

Harnessing big data for medical product safety surveillance: The experience of the Sentinel Initiative

**Judith C. Maro, PhD
Harvard Medical School
and Harvard Pilgrim Health Care Institute, Boston, MA**

Agenda

- Sentinel 101
- Answering Safety Questions in Sentinel
- Selected Methods Projects
- Selected FDA-Catalyst Projects

Sentinel and the United States Food and Drug Administration's (FDA) Mandate

Section 905

Mandates creation of Sentinel

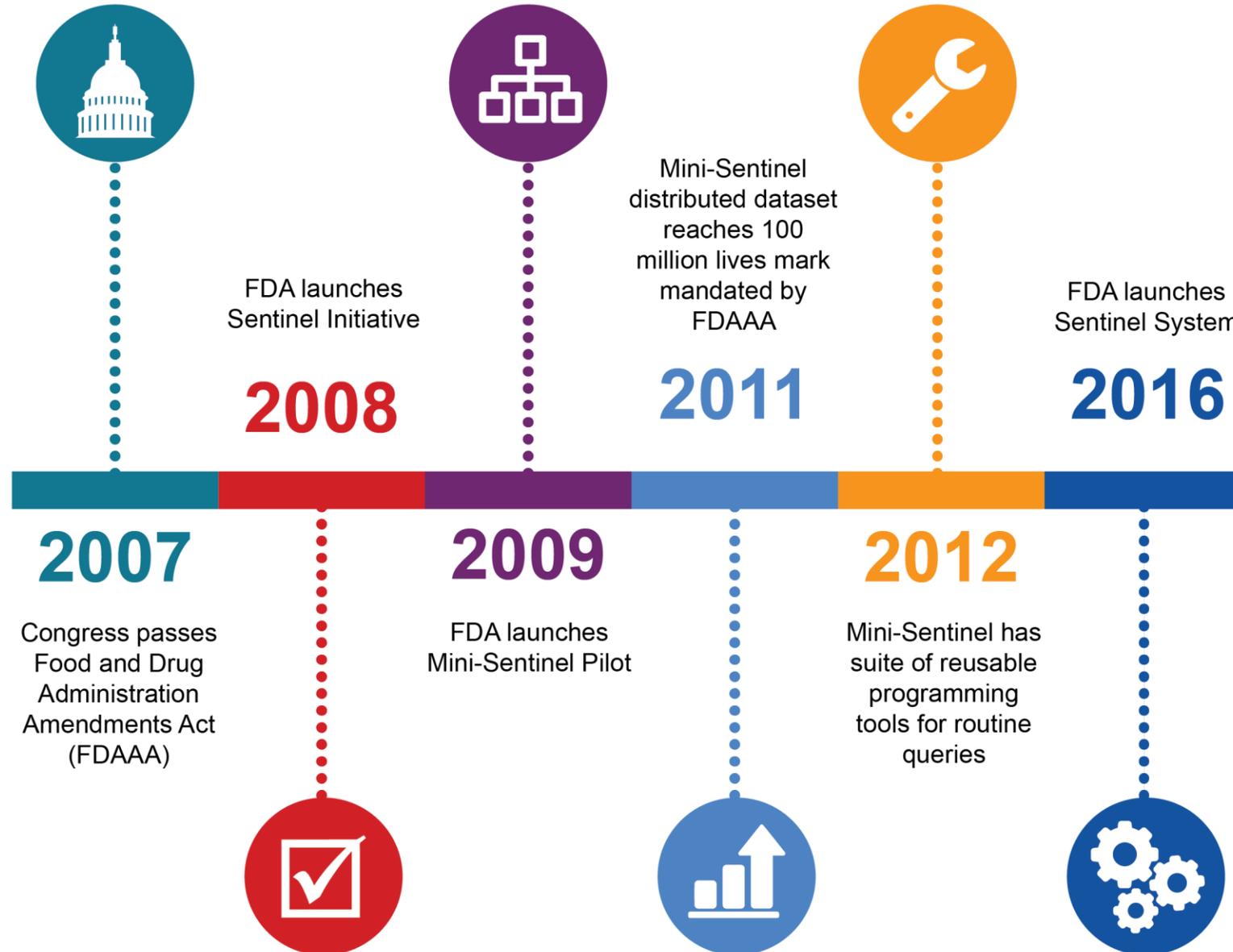


Section 901

New Food and Drug Administration Amendments Act (FDAAA) Postmarketing Requirements (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).”

Creation and Evolution of Sentinel

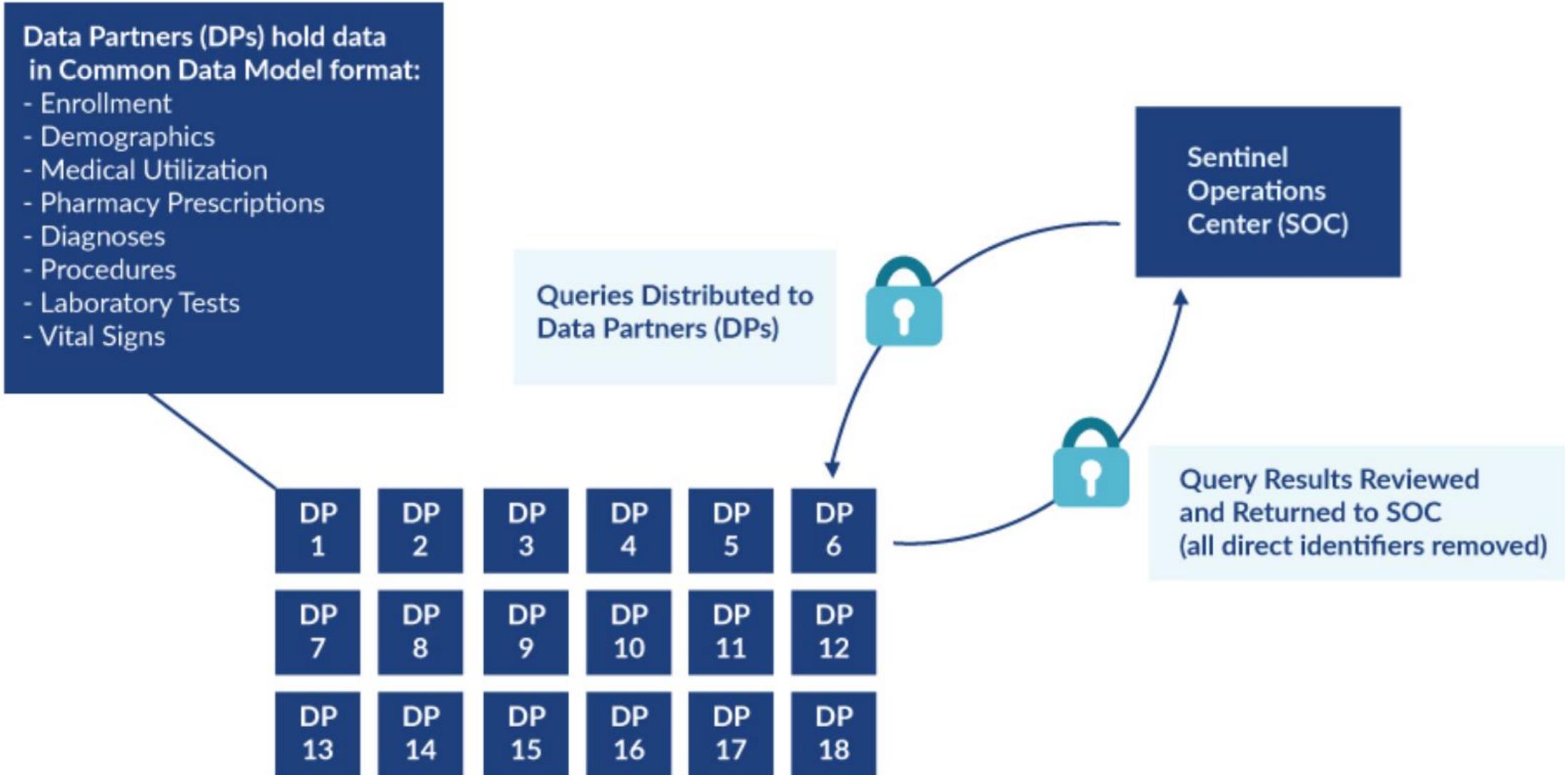


Sentinel Design Requirements



- Electronic health data for >100M persons
 - Include special populations (pregnant women, elderly)
 - Ability to link to external sources, e.g., National Death Index
 - Ability to access full text medical records
- Expertise in the way health care delivery and payment influence electronic healthcare data
- Rapid answers to many FDA safety questions
- Accuracy sufficient to support regulatory decision making
- Federal Information Security Management Act (FISMA)-compliant data security
- Ability to protect non-public information and to keep records on all data requests for public record-keeping

