



ICPE 2022 Symposium

**What do real world data validation best practices look like?
Operationalizing guidance for real-world studies intended for
decision-making**

Presented at the 38th International Conference on Pharmacoepidemiology & Therapeutic Risk Management

Sentinel experience in validating study endpoints

Darren Toh, ScD

DPM Endowed Professor

Department of Population Medicine

Harvard Medical School and Harvard Pilgrim Health Care Institute

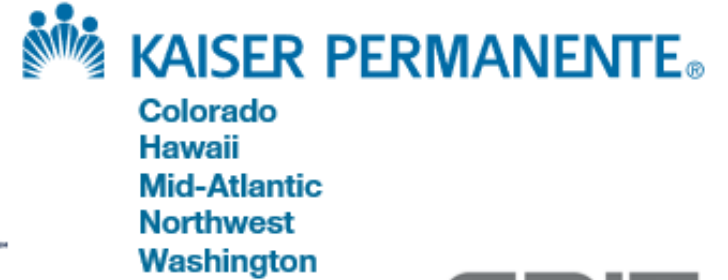
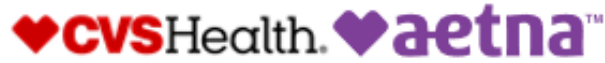
August 28, 2022

I am a consultant for Pfizer, Inc. and TriNetX, LLC

My spouse is an employee of Biogen, Inc.

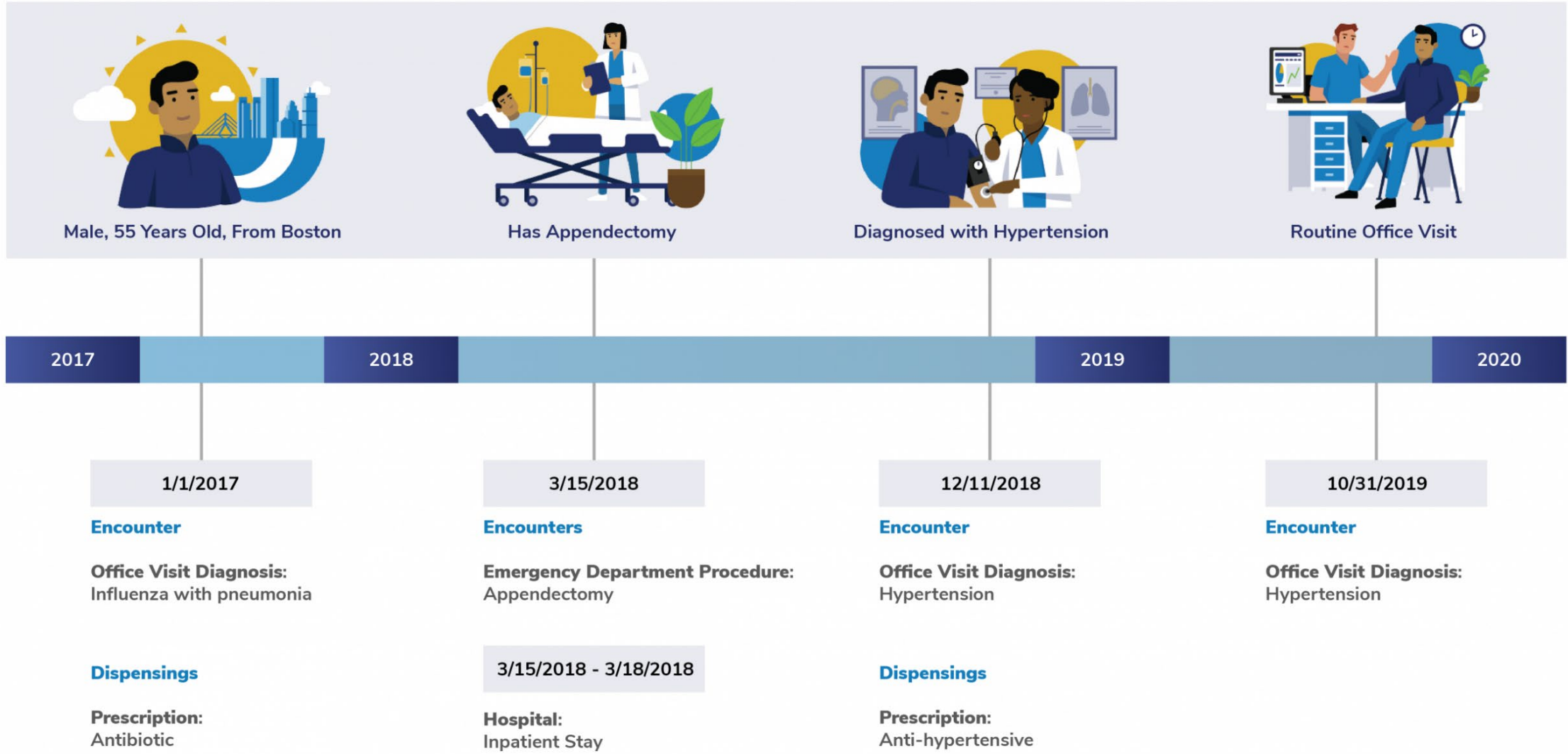
The work presented here was funded by the U.S. Food and Drug Administration. The views expressed in this presentation are mine and are not intended to convey official U.S. Food and Drug Administration policy or guidance.

