Drug Repurposing Roadmap that Optimizes Patient Impact through Collaboration

Nov 15, 2022

CDRC
CURE DRUG REPURPOSING COLLABORATORY

CURE ID
Challenging cases... New approaches

An Introduction

- "Smitty" Heavner
  - Pronouns: He/They
- Registered Nurse
  - 10 years, specialized in ED
- PhD in Evaluation Science
- Scientific Director
  - CURE Drug Repurposing Collaboratory
- Adjunct Faculty
  - Clemson University – Public Health Sciences
  - University of South Carolina School of Medicine Greenville

Scan for contact

Disclosures

- CDRC is a public-private partnership with FDA and NIH National Center for Advancing Translational Science
- Critical Path Institute is supported by the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS) and is 54.2% funded by the FDA/HHS, totaling $13,239,950, and 45.8% funded by non-government source(s), totaling $11,196,634.
- The views expressed are those of the presenter and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS NIH or the U.S. Government.

Outline

- CDRC Overview
- Drug Repurposing
- Real-World Data
- The Edge Tool
  - Use Case: COVID-19
- Real-World Evidence
- Future Directions

CDRC Team

Marco Schito
Executive Director

Smith Heavner
Scientific Director

Chandler Birch
Sr Project Manager

Pamela Dasher
Project Manager

Stacey Coe
Sr Clin Res Coordinator

Kerry Howard
Post Doc Fellow

Claire Bassetti
Comms & Patient
Engagement Manager

Kitty Bogy
Sr Project
Coordinator

Jennifer Stephens
Project Coordinator

Interns
Zeba Idris
Nalini Oliver
Bisma Ali

Cross Functional Groups

DCC

Keith Scollick
Cloud Platform

Will Roddy
Sr Data Engineer

Quant Med

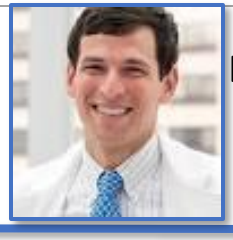
Jagdeep Podichetty
Director Predictive

Wesley Anderson
Post Doc Fellow

LotosNile

Roxan Olivas
Communications

CDRC Team Collaborators



David Fajgenbaum
CDRC Co-director

FDA OMP

Leonard Sacks
(Team lead)

Heather Stone
(Policy Analyst)

Milli Duggal
(Special Pop)

Shira Strongin
(Rare Dis)

Maya
Younoszai
(PE Grant)

Reema Charles
(Infectious Dis)

Tahsin Farid
(Clinical Fellow)

FDA OCE

Julie Schneider
(Team lead)

Leslie Doros
(Med Officer)

Sonia Singh
(Med Officer)

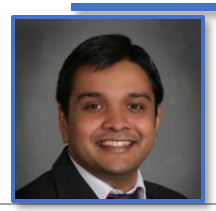
NIH NCATS

Ewy Mathe
(Informaticist)

Tim Sheils
(App Developer)

Hyun Cho
(App Designer)

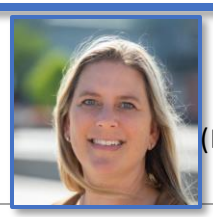
Consultants



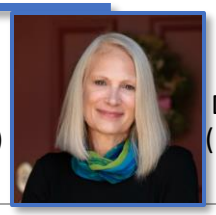
Raghav Tirupathi
(ID doc)

Nitin Gupta
(ID doc)

Ruth Kurtycz
(Statistician)

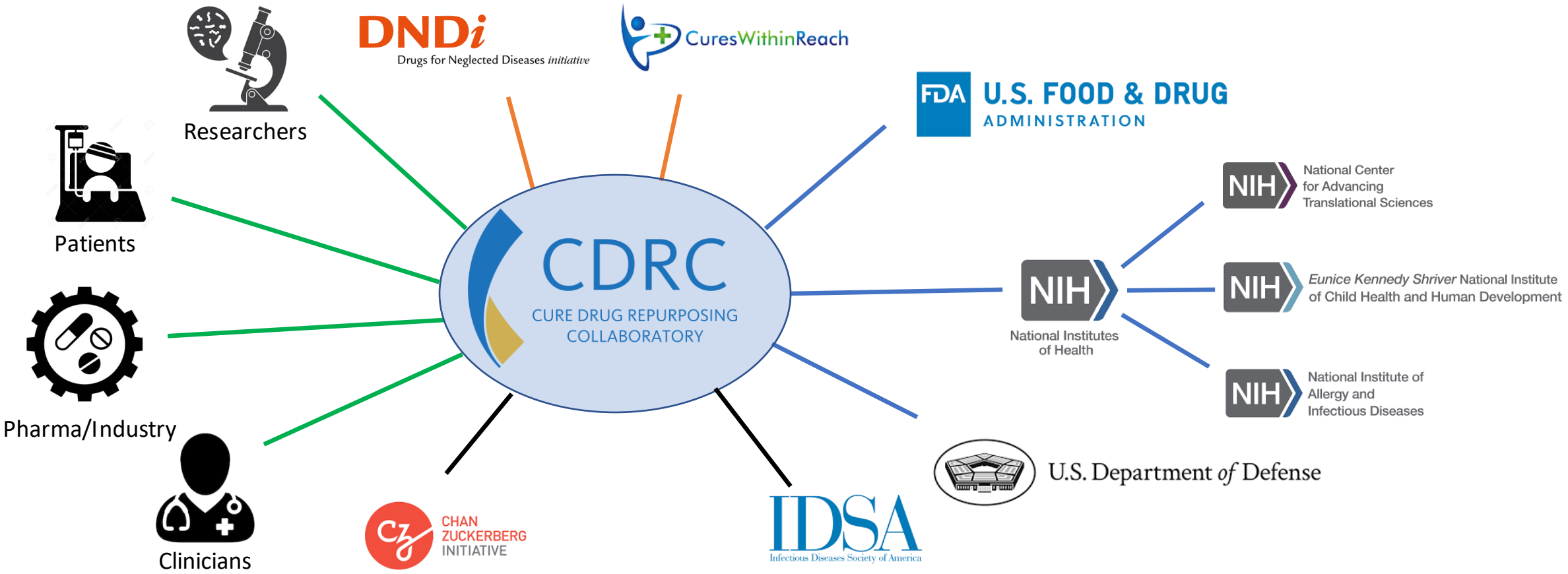


Lisa Schill
(Patient Advoc)