Sentinel Distributed Database



Sentinel Partner Organizations



Lead – HPHC Institute

DEPARTMENT OF POPULATION MEDICINE



Data & scientific partners



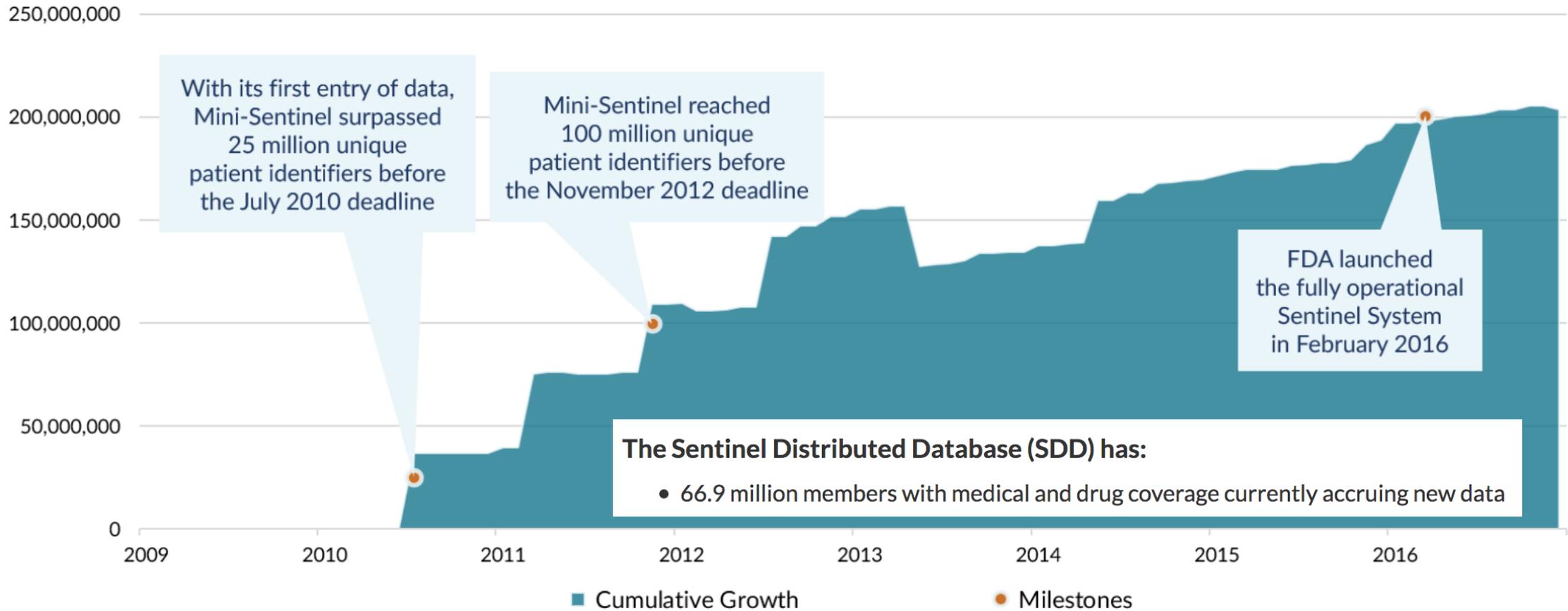
Scientific partners



Sentinel Distributed Database Characteristics



Growth of the Sentinel Distributed Database



The area above depicts the cumulative number of unique patient identifiers in the Sentinel Distributed Database from 2010 to present. If patients move health plans, they may have more than one patient identifier.

Sentinel Common Data Model Guiding Principles



- Includes claims, electronic health record (EHR), and registry data and flexible enough to accommodate new data domains (e.g., free text).
 - Typically, we do not include empty tables – we expand as needed when fit for purpose.
- Data are stored at most **granular/raw level possible** with minimal mapping.
 - Distinct data types should be kept separate (e.g., prescriptions, dispensings)
 - Construction of medical concepts (e.g., outcome algorithms) from these elemental data is a **project-specific** design choice.
 - Sentinel stores these algorithms in a library for future use.
- Appropriate use and interpretation of local data requires the Data Partners' local knowledge and data expertise.
 - Not all tables are populated by all Data Partners → site-specificity is allowed.
- Designed to meet FDA needs for analytic flexibility, transparency, and control.

Sentinel Common Data Model v 6.0



Administrative

Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure
Person ID	Person ID	Person ID	Person ID	Person ID	Person ID
Enrollment start & end dates	Birth date	Dispensing date	Service date(s)	Service dates	Service date(s)
Drug coverage	Sex	National drug code (NDC)	Encounter ID	Encounter ID	Encounter ID
Medical coverage	Zip code	Days supply	Encounter type and provider	Encounter type and provider	Encounter type & provider
Medical record availability	Etc.	Amount dispensed	Facility	Diagnosis code & type	Procedure code & type
			Etc.	Principal discharge diagnosis	Etc.

Clinical

Lab Result
Person ID
Result and specimen collection dates
Test type, immediacy & location
Logical Observation Identifiers Names and Codes (LOINC®)
Test result & unit
Etc.

Vital Signs

Person ID
Measurement date & time
Height & weight
Diastolic & systolic BP
Tobacco use & type
Etc.

Registry

Death
Person ID
Death date
Source
Confidence
Etc.

Cause of Death

Person ID
Cause of death
Source
Confidence
Etc.

State Vaccine

Person ID
Vaccination date
Admission type
Vaccine code & type
Provider
Etc.

Inpatient

Inpatient Pharmacy
Person ID
Administration date & time
Encounter ID
National Drug Code (NDC)
Route
Dose
Etc.

Inpatient Transfusion
Person ID
Administration start & end date & time
Encounter ID
Transfusion administration ID
Transfusion product code
Blood type
Etc.

Single Patient Example Data in Model



DEMOGRAPHIC

PATID	BIRTH_DATE	SEX	HISPANIC	RACE	zip
PatID1	2/2/1964	F	N	5	32818

DISPENSING

PATID	RXDATE	NDC	RXSUP	RXAMT
PatID1	10/14/2005	00006074031	30	30
PatID1	10/14/2005	00185094098	30	30
PatID1	10/17/2005	00378015210	30	45
PatID1	10/17/2005	54092039101	30	30
PatID1	10/21/2005	00173073001	30	30
PatID1	10/21/2005	49884074311	30	30
PatID1	10/21/2005	58177026408	30	60
PatID1	10/22/2005	00093720656	30	30
PatID1	10/23/2005	00310027510	30	15

ENROLLMENT

PATID	ENR_START	ENR_END	MEDCOV	DRUGCOV
PatID1	7/1/2004	12/31/2004	Y	N
PatID1	1/1/2005	12/31/2005	Y	Y

DEATH

PATID	DEATHDT	DTIMPUTE	SOURCE	CONFIDENCE
PatID1	12/27/2005	N	S	E

ENCOUNTER

PATID	ENCOUNTERID	ADATE	DDATE	ENCTYPE
PatID1	EncID1	10/18/2005	10/20/2005	IP

DIAGNOSIS

PATID	ENCOUNTERID	ADATE	PROVIDER	ENCTYPE	DX	DX_CODETYPE	PDX
PatID1	EncID1	10/18/2005	Provider1	IP	296.2		9 P
PatID1	EncID1	10/18/2005	Provider1	IP	300.02		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	305.6		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	311		9 P
PatID1	EncID1	10/18/2005	Provider1	IP	401.9		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	493.9		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	715.9		9 S

PROCEDURE

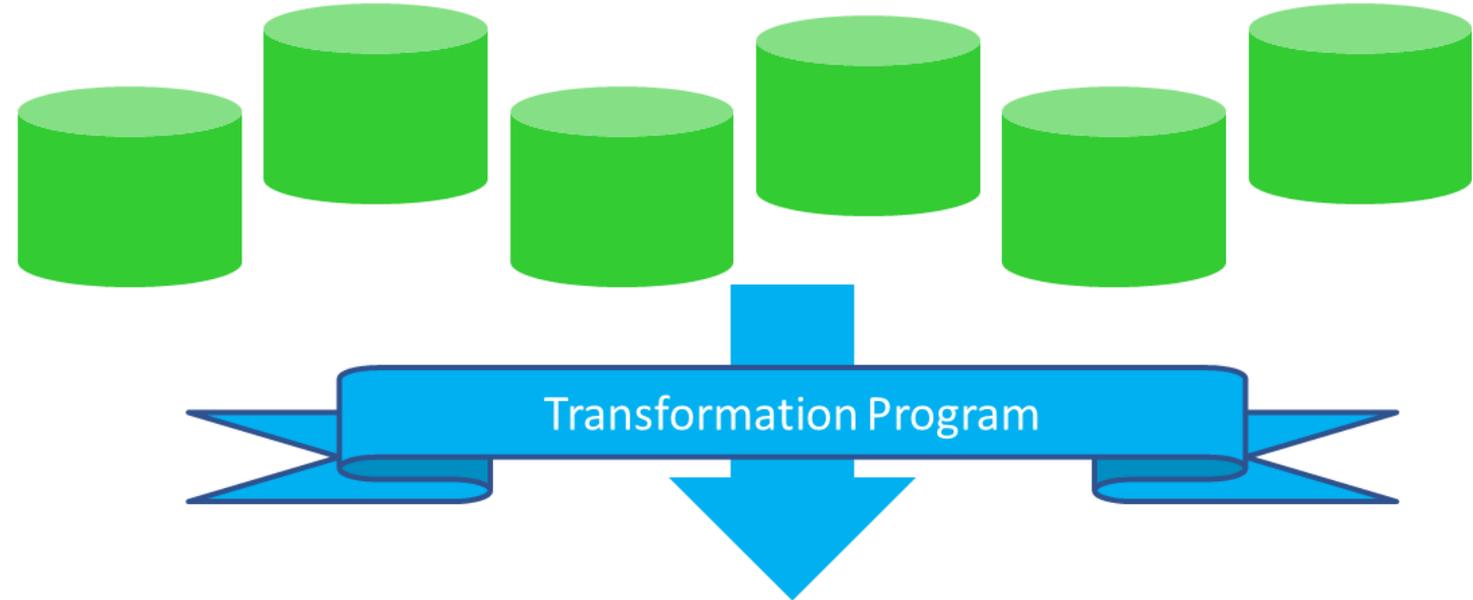
PATID	ENCOUNTERID	ADATE	PROVIDER	ENCTYPE	PX	PX_CODETYPE
PatID1	EncID1	10/18/2005	Provider1	IP	84443	C4
PatID1	EncID1	10/18/2005	Provider1	IP	99222	C4
PatID1	EncID1	10/18/2005	Provider1	IP	99238	C4
PatID1	EncID1	10/18/2005	Provider2	IP	27445	C4

CAUSE OF DEATH

PATID	COD	CODETYPE	CAUSETYPE	SOURCE	CONFIDENCE
PatID1	J18.0	10	U	S	E

Every Data Partner Transforms their Source Data into the Sentinel Common Data Model

Unique Data Partner's Source Database Structure



Data Partner's Database Transformed into SCDM Format (Refresh)

Administrative					
Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure
Person ID	Person ID	Person ID	Person ID	Person ID	Person ID
Enrollment start & end dates	Birth date	Dispensing date	Service date(s)	Service dates	Service date(s)
Drug coverage	Sex	National drug code (NDC)	Encounter ID	Encounter ID	Encounter ID
Medical coverage	Zip code	Days supply	Encounter type and provider	Encounter type and provider	Encounter type & provider
Medical record availability	Etc.	Amount dispensed	Facility	Diagnosis code & type	Procedure code & type
			Etc.	Principal discharge diagnosis	Etc.