DEPARTMENT OF POPULATION MEDICINE



Booz | Allen | Hamilton





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

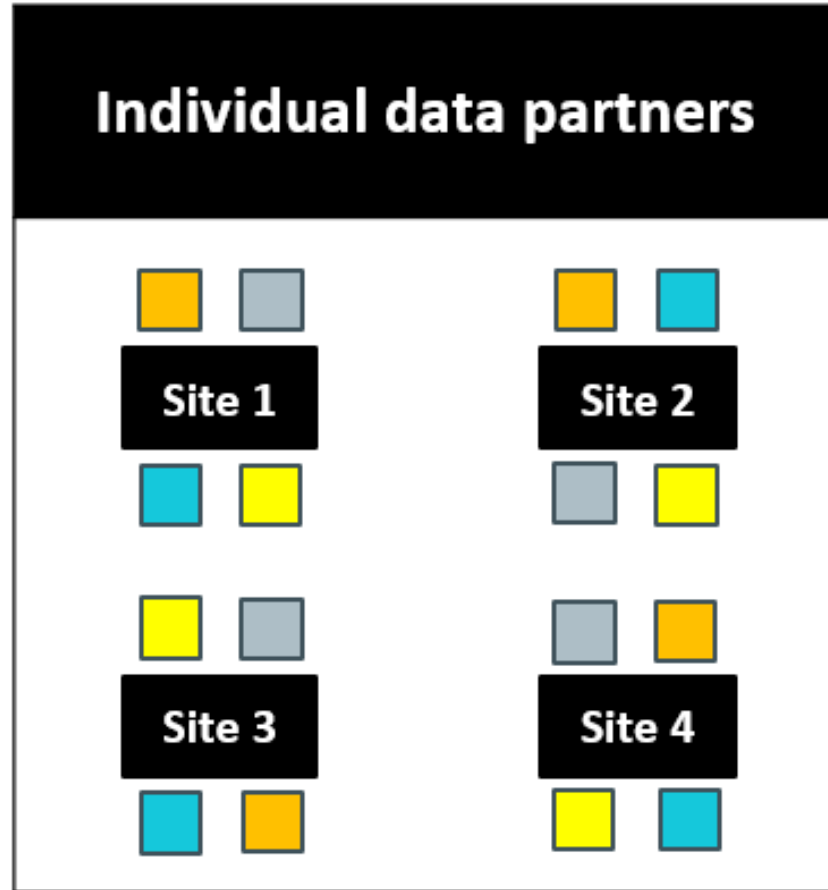
PATID	ENCOUNTERID	ADATE	PROVIDER	ENCTYPE	DX	DX_CODETYPE	PDX
PatID1	EncID1	10/18/2005	Provider1	IP	296.2		9 P
PatID1	EncID1	10/18/2005	Provider1	IP	300.02		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	305.6		9 S
PatID1	EncID1	10/18/2005	Provider1	IP	311		9 P
PatID1	EncID1	10/18/2005	Provider1	IP	401.9		9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

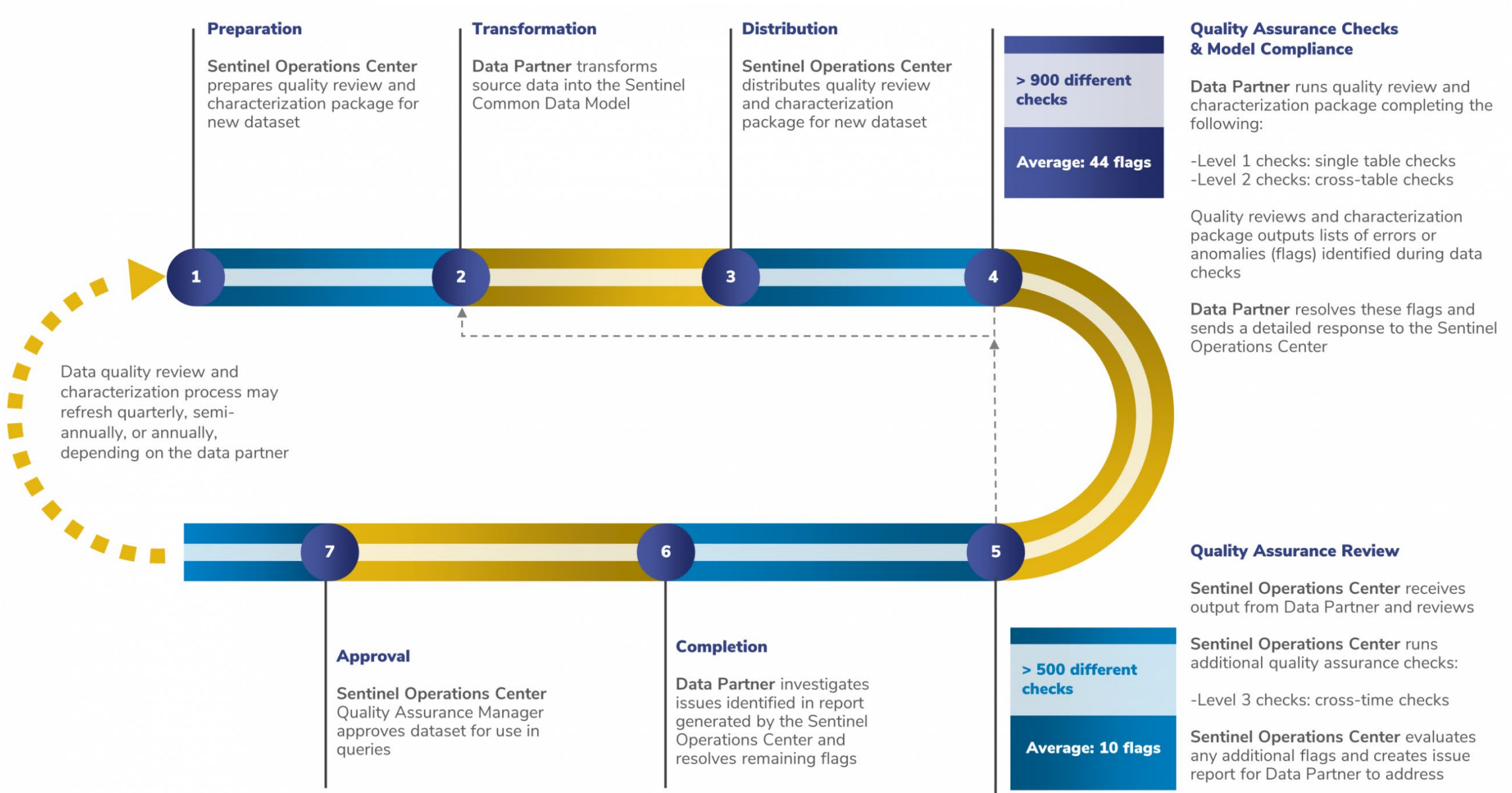
PROCEDURE

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





Guidance for Industry and FDA Staff
Best Practices for Conducting
and Reporting
Pharmacoepidemiologic Safety
Studies Using Electronic
Healthcare Data