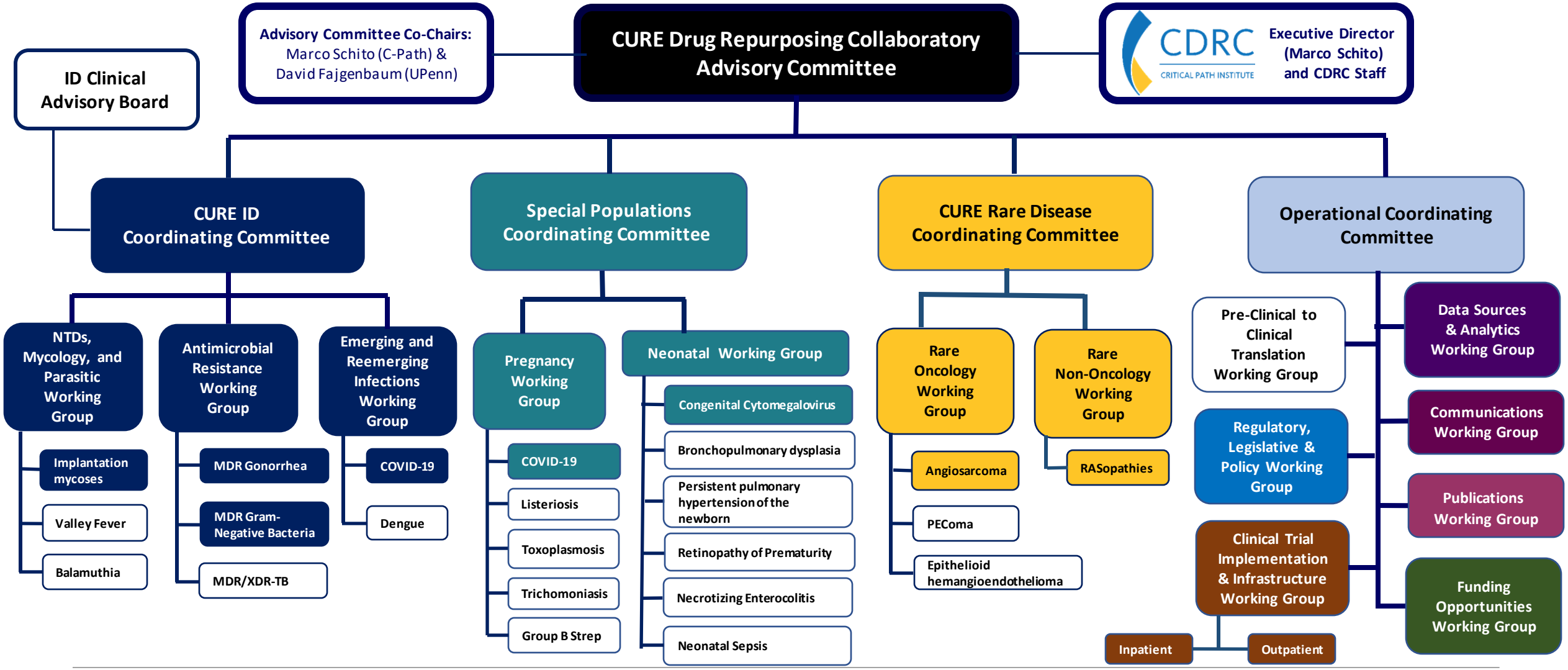
Denise Reinke
(Patient Advoc)

CDRC: Advisory committee members



Goal: Facilitate a drug's demonstrated utility in clinical practice, harnessing a combination of traditional trials and real-world evidence

CDRC Structure



Drug Repurposing

- Significant portion of conditions have no FDA approved treatments
- Traditional drug development and labeling is slow and expensive.
- Repurposing may be needed in diseases which are:
  - rapidly emerging
  - extremely rare
  - impacting vulnerable groups
  - treated by standardized guidelines

Infectious disease examples

Examples of Diseases where Drug Repurposing may be important

Antimicrobial-Resistant Organisms

- MDR Gram Negative bacteria
- DR Gonorrhea and Mycoplasma treatment failures
- Resistant fungal infections

Emerging Infectious Threats

- COVID-19
- Lassa Fever
- Nipah Virus
- Powassan Virus

Unusual Infections

- Balamuthia mandrillaris
- Acinetobacter baumanii
- Mycobacterium abscessus
- Fusarium

Infections in Challenging Sites

- Osteomyelitis
- Infective endocarditis
- Meningitis
- Endophthalmitis

The best use case for CURE ID is for diseases where clinical trials are not or cannot be conducted.

In some instances (e.g., start of outbreak), it may be helpful to capture data that could later be used to inform the design and conduct of RCTs.

Real-World Data

- Explore clinical practice
- Develop and refine hypotheses
- Provide external controls
- Observational research

Randomized, Interventional Study		Nonrandomized, Interventional Study	Nonrandomized, Noninterventional Study
Traditional randomized trial using RWD in planning	**Trial in clinical practice settings, with pragmatic elements**	**Externally controlled trial**	**Observational study**
RWD used to assess enrollment criteria and trial feasibility RWD used to support selection of trial sites	Selected outcomes identified using, e.g., health records data, claims data, or data from digital health technologies RCT conducted using, e.g., electronic case report forms for health records data or claims data	Single-group trial with external control group derived from RWD	Cohort study Case–control study Case–crossover study
	Generation of RWE		
Increasing reliance on RWD			

Reliance on RWD in Representative Types of Study Design.

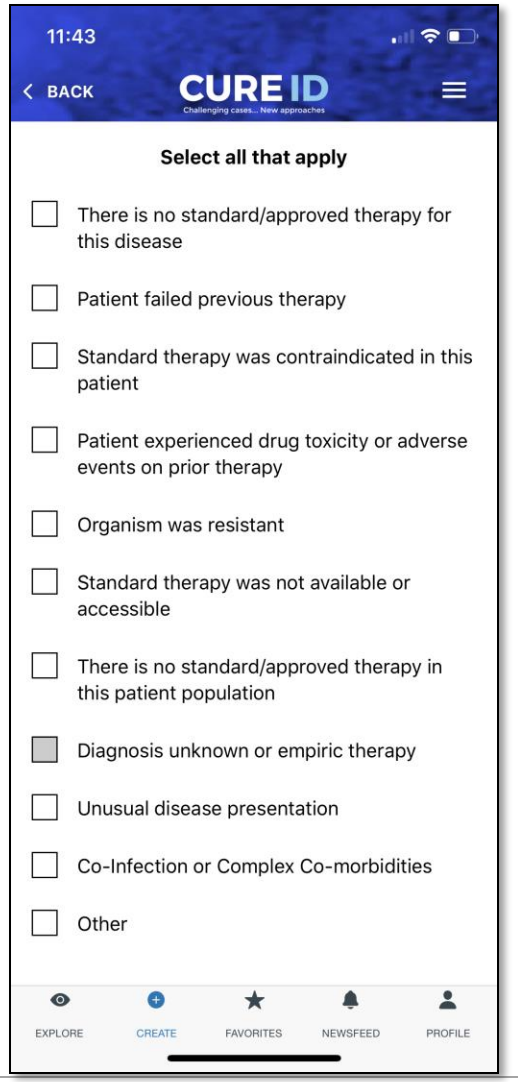
RCT denotes randomized, controlled trial; RWD real-world data; and RWE real-world evidence.