Clinical		Registry			Inpatient	
Lab Result	Vital Signs	Death	Cause of Death	State Vaccine	Inpatient Pharmacy	Inpatient Transfusion
Person ID	Person ID	Person ID	Person ID	Person ID	Person ID	Person ID
Result and specimen collection dates	Measurement date & time	Death date	Cause of death	Vaccination date	Administration date & time	Administration start & end date & time
Test type, immediacy & location	Height & weight	Source	Source	Administration type	Encounter ID	Encounter ID
Logical Observation Identifiers Names and Codes (LOINC*)	Diastolic & systolic BP	Confidence	Confidence	Vaccine code & type	National Drug Code (NDC)	Transfusion administration ID
Test result & unit	Tobacco use & type	Etc.	Etc.	Provider	Route	Transfusion product code
Etc.	Etc.			Etc.	Dose	Blood type
					Etc.	Etc.

Guidance for Industry and FDA Staff

**Best Practices for Conducting
and Reporting**

Pharmacoepidemiologic Safety

Studies Using Electronic

Healthcare Data

Project-Specific v. System Data Curation

Project-Specific	System
“As needed / as-you go”	“Always Ready”
<i>Ad hoc</i>	Repeatable, systematic
Variable amount of data cleaning	1200+ checks to pass each dataset
Burden on study team	Burden on Quality Assurance team
Cost is included in the cost of a study	Cost is front-loaded for studies that use system

Take-home message: “Making data fit for purpose” at scale entails cost and time trade-offs.

Submit Comment

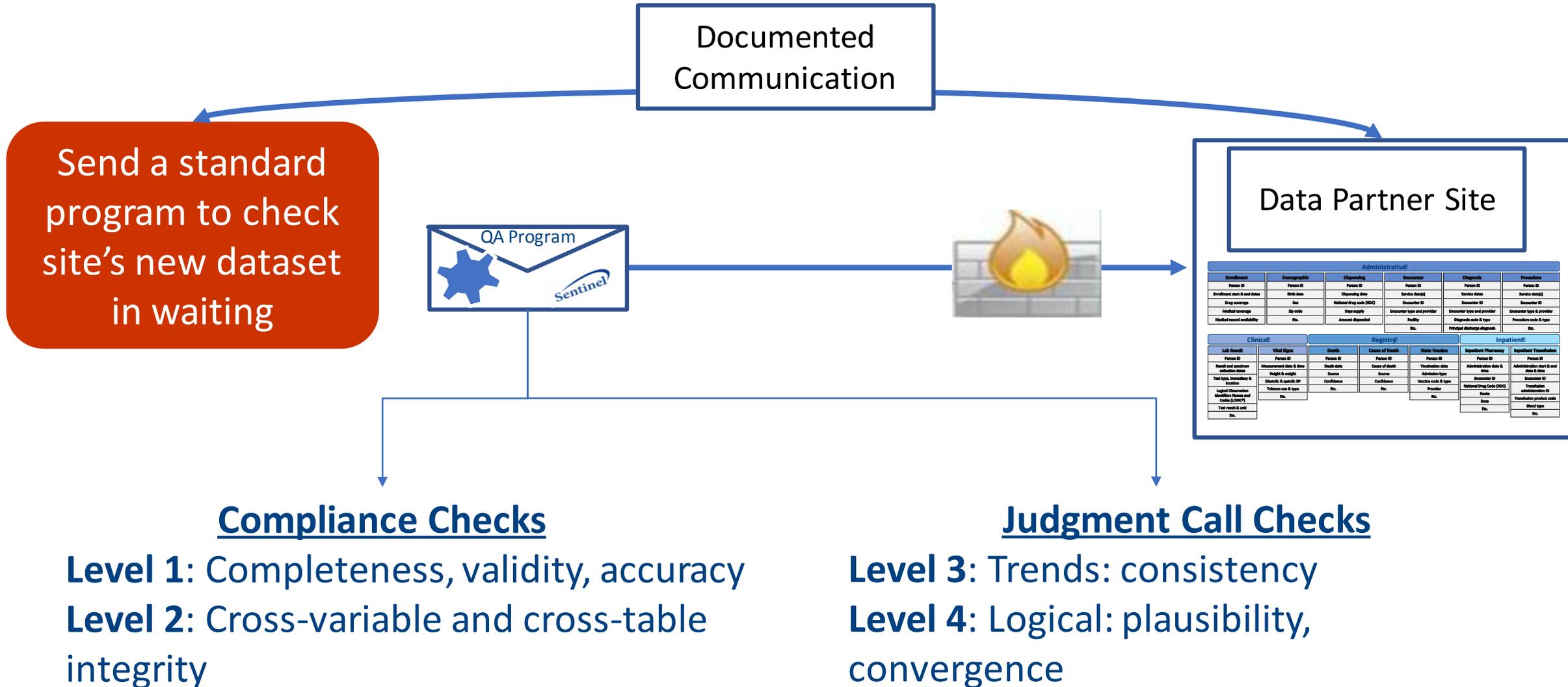
Sentinel Data Quality Assurance Practices

Project Title	Sentinel Data Quality Assurance Practices
Date Posted	<i>Thursday, March 23, 2017</i>
Status	Complete
Deliverables	Sentinel Data Quality Assurance Practices
Description	<p>The Food and Drug Administration (FDA) set forth its current recommendations for data quality assurance (QA) in the following document: “Guidance for Industry and FDA Staff: Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data” (Guidance), section IV.E “Best Practices – Data Sources: Quality Assurance (QA) and Quality Control (QC),” in May 2013. This Guidance describes best practices that particularly apply to observational studies designed to assess the risk associated with a drug exposure using electronic healthcare data.</p>

Data Quality Review and Characterization Programs v4.1.0

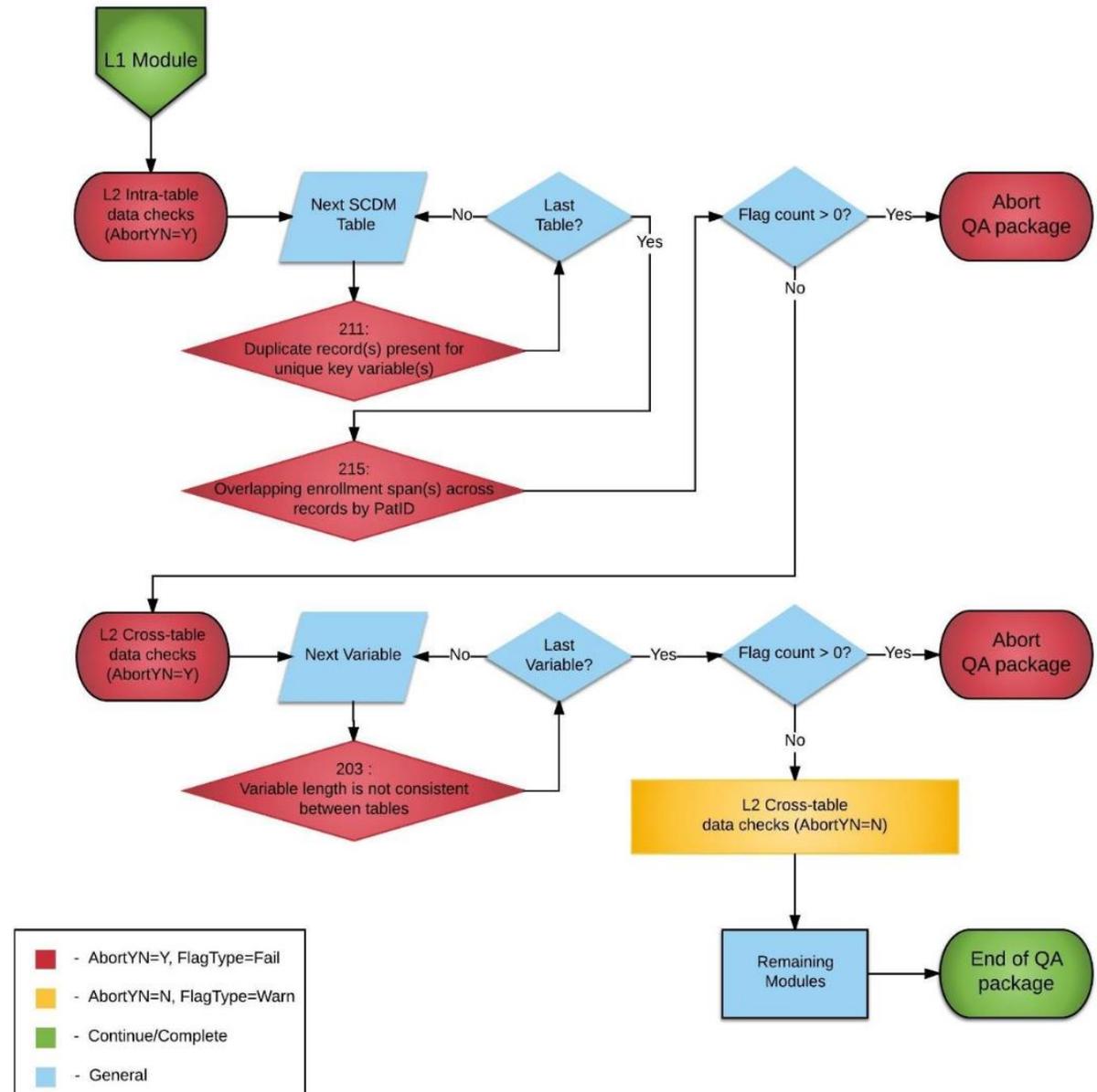
Project Title	Data Quality Review and Characterization Programs v4.1.0
Description	The Sentinel Data Quality Review and Characterization Programs are used by the Sentinel Operations Center (SOC) for data quality review and characterization of the Sentinel Distributed Database (SDD). To create the SDD, each Data Partner transformed local source data into the Sentinel Common Data Model (SCDM) format. The SOC created a set of data quality review and characterization programs to ensure that the SDD meets reasonable standards for data transformation consistency and quality and that the SDD data meets expectations needed for a distributed health data network.
Link	Sentinel Data Quality Review and Characterization Programs v4.1.0 – Overview Sentinel Data Quality Review and Characterization Programs v4.1.0 – Appendix A Sentinel Data Quality Review and Characterization Programs v4.1.0 – Appendix B Sentinel Data Quality Review and Characterization Programs v4.1.0 – SAS Programs View more details here.