SENTINEL DATA QUALITY ASSURANCE
PRACTICES

**COMPLIANCE WITH “GUIDANCE FOR INDUSTRY AND FDA STAFF: BEST
PRACTICES FOR CONDUCTING AND REPORTING
PHARMACOEPIDEMIOLOGIC SAFETY STUDIES USING ELECTRONIC
HEALTHCARE DATA”**

Sentinel Common Data Model

Administrative Data							Mother-Infant Linkage Data	Auxiliary Data	
Enrollment	Demographic	Dispensing	Encounter	Diagnosis	Procedure	Prescribing	Mother-Infant Linkage	Facility	Provider
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Mother ID	Facility ID	Provider ID
Enrollment Start & End Dates	Birth Date	Provider ID	Encounter ID & Type	Encounter ID & Type	Encounter ID & Type	Encounter ID	Mother Birth Date	Facility Location	Provider Specialty & Specialty Code Type
Medical Coverage	Sex	Dispensing Date	Service Date(s)	Provider ID	Provider ID	Provider ID	Encounter ID & Type		
Drug Coverage	Postal Code	Rx	Facility ID	Service Date(s)	Service Date(s)	Order Date	Mother Admission & Discharge Date		
Medical Record Availability	Race	Rx Code Type	Etc.	Diagnosis Code & Type	Procedure Code & Type	Rx	Child ID		
	Etc.	Days Supply		Principal Discharge Diagnosis	Etc.	Days Supply	Childbirth Date		
		Amount Dispensed				Rx Route of Delivery	Mother-Infant Match Method		
						Etc.	Etc.		

Registry Data			Inpatient Data		Clinical Data	Patient-Reported Measures (PRM) Data		
Death	Cause of Death	State Vaccine*	Inpatient Pharmacy	Inpatient Transfusion	Lab Result	Vital Signs	PRM Survey	PRM Survey Response
Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Patient ID	Measure ID	Patient ID
Death Date	Cause of Death	Vaccination Date	Encounter ID	Encounter ID	Result & Specimen Collection Dates	Measurement Date & Time	Survey ID	Encounter ID
Date Imputed Flag	Source	Admission Date	Rx Administration Date & Time	Transfusion Administration ID	Test Type, Immediacy & Location	Height & Weight	Question ID	Measure ID
Source	Confidence	Vaccine Code & Type	National Drug Code (NDC)	Administration Start & End Date & Time	Logical Observation Identifiers Names and Codes (LOINC®)	Diastolic & Systolic BP	Etc.	Survey ID
Confidence	Etc.	Provider	Rx ID	Transfusion Product Code	Etc.	Tobacco Use & Type		Question ID
Etc.		Etc.	Route	Blood Type		Etc.		Response Text
			Dose	Etc.				Etc.
			Etc.					

*The State Vaccine table has not been in use since SCDM v6.0.

Sentinel's Multi-Modal Response System

Claims (with Limited EHR Network)

Active Risk Identification and Analysis (ARIA)*

Sentinel Distributed Database

Merative™ MarketScan® Research Databases

- Sentinel Common Data Model
- Sentinel Analytic Tools
- Access to Medical Records within the Sentinel Distributed Database

EHR Data Aggregators

TriNetX

IBM Watson Health

- Proprietary Common Data Models
- Web-Based Query Interface & Custom Programming
- Access to Medical Records varies by Source

EHR Data Warehouses

HCA Healthcare

Veradigm

- Data Warehouses for Multiple Healthcare Organizations in a System
- Custom Programming
- Access to Medical Records

EHR Networks

PCORnet

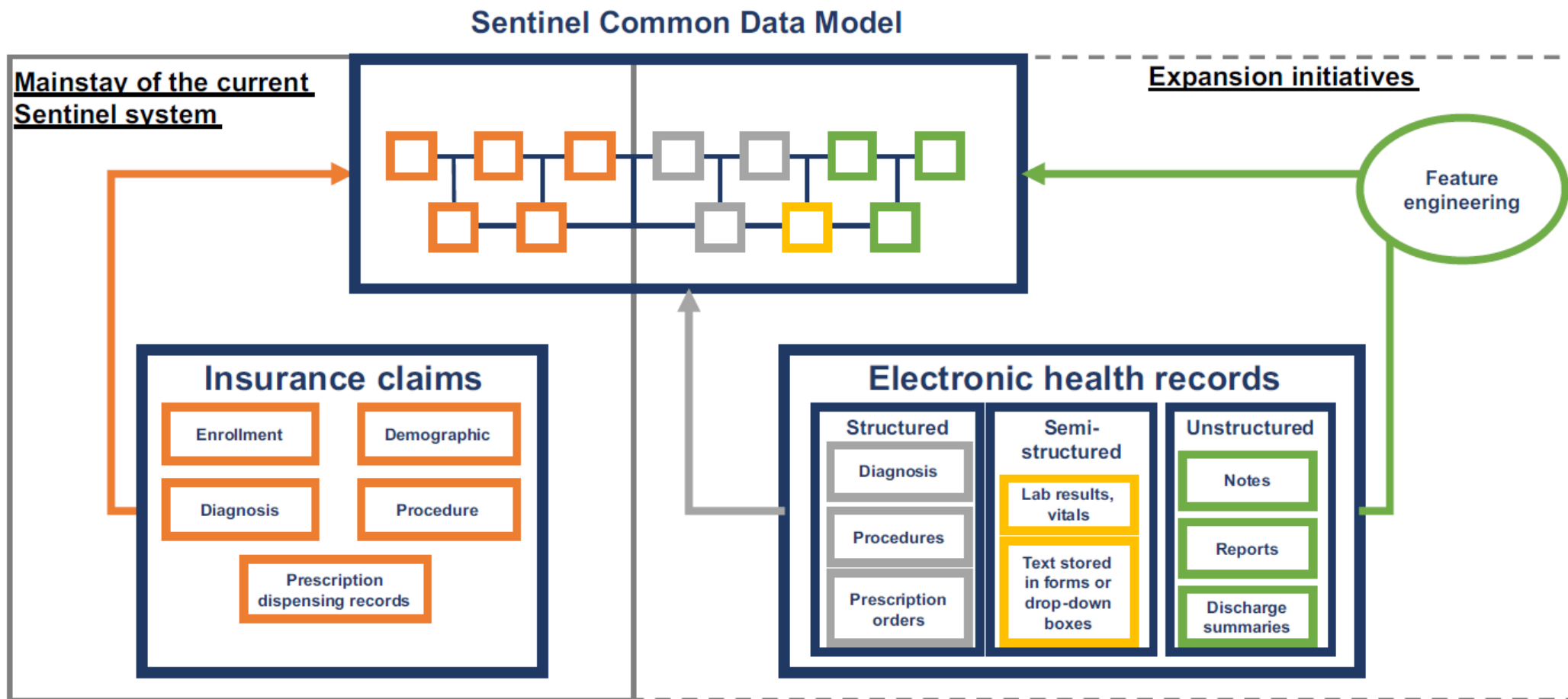
- PCORnet Common Data Model
- PCORnet Analytic Tools
- Access to Medical Records