Concato J, Corrigan-Curay J. Real-World Evidence — Where Are We Now? N Engl J Med. 2022;386(18):1680-1682. doi:10.1056/NEJMp2200089

CURE ID data sources

CRITICAL PATH INSTITUTE

- Case reports from literature
- Clinician submitted infectious disease case reports
  1. What disease did your patient have?
  2. How did you make the diagnosis?
  3. **What made your patient's infection difficult to treat?**
  4. What drug(s) did you use to address this difficult to treat infection?
  5. What was the patient's outcome?
  6. Did the patient experience any adverse events?



Download the CURE ID App Today!

Download on the App Store | GET IT ON Google Play

https://cure.ncats.io

@id_cure

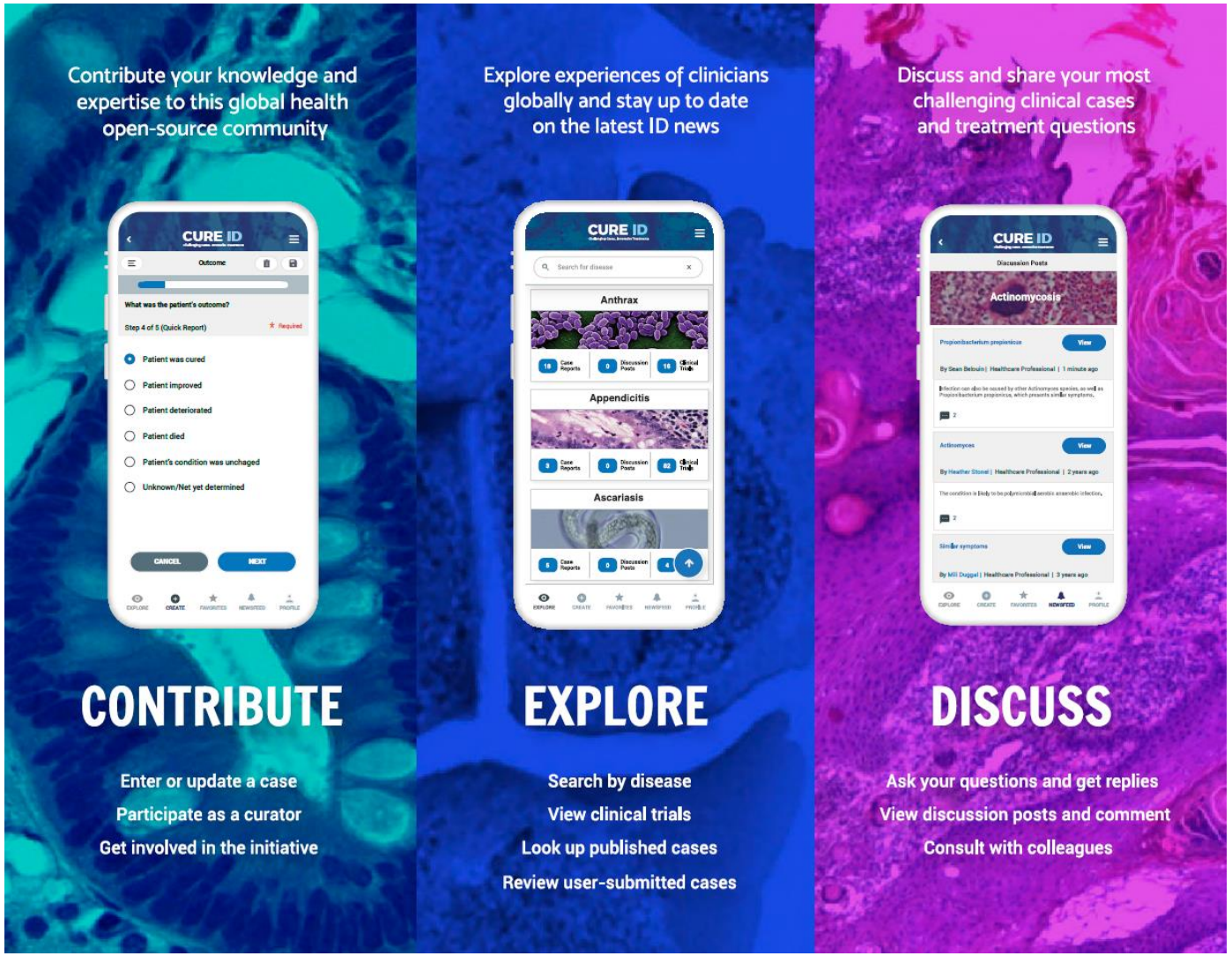
curesupport@nih.gov

FDA U.S. FOOD & DRUG ADMINISTRATION

NIH National Center for Advancing Translational Sciences

A Platform to Capture Novel Uses of Existing Drugs

- Web-based tool
  - Computer, smartphone or mobile device
- Capture and share real-world experiences treating patients through a simple online case report form
- HIPAA compliant, contains no PII
- Newsfeed
- Link to www.clinicaltrials.gov



https://cure.ncats.io/explore

Real-World Data: Challenges

- Unstructured data
- Defining outcomes
- Privacy concerns
- Probabilistic linkage
- Data harmonization

Real-World Data: Data Harmonization

EXTRACT

TRANSFORM

LOAD

ETL

Real-World Data: Data Harmonization

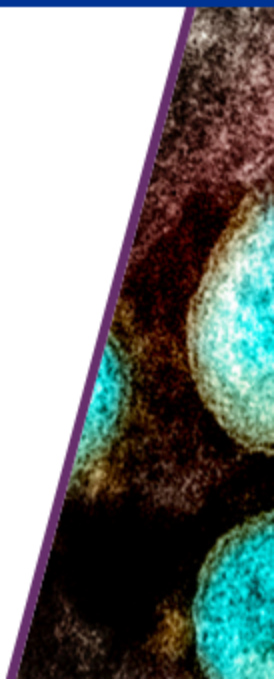
- Centralized harmonization
- Sites use common data model at baseline
- Only for US data
- Limited to COVID-19

National COVID Cohort Collaborative (N3C)

The N3C is a partnership among the NCATS-supported **Clinical and Translational Science Awards (CTSA) Program** hubs, the **National Center for Data to Health (CD2H)**, and NIGMS-supported **Institutional Development Award Networks for Clinical and Translational Research (IDeA-CTR)**, with overall stewardship by NCATS. Collaborators will contribute and use COVID-19 clinical data to answer critical research questions to address the pandemic.