Data Quality Review and Characterization Process



Quality Review and Characterization Program Logic

- Compliance checks for all tables are mandatory.
- Quality Review and Characterization Program will abort after it runs through all compliance checks, producing an automatically created report on failures.



Sentinel is a National Medical Product Monitoring System

LEARN MORE



ABOUT

- Background
- Coordinating Center
- Privacy and Security
- The Sentinel System Story



MEDICAL PRODUCT ASSESSMENTS

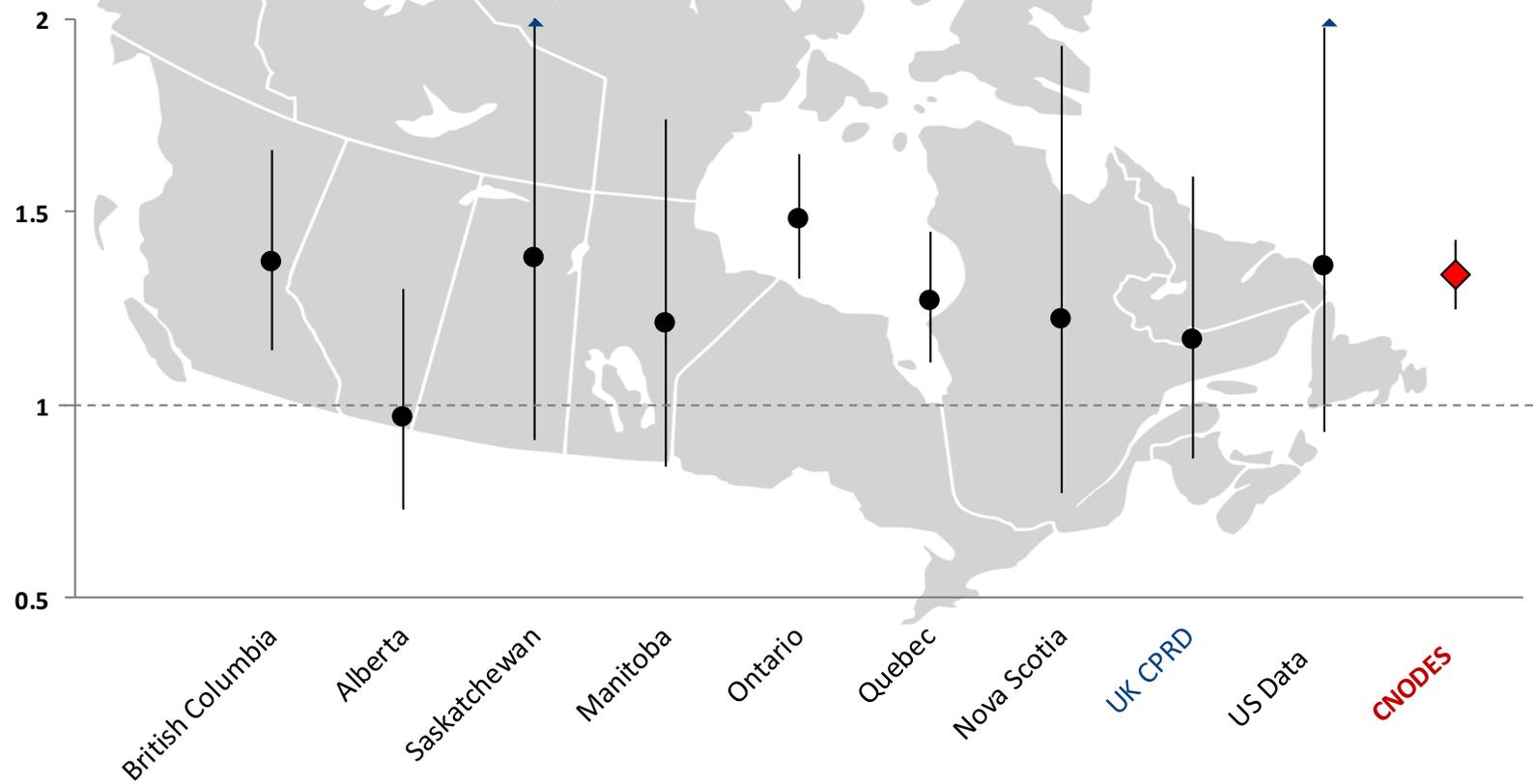
- Active Risk Identification and Analysis System
- Ongoing ARIA Assessments
- Assessments of Drugs

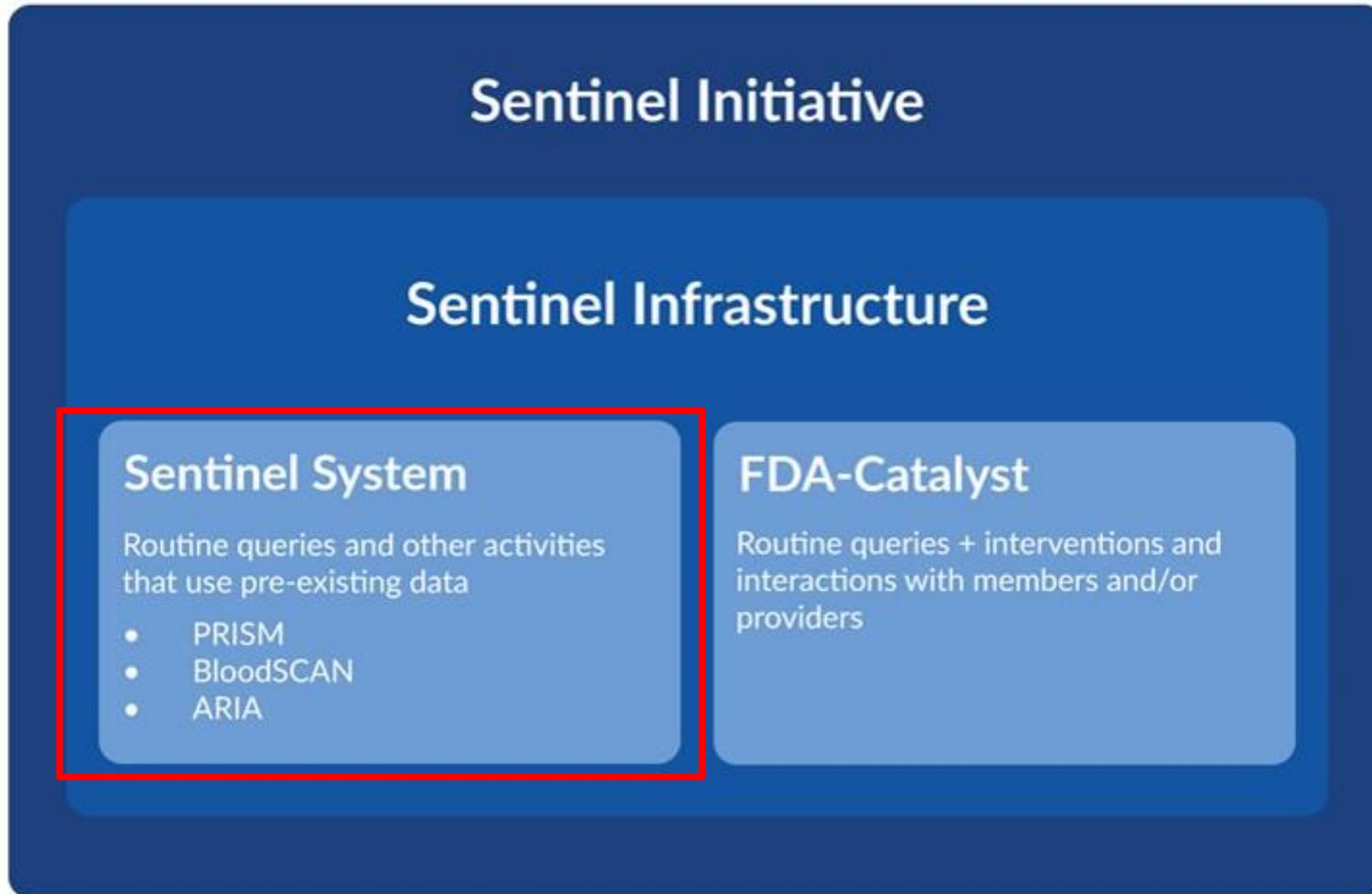
Featured Postings

SPOT

- Medicare Claims Synthetic Public Use Files in Sentinel Common Data Model Format: User Documentation, Demonstration Routine Querying Package, and Data Files
Wed, 04/25/2018