***Note:** The Active Risk Identification and Analysis (ARIA) System is comprised of the Sentinel Distributed Database, the Sentinel Common Data Model, and Sentinel analytic tools.

Broadening the reach of the FDA Sentinel system: A roadmap for integrating electronic health record data in a causal analysis framework

Rishi J. Desai^{1,✉}, Michael E. Matheny^{1,2}, Kevin Johnson², Keith Marsolo³, Lesley H. Curtis³, Jennifer C. Nelson⁴, Patrick J. Heagerty⁵, Judith Maro⁶, Jeffery Brown⁶, Sengwee Toh⁵, Michael Nguyen⁷, Robert Ball⁷, Gerald Dal Pan⁷, Shirley V. Wang¹, Joshua J. Gagne^{1,8} and Sebastian Schneeweiss¹

npj Digital Medicine (2021) 170



This presentation will focus on validation of study endpoints, although Sentinel had also validated other study variables (e.g., exposures)

User-defined
claims-based
outcome
algorithms

Data-driven
claims-based
outcome
algorithms

User-defined
structured EHR-
based outcome
algorithms

User-defined
structured and
unstructured
EHR-based
outcome
algorithms

Data-driven
structured and
unstructured
EHR-based
outcome
algorithms

Increasing use of unstructured data

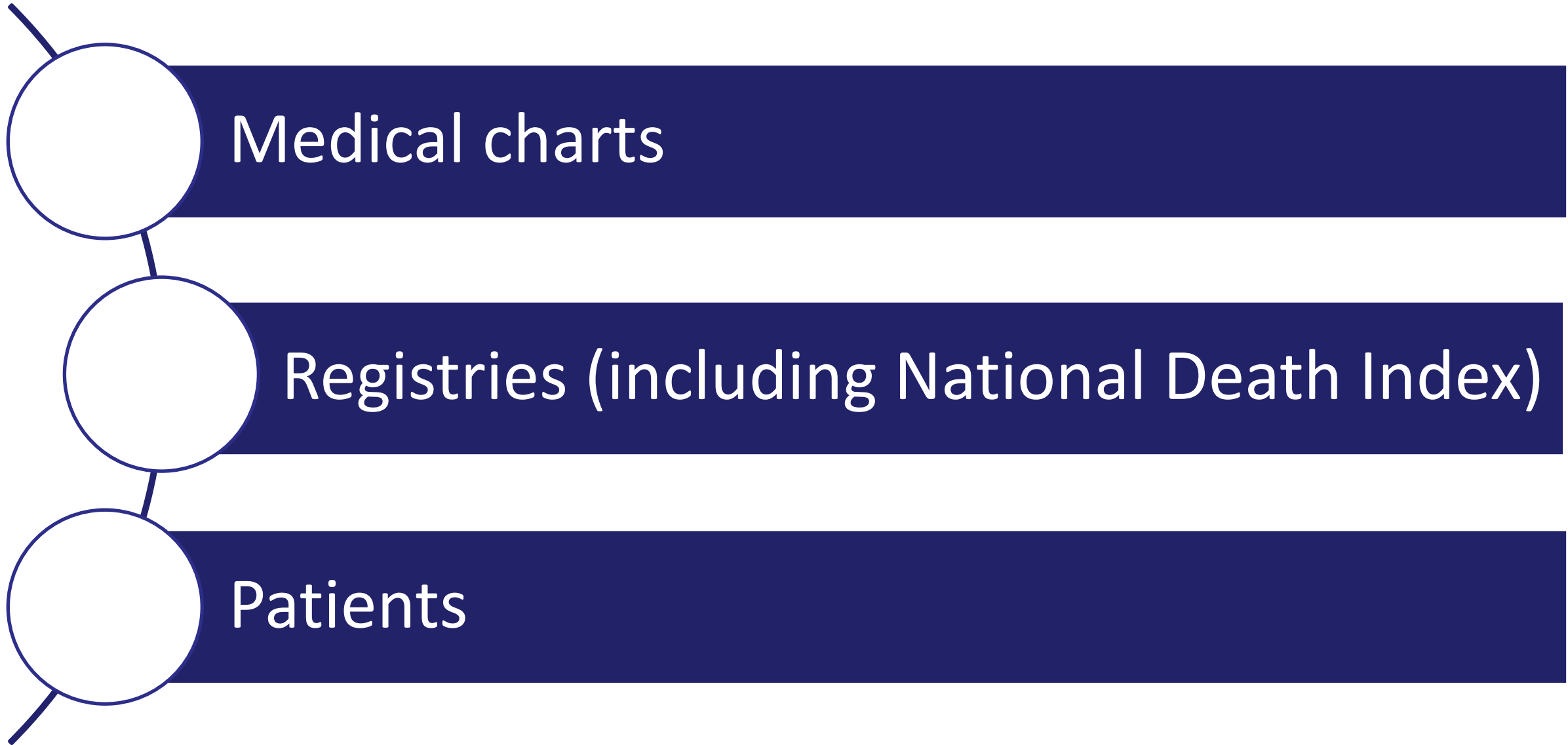
Application of algorithms from published literature

Validation of select potential cases from populations outside of the study

Validation of select potential cases from study cohort

Validation of all potential cases from study cohort

Toward complete verification



ORIGINAL REPORT

A systematic review of validated methods for identifying cerebrovascular accident or transient ischemic attack using administrative data