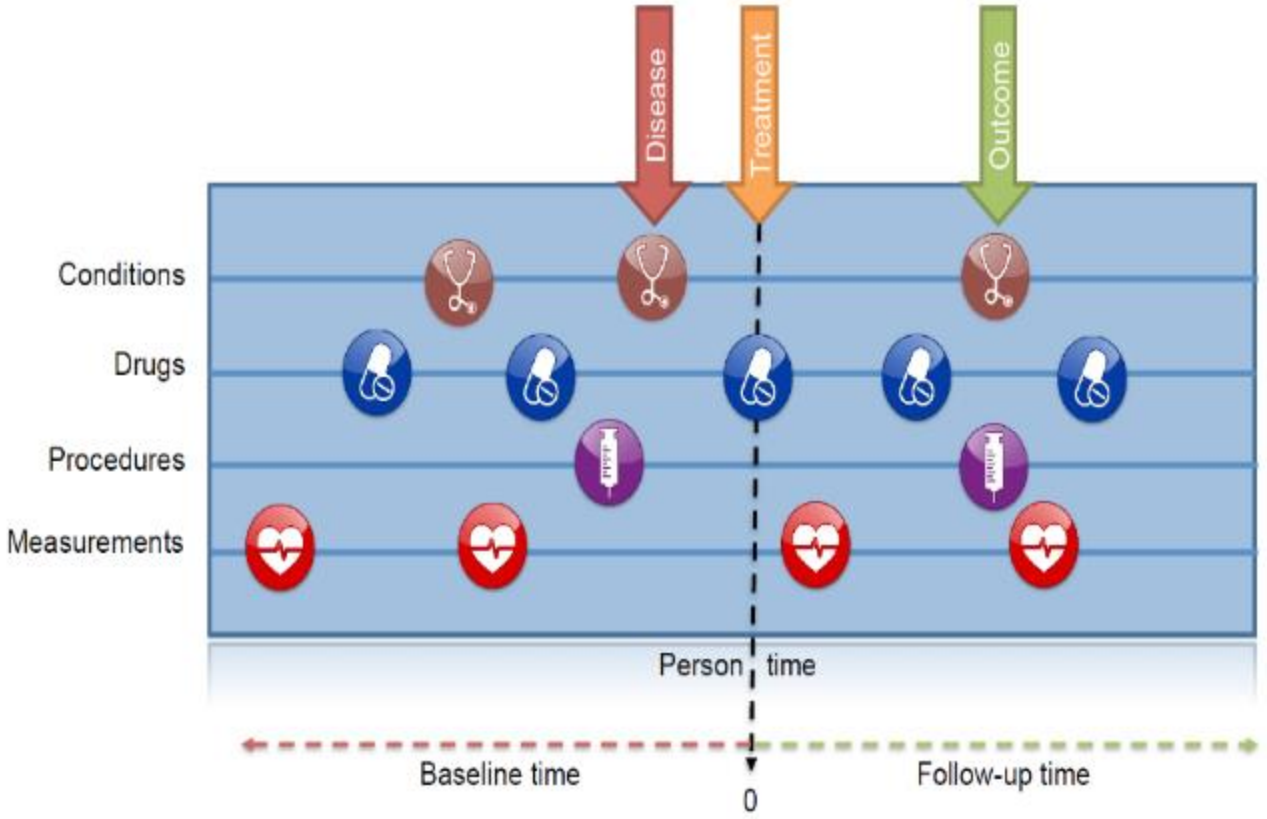
Scientists Use N3C Data to Identify Common Features of Long COVID

NIH-supported researchers used electronic health record data from the National COVID Cohort Collaborative (N3C) Data Enclave to identify people with long COVID and those likely to have it.



ncats.nih.gov/n3c

Observational Medical Outcomes Partnership (OMOP)



153 Controlled Vocabularies
9 Million concepts

ICD-10

NIH National Library of Medicine

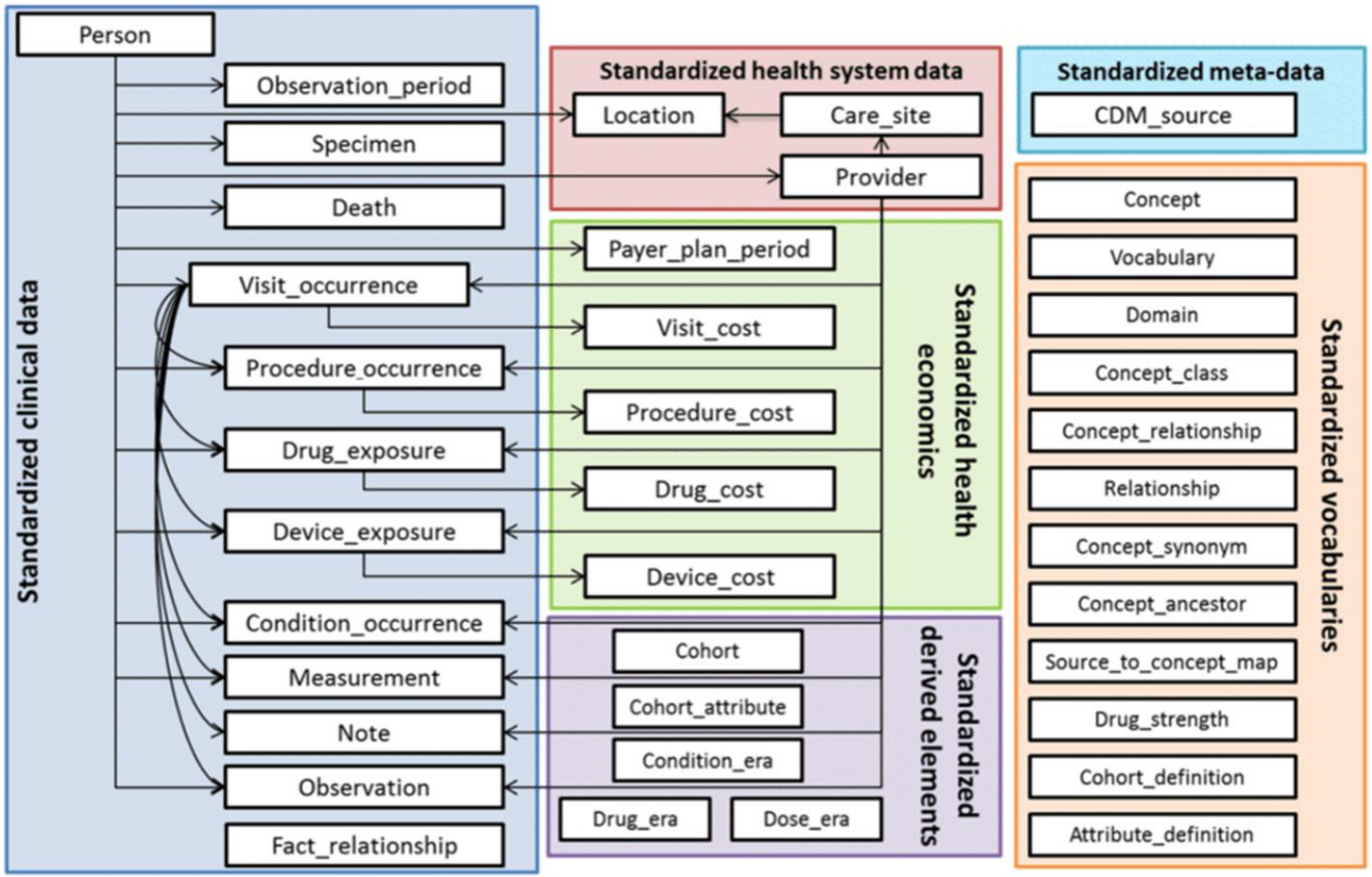
SNOMED International Leading healthcare terminology, worldwide

