate analyses on unprecedented numbers of patients, which can be particularly useful for evaluating very rare adverse outcomes, investigating heterogeneity across patient subgroups, or assessing outcomes

https://link.springer.com/article/10.1007%2Fs40264-015-0313-9

Create an Analytic Platform that Facilitates Highest Quality Evidence at Scale

"…CDMs and standardized analytic tools developed to interface with them must enable investigators to implement the most appropriate design and analysis plans for given drug–outcome pairs."

"To the extent that CDMs facilitate scaling of the most rigorous design and analysis plans, bigger will be better. However, scaling of inappropriate design and analysis methods will lead to more results that are precisely wrong."

Gagne J. Common Models, Different Approaches. Drug Saf. 2015 Aug;38(8):683-6.

Potential Guiding Principles from an FDA Perspective*

- Build the system around your primary objective
  - Generate the highest quality evidence to support medical product regulation
  - Provide safety information that builds upon clinical trials
- Build the system around your end-user and stakeholder
  - End users: clinicians, epidemiologists, statisticians at a regulatory Agency
  - Stakeholders: regulators, industry, patients, advocacy groups
- Create a CDM and analytic platform that allows the investigator to implement the most appropriate study design and parameters for a specific question
- Minimize vocabulary mapping to best understand data provenance
- Enable traceback to the individual level medical record, when needed
- Implement data protection and control safeguards (e.g., distributed database)
- Ensure transparency and reproducibility of data, tools, study design

* To help achieve EMA's meeting objective


FDA
U.S. FOOD & DRUG
ADMINISTRATION