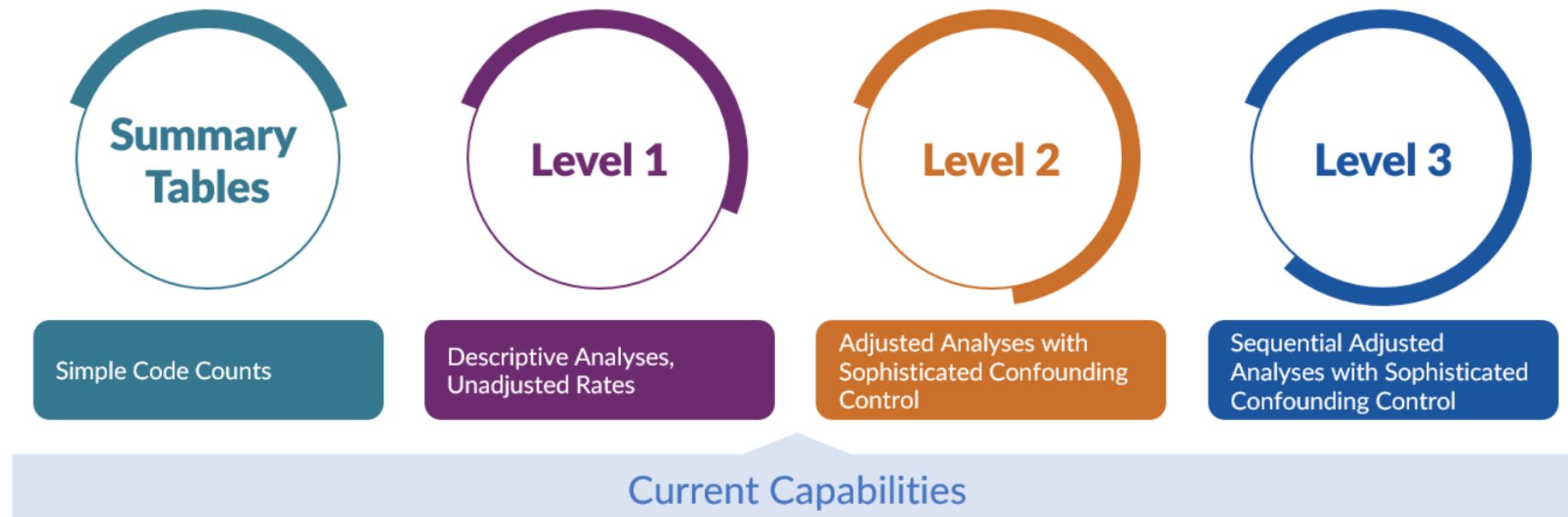
Adding Data Sources from Across Canada





Active Risk Identification and Analysis (ARIA)

ARIA is FDA's active post-market risk identification and analysis system, which is comprised of pre-defined, parameterized, reusable routine querying tools, combined with the electronic data in the Sentinel [Common Data Model](#). Because ARIA uses parameterized tools and a trusted multi-site distributed database that undergoes continuous quality checks and refreshes, safety analyses can be done more efficiently to conduct medical product safety surveillance to fulfill the mandate in the FDA Amendments Act of 2007.



Sentinel is a National Medical Product Monitoring System

LEARN MORE

DATA AND TOOLS

- Routine Querying System Documentation (version 5.4.4)
Fri, 08/17/2018
- Data Quality Review and Characterization Programs v4.1.0
Wed, 02/28/2018
- Sentinel Common Data Model v6.0.2
Wed, 10/04/2017



COMMUNICATIONS

- [FDA Safety Communications](#)
- [Publications and Presentations](#)
- [Sentinel Initiative Events](#)
- [Report Finder](#)

Public Sentinel Training at FDA - Day 2 of the Tenth Annual Sentinel Initiative Public Workshop

Recordings of the presentations are available via the following links:

[Welcome, Introduction, Agenda, Learning Objectives](#)

[Review of Sentinel Capabilities](#) (*skip ahead to 14:50*)

[Propensity Score Analysis Tool](#) (*skip ahead to 28:08*)

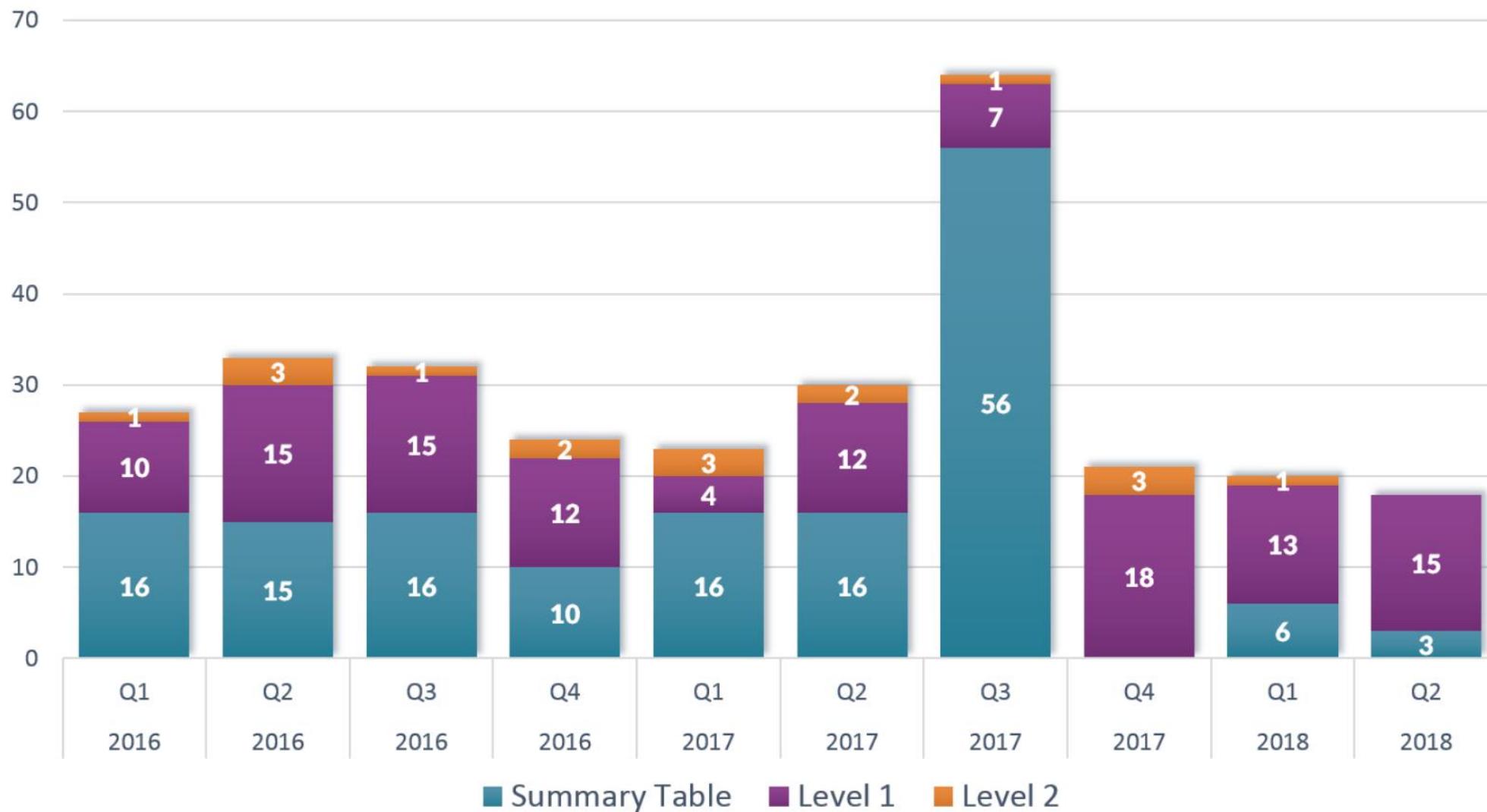
[Self-Controlled Risk Interval Tool](#)

[TreeScan Analyses](#)

[Closing Remarks](#) (*skip ahead to 57:18*)

[Sentinel Initiative Public Workshop Training Slides](#)

ARIA Analyses by Quarter (N = 292)



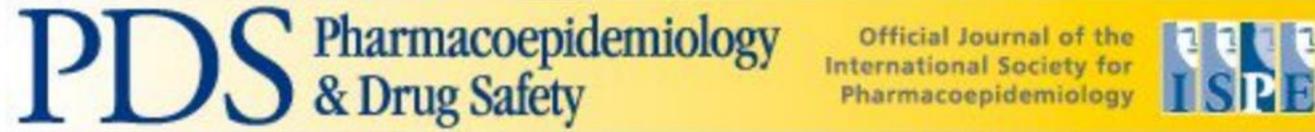
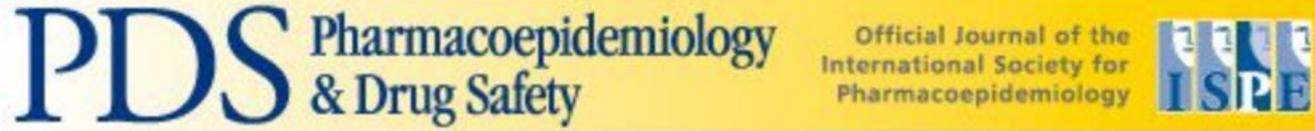
DRUGS

- [About CDER](#)
- [Assessments](#)
- [Ongoing ARIA Assessments](#)
- [How ARIA Analyses Have Been Used by FDA](#)

How ARIA Analyses Have Been Used by FDA

This page summarizes how select analyses conducted in Sentinel's [Active Risk Identification and Analysis \(ARIA\)](#) system have been used by FDA since Sentinel's official launch in February 2016. ARIA can contribute to FDA's regulatory process in a variety of ways, such as contributing evidence to support a label change, respond to a Citizens Petition, or become part of an Advisory Committee deliberation. Information from ARIA can also provide evidence that alleviates concerns about a particular safety issue and might lead FDA to determine that no regulatory action is necessary based on the available information.