OHDSI Observational Health Data Sciences and Informatics

LOINC® from Regenstrief

The Edge Tool

Proprietary Data Model



OMOP CDM V5.3.1

The Edge Tool

- Web-based decision support for concept mapping
- Base configuration settings for major EMR Vendors.
- Configuration management documentation tool
- Inspection Report of DevOps on ETL processes
- Data Quality Dashboard framework of 3,000+ data quality tests
- Collaborative cohort subset definition
- Perform de-identification and submission
- **All open-source resources**

The Edge Tool on Azure

CDRC CURE DRUG REPURPOSING COLLABORATORY | CURE ID Challenging cases... New approaches

ETL

- Perseus
- EMR base config
- Usagi
- White Rabbit
- Rabbit in the Hat

DevOps

- Data Quality Dashboard
- Documentation Engine
- Submission extraction
- Change control

Analysis

- Atlas
- WebAPI
- Hades
- R-Studio
- Methods Library

Microsoft Azure

Azure SQL Server
OMOP Data Model
Vocabulary Management
Authentication and Authorization

The Edge Tool: De-identification

- Reassignment of Person IDs: Person IDs are regenerated sequentially from a sorted copy of the Person table. These new Person IDs are carried throughout the CDM to all tables that reference it.
- Date Shifting:
  - Each person is assigned a random date shift value between -186 and +186 days. All dates for that person are then shifted by that amount.
  - Birthdays: After date shifting a person's birthday, the day is then set to the first of the new birth month. If the person would be > 89 years old then they are assigned a random birth year that would make them 90-99 years old.
- Date Truncation:
  - A user-defined Start and End date are used to exclude any date shifted data that falls outside of the target date range (E.G. Procedures, conditions occurrences, etc. Does not include Birthdates).
- Removal of other identifiers:
  - Other potentially identifying datapoints are removed from the dataset such as `location_id`, `provider_id`, and `care_site_id`

Preserving temporal relations in clinical data while maintaining privacy

George Hripcsak,[1] Parsa Mirhaji,[2] Alexander FH Low,[3] and Bradley A Malin[4,5]

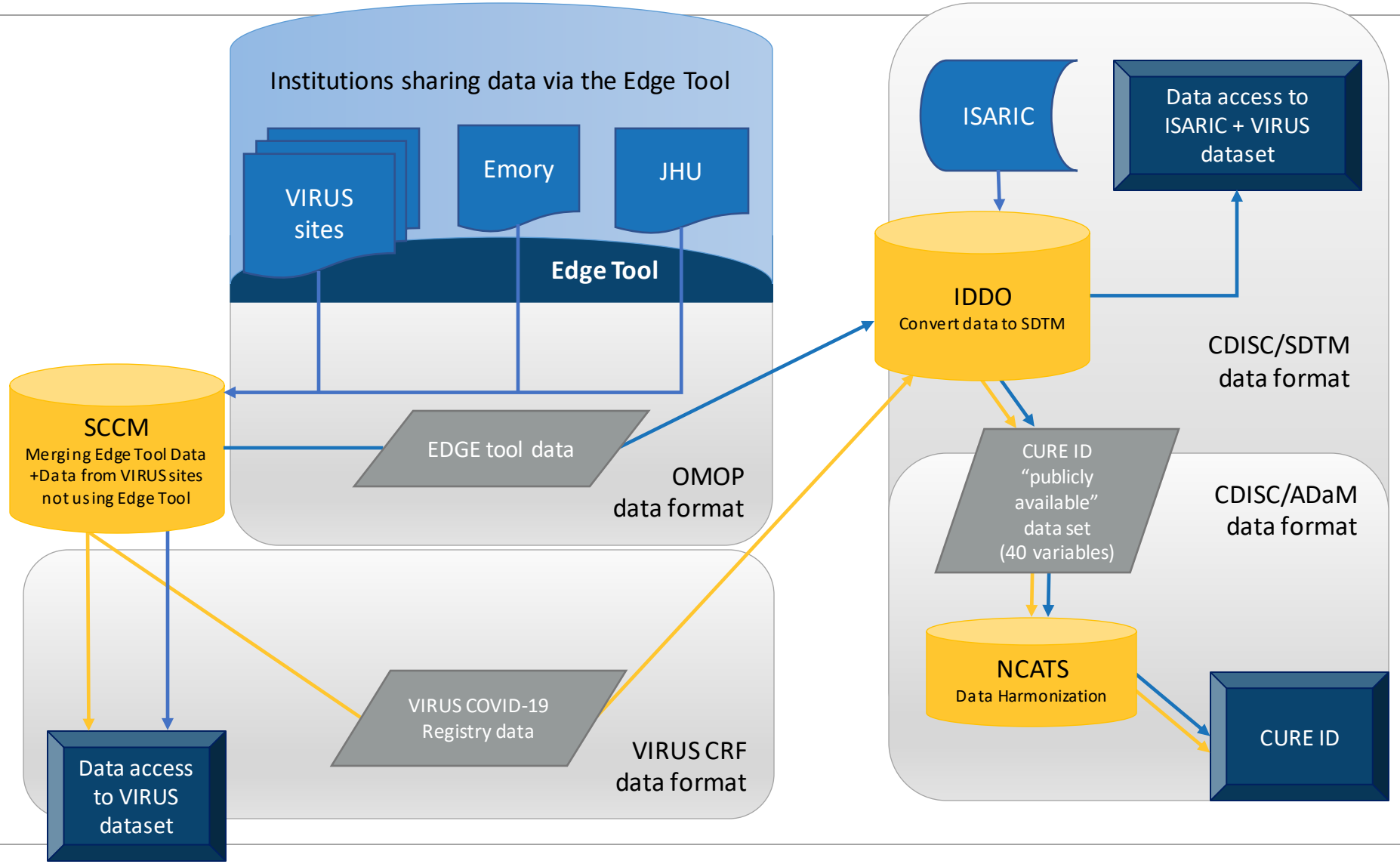
RECEIVED 30 September 2015
REVISED 4 January 2016
ACCEPTED 6 January 2016
PUBLISHED ONLINE FIRST 24 March 2016