Susan E. Andrade*, Leslie R. Harrold, Jennifer Tjia, Sarah L. Cutrona, Jane S. Saczynski, Katherine S. Dodd, Robert J. Goldberg and Jerry H. Gurwitz

Meyers Primary Care Institute (Reliant Medical Group, Fallon Community Health Plan, and University of Massachusetts Medical School), Worcester, MA, USA

ORIGINAL REPORT

A systematic review of validated methods for identifying suicide or suicidal ideation using administrative or claims data

James T. Walkup^{1*}, Lisa Townsend², Stephen Crystal^{2,3} and Mark Olfson⁴

¹*Institute for Health, Health Care Policy and Aging Research, Rutgers University, New Brunswick, NJ, USA*

²*School of Social Work, Rutgers University, New Brunswick, NJ, USA*

³*Chronic Disease Management and Outcomes, Center for Health Services Research on Pharmacotherapy, New Brunswick, NJ, USA*

⁴*Department of Psychiatry, Columbia University, New York, New York, USA*

ORIGINAL REPORT

A systematic review of validated methods for identifying heart failure using administrative data

Jane S. Saczynski*, Susan E. Andrade, Leslie R. Harrold, Jennifer Tjia, Sarah L. Cutrona, Katherine S. Dodd, Robert J. Goldberg and Jerry H. Gurwitz

Division of Geriatric Medicine and Meyers Primary Care Institute, University of Massachusetts Medical School, Worcester, MA, USA

ORIGINAL REPORT

A systematic review of validated methods for identifying pancreatitis using administrative data

Kevin Moores^{1,2*}, Bradley Gilchrist^{1,2}, Ryan Carnahan³ and Thad Abrams^{4,5}

¹*Division of Drug Information Service, The University of Iowa College of Pharmacy, Iowa City, IA, USA*

²*Iowa Drug Information Service, The University of Iowa College of Pharmacy, Iowa City, IA, USA*

³*Department of Epidemiology, University of Iowa College of Public Health, Iowa City, IA, USA*

⁴*Department of Internal Medicine, Division of General Internal Medicine, University of Iowa Carver College of Medicine, Iowa City, IA, USA*

⁵*Center for Implementation of Innovative Strategies in Practice, Iowa City Veterans Affairs Medical Center, Iowa City, IA, USA*

ORIGINAL REPORT

Validation of acute myocardial infarction in the Food and Drug Administration's Mini-Sentinel program

Sarah L. Cutrona^{1*}, Sengwee Toh², Aarthi Iyer², Sarah Foy¹, Gregory W. Daniel⁵, Vinit P. Nair⁶, Daniel Ng⁷, Melissa G. Butler⁸, Denise Boudreau⁹, Susan Forrow², Robert Goldberg¹, Joel Gore³, David McManus³, Judith A. Racoosin⁴ and Jerry H. Gurwitz¹

ORIGINAL REPORT

Validation of anaphylaxis in the Food and Drug Administration's Mini-Sentinel

Kathleen E. Walsh^{1*}, Sarah L. Cutrona^{1,2}, Sarah Foy¹, Meghan A. Baker^{3,4}, Susan Forrow⁴, Azadeh Shoaibi⁵, Pamala A. Pawloski⁶, Michelle Conroy⁷, Andrew M. Fine⁸, Lise E. Nigrovic⁸, Nandini Selvam⁹, Mano S. Selvan¹⁰, William O. Cooper¹¹ and Susan Andrade¹

ORIGINAL REPORT

Validity of diagnostic codes to identify cases of severe acute liver injury in the U.S. Food and Drug Administration's Mini-Sentinel Distributed Database

Vincent Lo Re III^{1,2*}, Kevin Haynes², David Goldberg^{2,3}, Kimberly A. Forde^{2,3}, Dena M. Carbonari², Kimberly B. F. Leidl², Sean Hennessy², K. Rajender Reddy³, Pamala A. Pawloski⁴, Gregory W. Daniel^{5,6}, T. Craig Cheetham⁷, Aarthi Iyer⁸, Kara O. Coughlin⁸, Sengwee Toh⁸, Denise M. Boudreau⁹, Nandini Selvam⁵, William O. Cooper¹⁰, Mano S. Selvan¹¹, Jeffrey J. VanWormer¹², Mark I. Avigan¹³, Monika Houstoun¹³, Gwen L. Zornberg¹³, Judith A. Racoosin¹³ and Azadeh Shoaibi¹³



VALIDATION OF ACUTE KIDNEY INJURY CASES IN THE MINI-SENTINEL DISTRIBUTED DATABASE

Prepared by: Uptal D. Patel, MD,^{1,2} N. Chantelle Hardy, MPH,² David H. Smith, RPh, PhD,³ Jerry H. Gurwitz, MD,⁴ Chi-yuan Hsu, MD, MSc,⁵ Chirag R. Parikh, MD, PhD,⁶ Steven M. Brunelli, MD, MSCE,⁷ Meghan Baker, MD, ScD,⁸ Susan Forrow, BA,⁸ Carly Comins, BS,⁸ Denise M. Boudreau, PhD, RPh,⁹ Chunfu Liu, ScD,¹⁰ Pamala A. Pawloski, PharmD,¹¹ Nandini Selvam, PhD, MPH,¹⁰ Mano S. Selvan, PhD,¹² Shannon Stratton, BS,¹³ Jeffrey J. VanWormer, PhD,¹⁴ George Aggrey, MD, MPH,¹⁵ Melanie Blank, MD,¹⁵ Patrick Archdeacon, MD¹⁵

Validation of an electronic algorithm for Hodgkin and non-Hodgkin lymphoma in ICD-10-CM

Mara M. Epstein^{1,2} | Sarah K. Dutcher³ | Judith C. Maro⁴ |
Cassandra Saphirak^{1,2} | Sandra DeLuccia⁴ | Muthalagu Ramanathan⁵ |
Tejaswini Dhawale⁶ | Sonali Harchandani⁵ | Christopher Delude² | Laura Hou⁴ |
Autumn Gertz⁴ | Nina DiNunzio⁴ | Cheryl N. McMahon-Walraven⁷ |
Mano S. Selvan⁸ | Justin Vigeant⁴ | David V. Cole⁴ | Kira Leishear³ |
Jerry H. Gurwitz^{1,2} | Susan Andrade^{1,2} | Noelle M. Cocoros⁴