Recent Published Safety Studies



ORIGINAL REPORT

Safety asses
Administrati

Original Investigation

ONLINE FIRST

ORIGINAL REPORT

October 1, 2018

Joshua J. Gagne ,
Sengwee Toh

**Prospective surveillance Association of Risk for Venous
Food and Drug Admini Thromboembolism With Use of Low-Dose
Extended- and Continuous-Cycle Combined
Oral Contraceptives**
A Safety Study Using the Sentinel Distributed
Database

Elizabeth A. Chrischilles , Joshua
Azadeh Shoaibi, Marsha E. Reichn

First published: 10 January 2018

Jie Li, PhD¹; Genna Panucci, SM²; David Moeny, RPh¹; et al

» Author Affiliations

Recently Published Descriptive Work

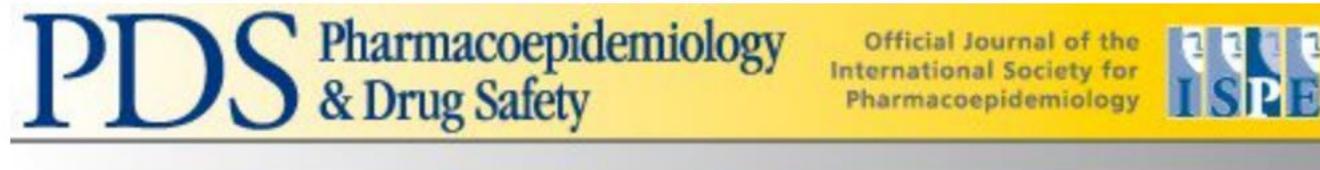


SHORT ARTICLE

Outpatient influenza s

Noelle M. Cocoro

First published: 2



ORIGINAL REPORT

Utilization of drugs with pregnancy exposure registries during pregnancy

Onyekachukwu
Kate Gelperin, I

First published:

Journal of Clinical Psychopharmacology. 38(5):505–508, OCT 2018
DOI: 10.1097/JCP.0000000000000939, PMID: 30102629
Issn Print: 0271-0749
Publication Date: 2018/10/01



 Print

Incidence of Heart Failure and Cardiomyopathy Following Initiation of Medications for Attention-Deficit/Hyperactivity Disorder: A Descriptive Study

Andrew D. Mosholder; Lockwood Taylor; Glenn Mannheim; Lisa Ortendahl; Tiffany S. Woodworth; Sengwee Toh

Relative Performance of Propensity Score Matching Strategies for Subgroup Analyses

Shirley V Wang ✉, Yinzhu Jin, Bruce Fireman, Susan Gruber, Mengdong He, Richard Wyss, HoJin Shin, Y Sara Karami, Jacqueline M Major, Sebastia Joshua J Gagne

American Journal of Epidemiology, Volume
Pages 1799–1807, <https://doi.org/10.1093/>

Published: 15 March 2018 **Article histo**

Extension of Disease Risk Score–Based Confounding Adjustments for Multiple Outcomes of Interest: An Empirical Evaluation FREE

Rishi J Desai ✉, Richard Wyss, Yinzhu Jin, Justin Bohn, Sengwee Toh, Austin Cosgrove, Adee Kennedy, Jessica Kim, Clara Kim, Rita Ouellet-Hellstrom, ... [Show more](#)

American Journal of Epidemiology, kwy130, <https://doi.org/10.1093/aje/kwy130>

Published: 26 June 2018 **Article history** ▼

Methods Development: Privacy Preserving Regression

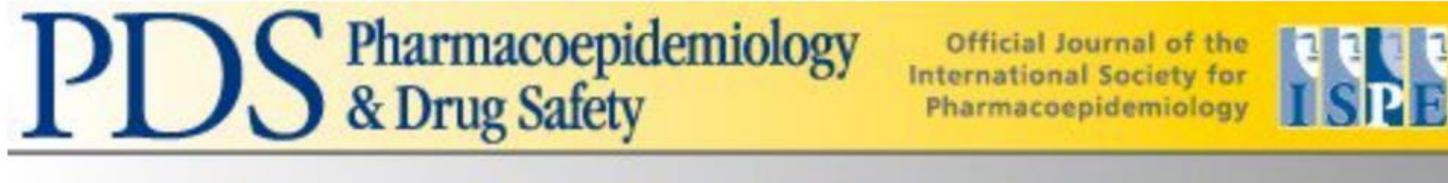
eGEMs The Journal for Electronic
AcademyHealth Health Data and Methods



Model / Framework

A Query Workflow Design to Perform

Automata
in Large E



Authors: Qou
Jessica Young

ORIGINAL REPORT

Comparison of privacy-protecting analytic and data-sharing methods: A simulation study

Kazuki Yoshida ✉, Susan Gruber, Bruce H. Fireman, Sengwee Toh

First published: 18 July 2018 | <https://doi.org/10.1002/pds.4615>

Epidemiology. 29(6):895–903, NOV 2018

DOI: 10.1097/EDE.0000000000000907, PMID: 30074538

Issn Print: 1044-3983

Publication Date: 2018/11/01

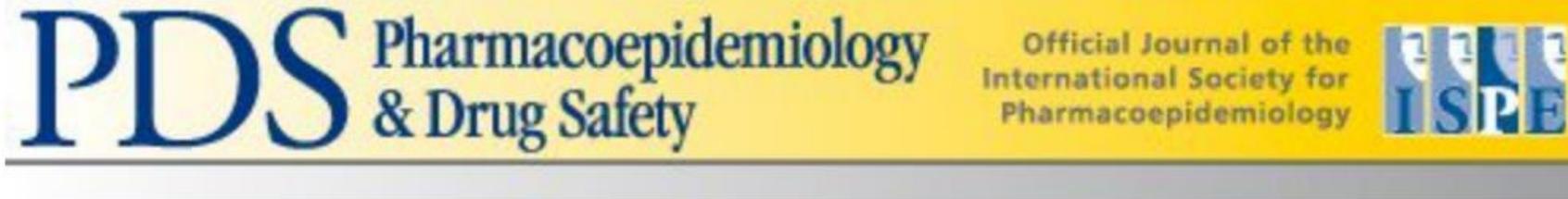


 Print

Data Mining for Adverse Drug Events With a Propensity Score-matched Tree-based Scan Statistic

Shirley V. Wang; Judith C. Maro; Elande Baro; Rima Izem; Inna Dashevsky; James R. Rogers; Michael Nguyen; Joshua J. Gagne; Elisabetta Patorno; Krista F. Huybrechts; Jacqueline M. Major; Esther Zhou; Megan Reidy; Austin Cosgrove; Sebastian Schneeweiss; Martin Kulldorff

[+ Author Information](#)

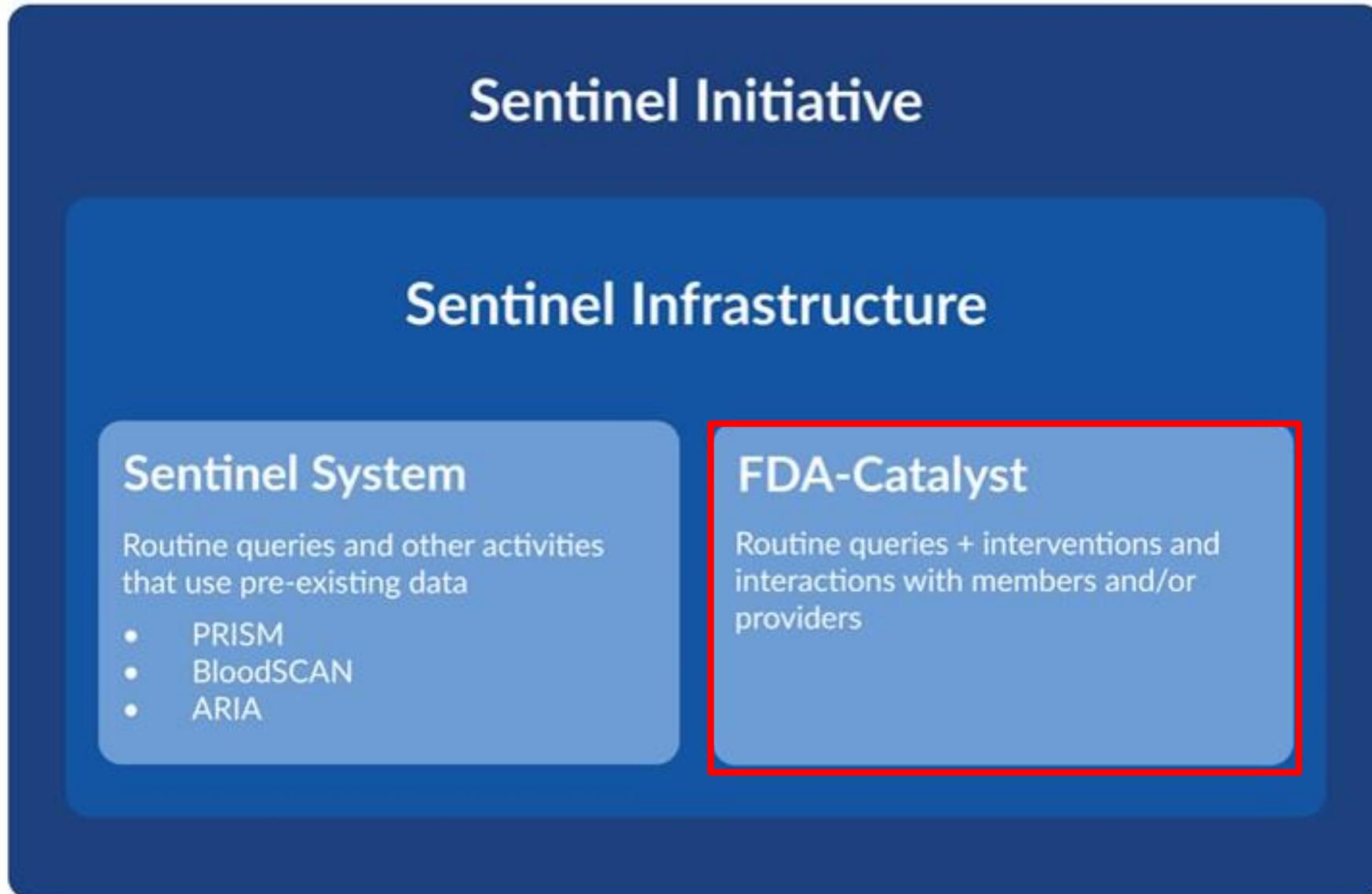


ORIGINAL REPORT

Evaluating automated approaches to anaphylaxis case classification using unstructured data from the FDA Sentinel System