ABSTRACT

The Edge Tool Use Case: COVID-19

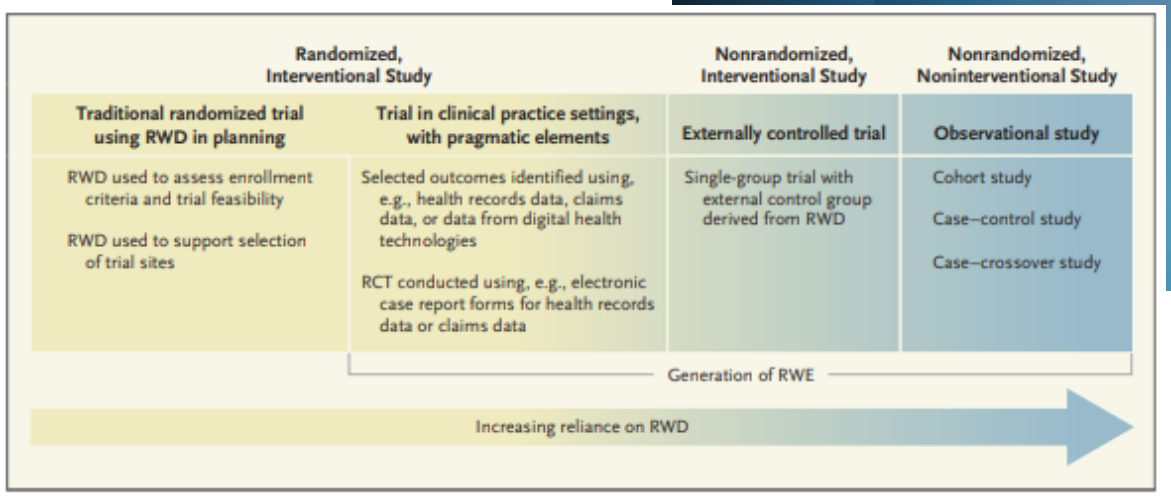
- Drug Repurposing
- Broad Impact
- Identifiable Cases
- Discreet Data
- Acute Disease
- Definitive Outcomes

Data Flow


Institutions sharing data via the Edge Tool
VIRUS sites
Emory
JHU
Edge Tool
EDGE tool data
OMOP data format
SCCM
Merging Edge Tool Data +Data from VIRUS sites not using Edge Tool
VIRUS COVID-19 Registry data
VIRUS CRF data format
Data access to VIRUS dataset
ISARIC
Data access to ISARIC + VIRUS dataset
IDDO
Convert data to SDTM
CDISC/SDTM data format
CURE ID "publicly available" data set (40 variables)
CDISC/ADaM data format
NCATS
Data Harmonization
CURE ID

Real-World Evidence: COVID-19

- Replicate findings of clinical trials
- Evaluate key agents: dexamethasone, baricitinib, tocilizumab
- Demonstrate utility of Edge Tool



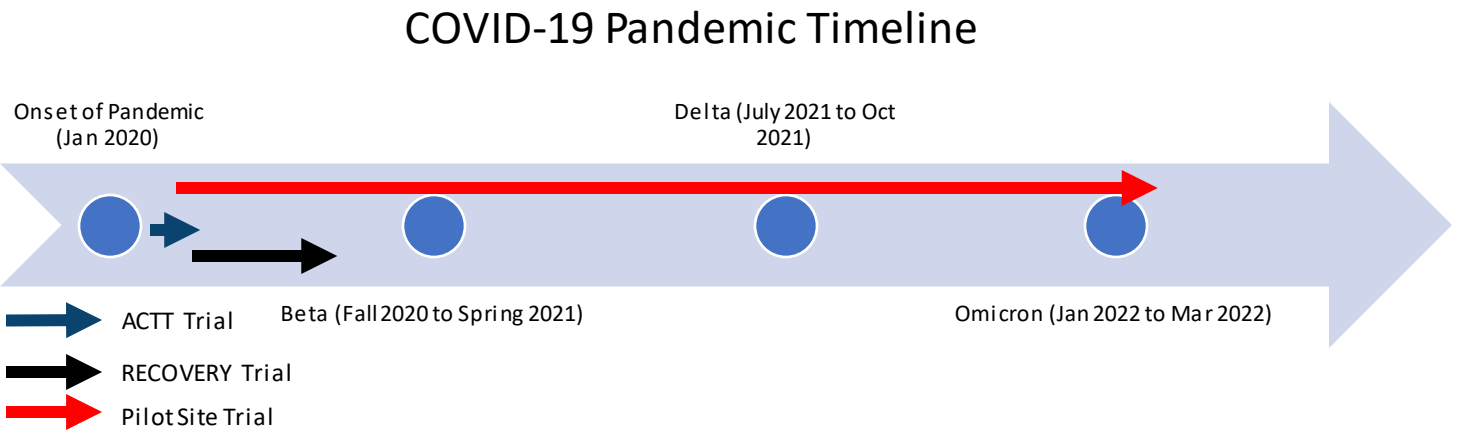
Reliance on RWD in Representative Types of Study Design.