Validity of ICD-10-CM diagnoses to identify hospitalizations for serious infections among patients treated with biologic therapies

Vincent Lo Re III^{1,2} | Dena M. Carbonari² | Jerry Jacob¹ | William R. Short¹ |
Charles E. Leonard² | Jennifer G. Lyons³ | Adele Kennedy³ | Jolene Damon³ |
Nicole Haug³ | Esther H. Zhou⁴ | David J. Graham⁴ |
Cheryl N. McMahon-Walraven⁵ | Lauren E. Parlett⁶ | Vinit Nair⁷ | Mano Selvan⁷ |
Yunping Zhou⁷ | Gaia Pocobelli⁸ | Judith C. Maro³ | Michael D. Nguyen⁴

Validation of an ICD-10-based algorithm to identify stillbirth in the Sentinel System

Susan E. Andrade¹ | Mayura Shinde² | Tiffany A. Moore Simas³ | Steven T. Bird⁴ |
Justin Bohn² | Kevin Haynes⁵ | Lockwood G. Taylor⁴ | Julianne R. Lauring³ |
Erin Longley⁶ | Cheryl N. McMahon-Walraven⁷ | Connie M. Trinacty⁸ |
Cassandra Saphirak¹ | Christopher Delude¹ | Sandra DeLuccia² | Tancy Zhang² |
David V. Cole² | Nina DiNunzio² | Autumn Gertz² | Elnara Fazio-Eynullayeva² |
Danijela Stojanovic⁴

Validation of diagnosis codes to identify hospitalized COVID-19 patients in health care claims data

Sheryl A. Kluberg¹ | Laura Hou¹ | Sarah K. Dutcher² | Monisha Billings² |
Brian Kit² | Sengwee Toh¹ | Sascha Dublin³ | Kevin Haynes⁴ |
Annemarie Kline⁵ | Mahesh Maiyani⁶ | Pamala A. Pawloski⁷ | Eric S. Watson⁸ |
Noelle M. Cocoros¹

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

FEBRUARY 6, 2014

VOL. 370 NO. 6

Intussusception Risk after Rotavirus Vaccination in U.S. Infants

W. Katherine Yih, Ph.D., M.P.H., Tracy A. Lieu, M.D., M.P.H., Martin Kulldorff, Ph.D., David Martin, M.D., M.P.H., Cheryl N. McMahon-Walraven, M.S.W., Ph.D., Richard Platt, M.D., Nandini Selvam, Ph.D., M.P.H., Mano Selvan, Ph.D., Grace M. Lee, M.D., M.P.H., and Michael Nguyen, M.D.

Vaccine 37 (2019) 4172–4176

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine



ELSEVIER



Validation of febrile seizures identified in the Sentinel Post-Licensure Rapid Immunization Safety Monitoring Program

Alison Tse Kawai^{a,*,1}, David Martin^b, Sarah E. Henrickson^c, Ashleigh Goff^a, Megan Reidy^a, Diana Santiago^a, Nandini Selvam^d, Mano Selvan^e, Cheryl McMahon-Walraven^f, Grace M. Lee^{a,8}

Journal of Thrombosis and Thrombolysis (2022) 53:264–272
<https://doi.org/10.1007/s11239-021-02610-4>

Self-controlled assessment of thromboembolic event (TEE) risk following intravenous immune globulin (IGIV) in the U.S. (2006–2012)

Eric M. Ammann^{1,7} · Elizabeth A. Chrischilles¹ · Ryan M. Carnahan¹ · Bruce Fireman² · Candace C. Fuller³ · Marin L. Schweizer¹ · Crystal Garcia¹ · Madelyn Pimentel³ · Charles E. Leonard⁴ · Meghan A. Baker³ · Adam Cuker⁴ · Enrique C. Leira^{1,5} · Jennifer G. Robinson¹ · Scott K. Winiacki^{6,8}

SENTINEL ASSESSMENT REPORT

PARENTERAL IRON AND ANAPHYLACTOID REACTIONS

Prepared by: Kathleen Walsh, MD, MSc;¹ Susan Andrade, ScD;² Noelle Cocoros, DSc, MPH;³ Susan Farrow, BA;³ Robert Kane, MD;⁴ Niko Lehman-White, BA;³ Mark Levenson, PhD;⁴ Lingling Li, PhD;³ Marsha Reichman, PhD;⁴ Qin Ryan, MD, PhD;⁴ Ryan Saliga, BA;³ Gayathri Sridhar, MBBS, MPH, PhD;⁵ Joann Wagner, MSW;² Diqiong Xie, MS, PhD;³ Cunlin Wang, MD, PhD⁴



Received: 31 July 2020 | Revised: 20 November 2020 | Accepted: 14 December 2020

DOI: 10.1111/trf.16251

TRANSFUSION COMPLICATIONS


TRANSFUSION

Medical chart validation of inpatient diagnosis codes for transfusion-related acute lung injury 2013-2015

Candace C. Fuller¹  | Vinod E. Nambudiri² | Caren Spencer-Smith³ |
Lesley H. Curtis⁴ | Mayura Shinde¹ | Austin Cosgrove¹ | Margaret Johnson¹ |
Jason Hickok³ | Stacey Honda⁵ | Heba Ismail⁶ | Richard Max Kaufman⁷  |
Adee Kennedy¹ | Karla M. Miller³ | David J. Mohlman⁸ | Russell E. Poland^{1,3} |
Robert Rosofsky⁹ | Kimberly Smith³ | Salim R. Surani¹⁰ | Meghan A. Baker¹

ORIGINAL REPORT

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

Robert Ball¹  | Sengwee Toh²  | Jamie Nolan² | Kevin Haynes³ | Richard Forshee⁴ | Taxiarchis Botsis⁴

Pharmacoepidemiol Drug Saf. 2018;**27**:1077–1084.