Robert Ball , Sengwee Toh, Jamie Nolan, Kevin Haynes, Richard Forshee, Taxiarchis Botsis

First published: 28 August 2018 | <https://doi.org/10.1002/pds.4645>



Sentinel System

Routine queries and other activities that use pre-existing data

- PRISM
- BloodSCAN
- ARIA

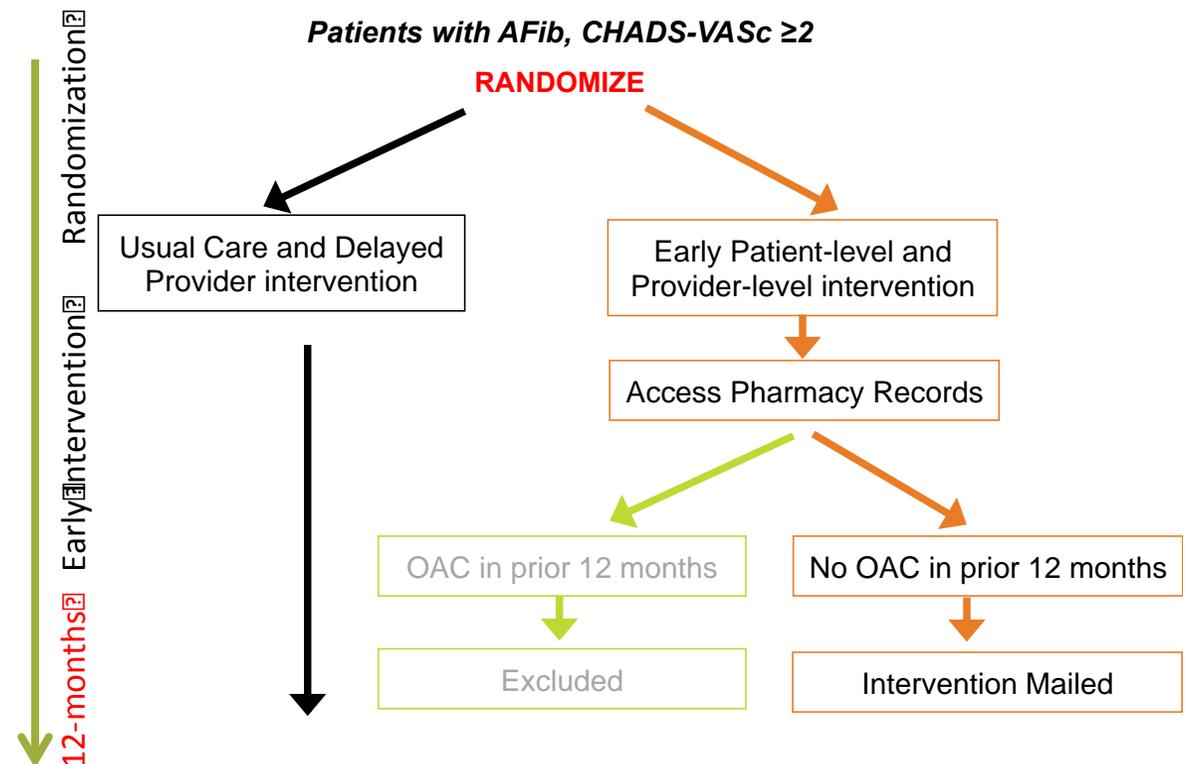
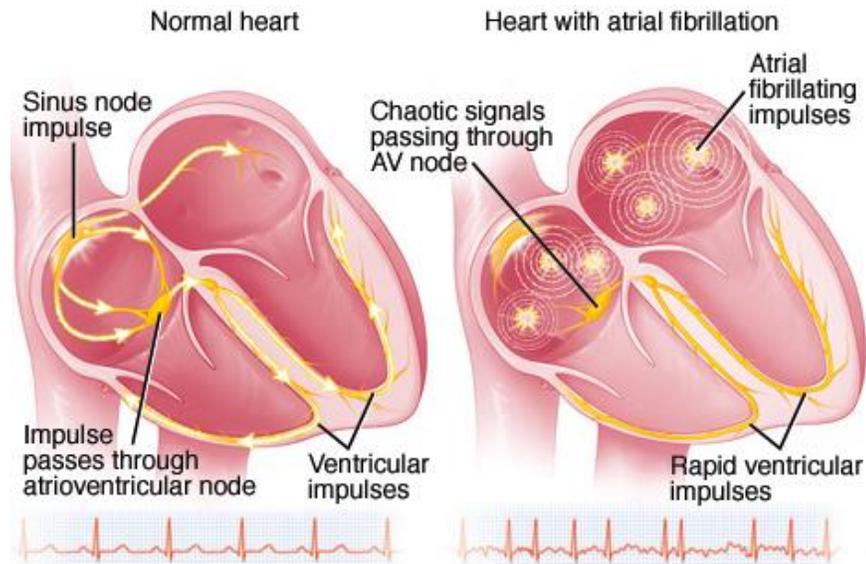
FDA-Catalyst

Routine queries + interventions and interactions with members and/or providers

Pragmatic Trial in Sentinel – IMPACT AFib

Implementation of a randomized controlled trial to improve treatment with oral AntiCoagulanTs in patients with Atrial Fibrillation

- Direct mailer to health plan members with AFib, high risk for stroke and no oral anticoagulant (OAC) treatment, and to their providers, to encourage consideration of OACs



39,230 Letters Mailed to Early Intervention Group

PROVIDER LETTER



IMPACT-AFib



[HEALTH PLAN LOGO]

Dear Provider:

As part of our effort to improve the use of oral anticoagulant medications for stroke prevention in patients with atrial fibrillation (AFib), we would like to introduce you to the IMPACT-AFib initiative. The objective of the IMPACT-AFib initiative is to increase awareness and education among patients and you. This FDA-sponsored initiative is being conducted by [HEALTH PLAN] in collaboration with researchers at Harvard and Duke.

Educational materials were sent to patient(s) who appear to have atrial fibrillation, have high stroke risk (CHA₂DS₂-VASc score ≥ 2), and have no record available to us of having filled a prescription for an anticoagulant in the past year. Please see the next page for a list of patients who received these materials.

Facts about atrial fibrillation

- Patients with AFib have a five times higher stroke risk relative to patients without AFib (*Circulation* 2011; 123(10):e269-367)
- More than two-thirds of strokes caused by AFib are preventable with anticoagulation (*Annals of internal medicine* 146.12 (2007): 857-867)
- 50% of patients with AFib and high stroke risk have not filled an anticoagulant prescription (*Circulation* 2014; 129 (15), 1568-1576)

Common misperceptions about stroke prevention

Aspirin is good enough

- Aspirin reduces stroke by < 20%, if at all, compared with 70% reduction with anticoagulation; therefore, aspirin is not sufficiently effective for stroke prevention¹

Patients with AFib are at greater risk of bleeding than stroke

- 30% of elderly patients fall in a year, but a patient would need to fall nearly every day before the risk of intracranial bleeding outweighs the benefits of anticoagulants.²
- The risk of recurrent GI bleeding averages 1.2% per year, but would have to exceed 10% before the risk of GI bleeding outweighs the benefit of anticoagulants.³

There are appropriate reasons for patients to not take an anticoagulant, including pregnancy and history of intracranial hemorrhage. A response mailer is enclosed for you to share these reasons, should they exist for your patient(s).

¹ European Heart Journal 2015; 36: 653-656 ² Arch Intern Med 1999; 159:677-685 ³ Arch Intern Med 2002; 162:541-550

MEMBER LETTER

IMPACT-AFib

[HEALTH PLAN LOGO]

IMPACT AFib address
IMPACT AFib address

[Date]
[Member Name]
[Member Address]
[Member City, St, zip]

Dear [Member Name],

You can lower your risk of stroke. Bring this letter and pocket card to your next doctor's appointment. Talk to your doctor about the use of anticoagulant medications to prevent stroke.

According to our records, you may have been diagnosed with atrial fibrillation. We know that managing your health can be a challenge, and hope this information about how to lower your risk for stroke will help.

People who have the heartbeat irregularity known as "atrial fibrillation" are at an increased risk of having a stroke.

Please visit www.IMPACT-AFib.org, to learn more about atrial fibrillation, stroke risk, and anticoagulant medications. More information about the IMPACT-AFib initiative is available by calling [XXX-XXX-XXXX] or emailing [name@duke/healthplan.ext]

If you have questions about your benefits, call the number on the back of your health plan ID card.

Talk to your doctor about anticoagulant medications.

This packet contains information about the benefits of taking anticoagulant medications, also called blood thinners, to lower your risk of having a stroke. We recommend that you bring this information packet to your next doctor's appointment. We sent similar information to your doctor.

Anticoagulant medications may not be right for all patients, but they might be right for you. Even if you have talked about this with your doctor in the past, we encourage you to have another conversation about these medications. New anticoagulant medications are safe and effective options for many patients.