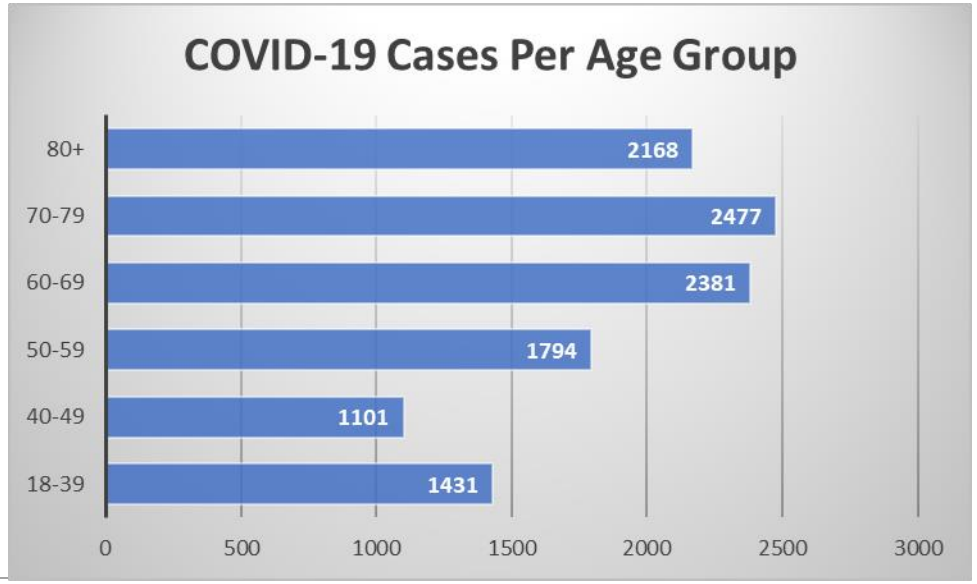
RCT denotes randomized, controlled trial; RWD real-world data; and RWE real-world evidence.

Concato J, Corrigan-Curay J. Real-World Evidence — Where Are We Now? N Engl J Med. 2022;386(18):1680-1682. doi:10.1056/NEJMp2200089

Pilot site data: (March 2020- March 2022)

- Methodology: Data cleaning
  - Observational Medical Outcomes Partnership Common Data Model (OMOP CDM)
  - Originally, N = 12,129 patients
    - Inclusion criteria: 18+, no missing demographic data (ethnicity, race)
      - N = 11,297 patients



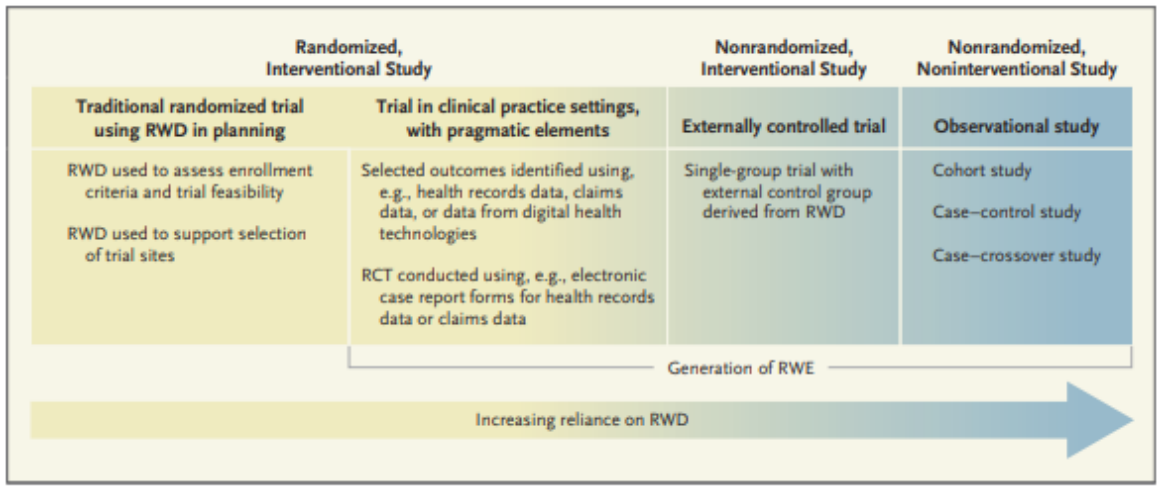


Limitations

- Unmeasured covariates
- Small sample size, control matching with replacement
- Single site
- Timing limitations
  - Shift-and-truncate
  - Data collection after dexamethasone was shown to be effective
- Covariates outside of data collected in trial not accounted for
- Context of treatment (timing and dosage) not accounted for

Future Directions: Beyond COVID

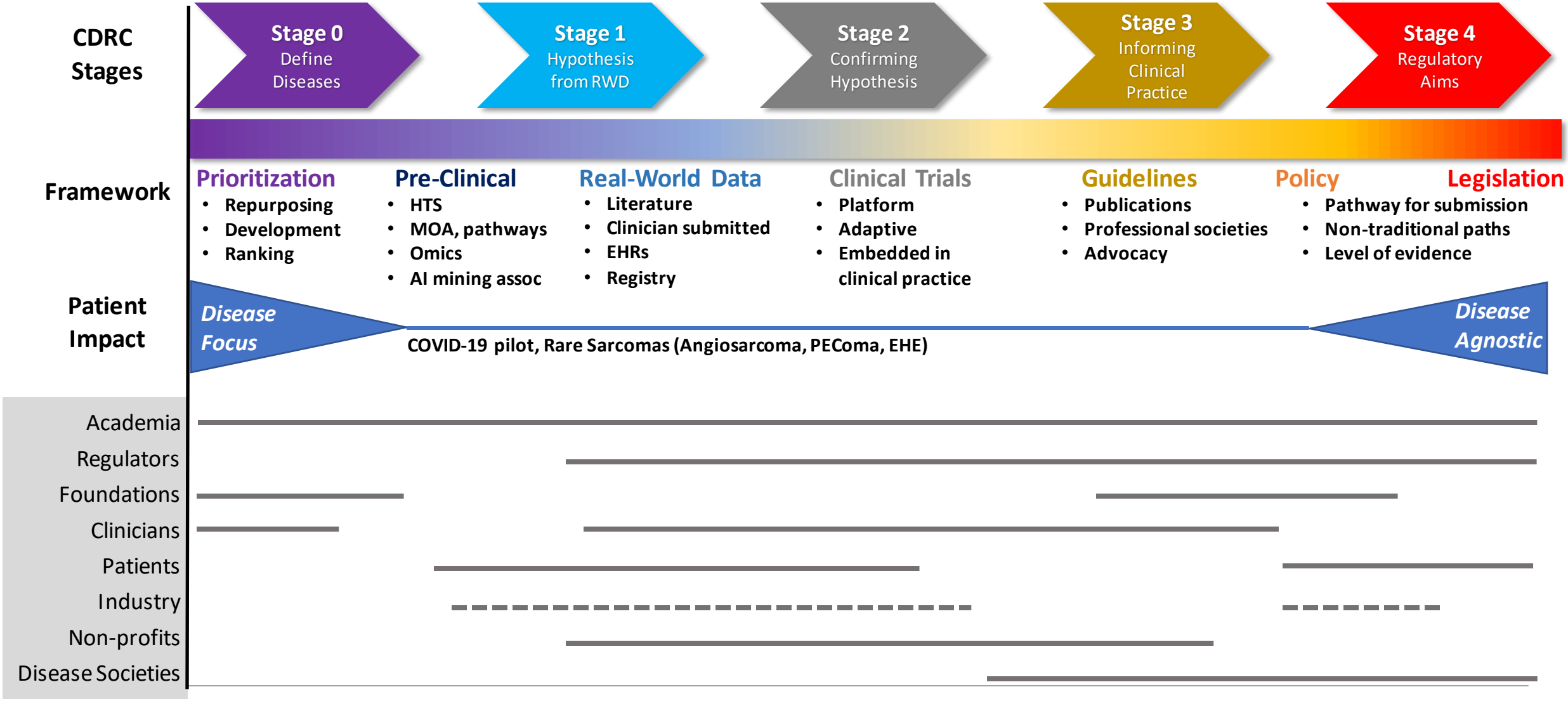
- Sepsis
  - Isolating organisms
  - Case definition
- Meningitis
  - Isolating organisms
  - Rarer cases of more interest
- Osteomyelitis
  - Isolating organisms
  - Linking encounters
  - Lost to follow up