Journal of the American Medical Informatics Association, 28(7), 2021, 1507–1517

doi: 10.1093/jamia/ocab036


Advance Access Publication Date: 13 March 2021

Research and Applications



Research and Applications

Electronic phenotyping of health outcomes of interest using a linked claims-electronic health record database: Findings from a machine learning pilot project

Teresa B. Gibson ^{1*} Michael D. Nguyen,² Timothy Burrell,¹ Frank Yoon,¹ Jenna Wong,³ Sai Dharmarajan,⁴ Rita Ouellet-Hellstrom,⁵ Wei Hua,² Yong Ma,⁶ Elande Baro,⁷ Sarah Bloemers,¹ Cory Pack,¹ Adele Kennedy,³ Sengwee Toh,³ and Robert Ball⁸

Use of a mobile app to capture supplemental health information during pregnancy: Implications for clinical research

Claire W. Rothschild¹ | Sascha Dublin^{1,2} | Jeffrey S. Brown^{3,4} | Predrag Klasnja² | Chayim Herzig-Marx^{3,4} | Juliane S. Reynolds^{3,4} | Zachary Wyner^{3,4} | Christina Chambers⁵ | David Martin⁶

JMIR RESEARCH PROTOCOLS

Fuller et al

Protocol

Developing a Standardized and Reusable Method to Link Distributed Health Plan Databases to the National Death Index: Methods Development Study Protocol

Candace C Fuller¹, MPH, PhD; Wei Hua², MSc, MHS, MD, PhD; Charles E Leonard³, PharmD, MSCE; Andrew Mosholder², MD, MPH; Ryan Carnahan⁴, PharmD, MS; Sarah Dutcher², PhD, MS; Katelyn King¹, BA; Andrew B Petrone¹, MPH; Robert Rosofsky⁵, MA; Laura A Shockro¹, BA; Jessica Young¹, PhD; Jea Young Min⁶, PharmD, MPH, PhD; Ingrid Binswanger⁷, MD, MPH, MS; Denise Boudreau⁸, RPh, PhD, MS; Marie R Griffin⁶, MD, MPH; Margaret A Adgent⁶, MSPH, PhD; Jennifer Kuntz⁹, MS, PhD; Cheryl McMahill-Walraven¹⁰, MSW, PhD; Pamala A Pawloski¹¹, PharmD; Robert Ball², MD, MPH, ScM; Sengwee Toh¹, ScD

VALIDATING TYPE 1 AND TYPE 2 DIABETES MELLITUS IN THE MINI-SENTINEL DISTRIBUTED DATABASE USING THE SURVEILLANCE, PREVENTION, AND MANAGEMENT OF DIABETES MELLITUS (SUPREME-DM) DATALINK

Prepared by: Marsha A Raebel, PharmD,^{1,2} Emily B Schroeder, MD, PhD,^{1,3} Glenn Goodrich, MS,¹ Andrea R Paolino, MA,¹ William Troy Donahoo,^{3,4} MD, Candace Fuller, PhD,⁵ Carlos Bell, MPH,⁶ Gregory A Nichols, PhD,⁷ Julie Schmittiel, PhD,⁸ Katherine M. Newton, PhD,⁹ Pamala Pawloski, PharmD,¹⁰ Azadeh Shoaibi, PhD, MHS, MS⁶