Protecting your health information

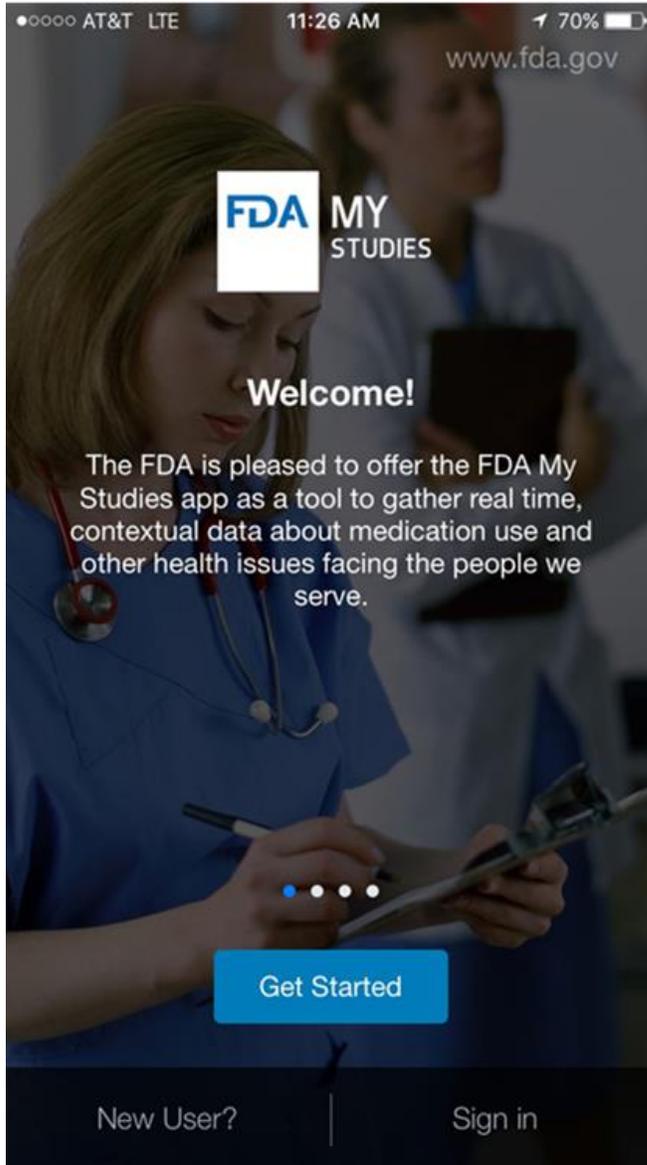
We take protecting your health information seriously. None of your health information has been shared with other health organizations. Only you and your doctor were sent this information.

Sincerely,

Chief Medical Officer
Enclosures

If you have any questions, please contact [name] at [phone #] or [email]

MyStudies App: Collecting Patient-reported Outcomes



- **Mobile App**
 - Standard frameworks - ResearchKit (iOS), ResearchStack (Android)
 - Gateway capability
- **Web-based configuration portal**
- **Secure Storage Environment**
 - FISMA complaint
 - Partitioned for distributed research
 - Responses can be downloaded in broadly compatible formats (e.g., for use in SAS, Excel, etc.)
- **Linked to the Sentinel Distributed Database**

Enroll and Consent Patients

← ELIGIBILITY Cancel

This study allows only pre-screened participants to join the study. If you are one, please enter the enrollment ID provided to you for this study.

Enter enrollment ID

Submit

← Validated! Cancel

Your ID has been validated. You are eligible to join the Study. Please click Continue to proceed to the Consent section.



Continue

●●●● AT&T LTE 2:19 PM 48% 

← Cancel



What will happen if I take part in this study?

You'll get about 1 new, 10-minute survey per week asking about any health conditions, medications take, and your pregnancy and childbirth history. You'll get a notification for each new survey. Researchers at KPWHRI will combine your responses with information in you and your baby's medical records for analysis. Your medical record will not be updated in any way and your doctor won't know you're in this study. The study is for research only and is not intended to provide medical advice.

●●●● AT&T LTE 2:17 PM 48% 

← Cancel

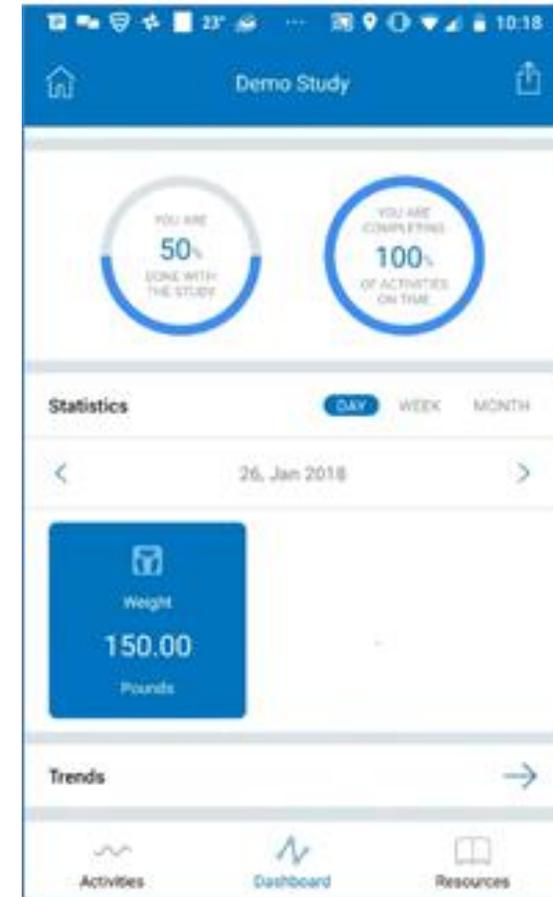
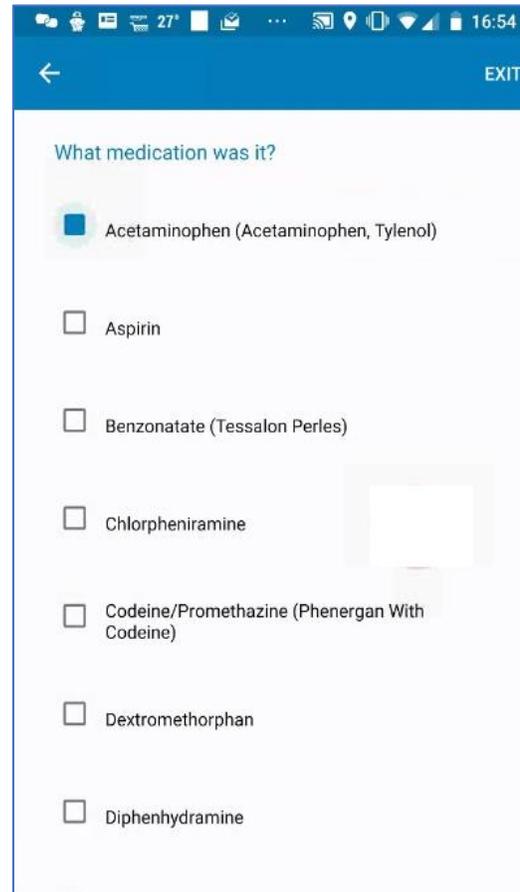
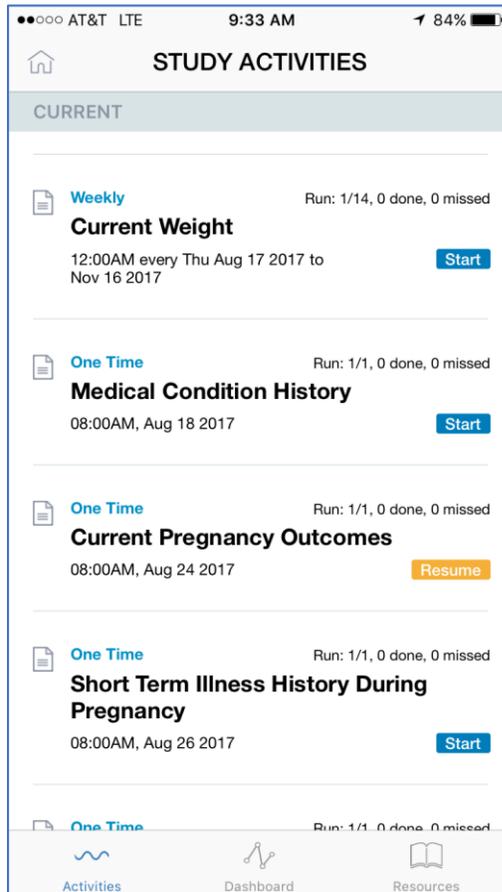


How will you protect my confidentiality?

All researchers have all completed training on how to protect your rights and privacy and your responses are only used for research. We won't use your name in study reports or add it to your survey answers from the mobile application. We won't tell your doctor whether or not you join this study or add information to your medical record.

[Learn more about how your privacy and identity are protected](#)

Next



- Data collected directly from patients
- Participants respond when they choose within the study schedule
- Study Dashboard displays progress as well as highlights from data collection

SENTINEL

- ▶ [About](#)
- ▶ [Data](#)
- ▶ [Surveillance Tools](#)
- [Methods](#)

[Submit Comment](#)

Data Infrastructure Expansion: Developing a Mother-Infant Linkage table in the Sentinel Common Data Model (SCDM)

Project Title

Data Infrastructure Expansion: Developing a Mother-Infant Linkage table in the Sentinel Common Data Model (SCDM)

Looking Ahead: New Challenges/Methods

- How to make best use of EHR data given the large degree of missingness?
 - Especially when missingness is likely Missing-Not-At-Random
- How to incorporate cutting edge machine learning techniques in a **distributed database environment** with a high degree of heterogeneity among databases?

Discussion

Acknowledgements

- Thanks to my many colleagues within the greater Sentinel Initiative including our many collaborating institutions
- Questions: info@sentinelsystem.org