Augmenting Date of Death & Cause of Death Ascertainment in Sentinel

Details

Additional Information

Contributors

Date Posted: Wednesday, March 16, 2022

Status: IN PROGRESS

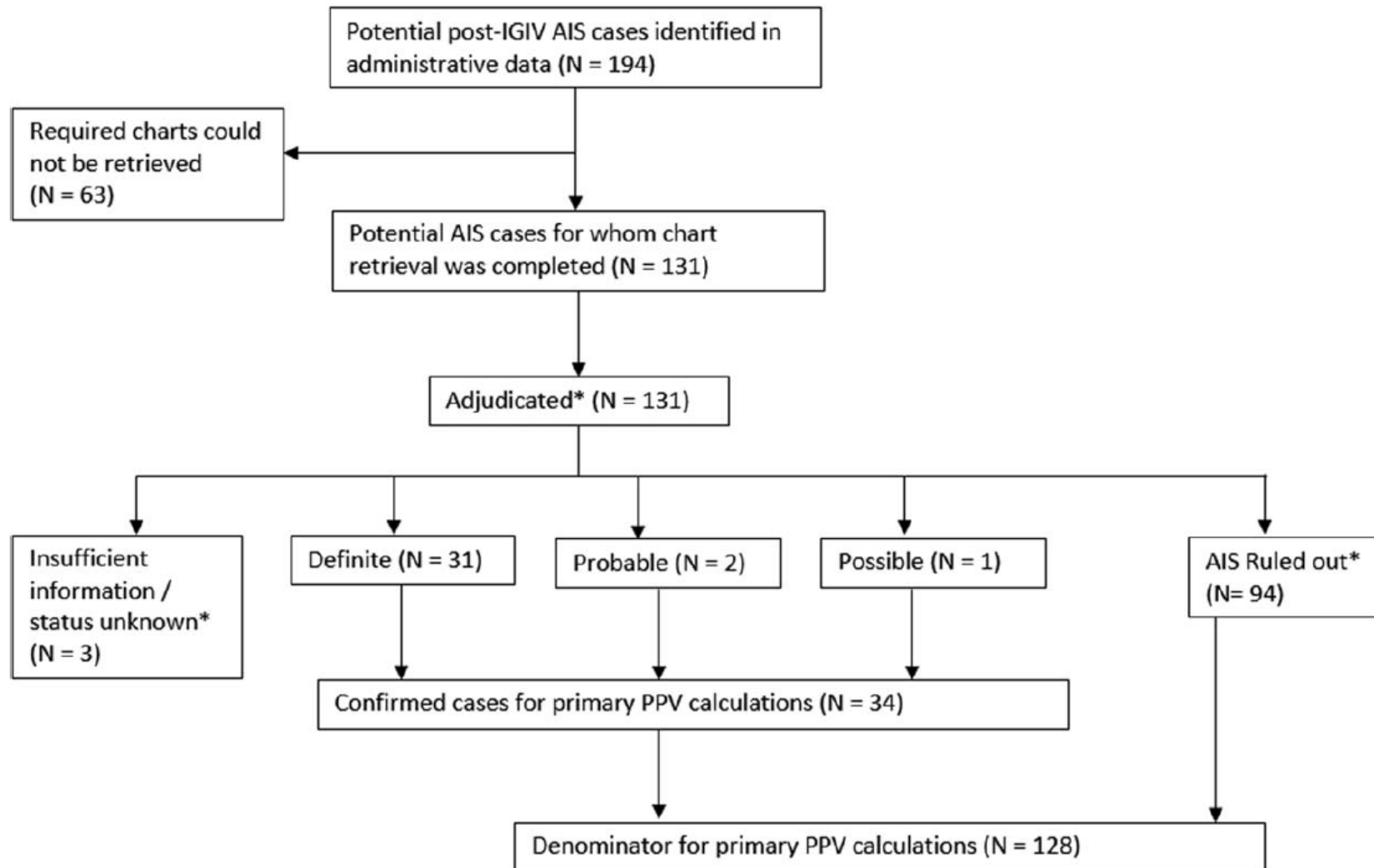


Figure 1. Disposition of potential acute ischemic stroke (AIS) cases identified in the SDD. In limited circumstances (see methods section), a potential case could be ruled out or classified as uninformative based on the judgment of the abstractor and was not physician-adjudicated. For AIS, this included 2 cases evaluated as insufficient information/status unknown and 3 cases as no AIS. *AIS=acute ischemic stroke, IGIV=intravenous immune globulin, PPV=positive predictive value.

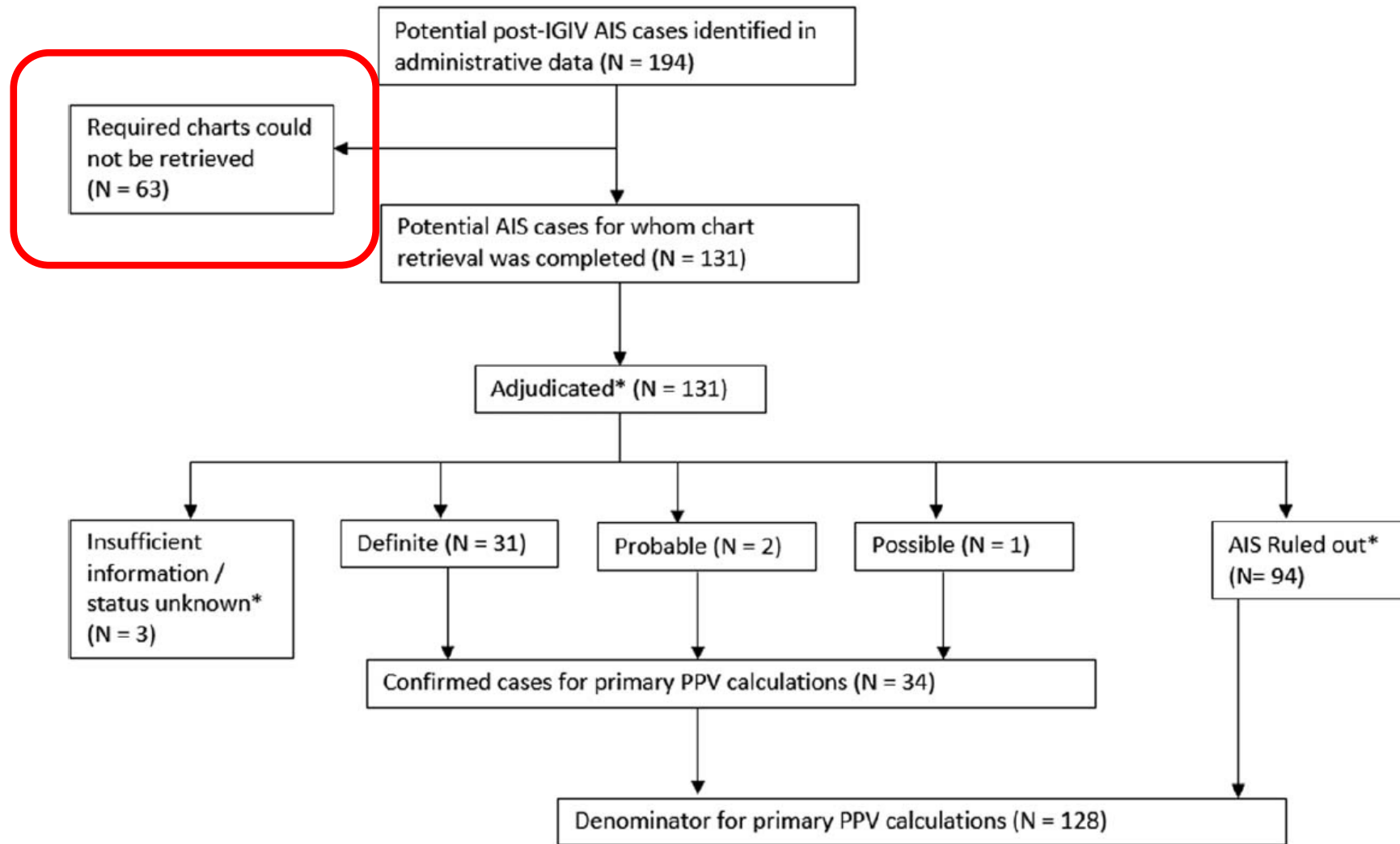


Figure 1. Disposition of potential acute ischemic stroke (AIS) cases identified in the SDD. In limited circumstances (see methods section), a potential case could be ruled out or classified as uninformative based on the judgment of the abstractor and was not physician-adjudicated. For AIS, this included 2 cases evaluated as insufficient information/status unknown and 3 cases as no AIS. *AIS=acute ischemic stroke, IGIV=intravenous immune globulin, PPV=positive predictive value.

Even if the goal is complete verification, it is often not possible to retrieve charts from all potential cases

And most studies do not confirm the case status of (all) individuals classified by the algorithm as not having the outcome events

As outcome algorithms are increasingly being defined within data sources conventionally used as a **reference standard (e.g., unstructured EHRs), it may become increasingly challenging to clearly identify when development ends and validation begins**

The question is
When does it matter?

Sentinel experience in validating study endpoints

Darren Toh, ScD



<https://www.sentinelinitiative.org/>



darren_toh@harvardpilgrim.org



[@darrentoh_epi](#)