Reliance on RWD in Representative Types of Study Design.
RCT denotes randomized, controlled trial; RWD real-world data; and RWE real-world evidence.

Concato J, Corrigan-Curay J. Real-World Evidence — Where Are We Now? N Engl J Med. 2022;386(18):1680-1682. doi:10.1056/NEJMp2200089

Developing partnerships and infrastructure to provide sustainable resources to impact patient treatments globally



Thank You

<- scan for contact information

CDRC
CRITICAL PATH INSTITUTE

CDRC
CURE DRUG REPURPOSING COLLABORATORY

CURE ID
Challenging cases... New approaches

References

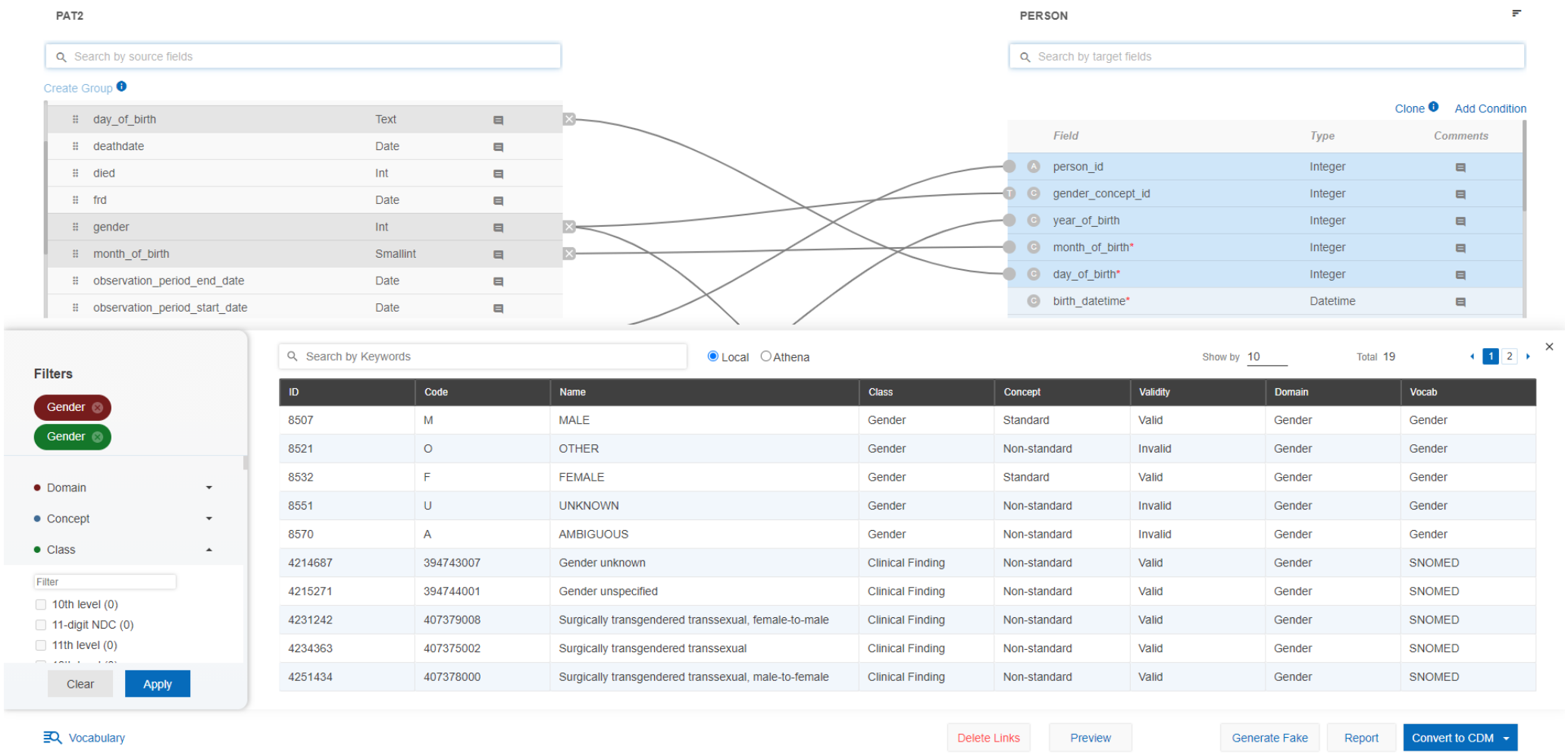
- Begley CG, Ashton M, Baell J, et al. Drug repurposing: Misconceptions, challenges, and opportunities for academic researchers. *Science Translational Medicine*. 2021;13(612):eabd5524. doi:10.1126/scitranslmed.abd5524
- Biedermann, P., Ong, R., Davydov, A. et al. Standardizing registry data to the OMOP Common Data Model: experience from three pulmonary hypertension databases. BMC Med Res Methodol 21, 238 (2021). https://doi.org/10.1186/s12874-021-01434-3
- Concato J, Corrigan-Curay J. Real-World Evidence — Where Are We Now? N Engl J Med. 2022;386(18):1680-1682. doi:10.1056/NEJMp2200089
- Food and Drug Administration. Framework for FDA's real-world evidence program. December 2018 (https://www.fda.gov/media/120060/download).
- Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, Estiri H, Goerg C, Holve E, Johnson SG, Liaw ST, Hamilton-Lopez M, Meeker D, Ong TC, Ryan P, Shang N, Weiskopf NG, Weng C, Zozus MN, Schilling L. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. EGEMS (Wash DC). 2016 Sep 11;4(1):1244. doi: 10.13063/2327-9214.1244. PMID: 27713905; PMCID: PMC5051581.
- Singh TU, Parida S, Lingaraju MC, Kesavan M, Kumar D, Singh RK. Drug repurposing approach to fight COVID-19. *Pharmacol Rep*. 2020;72(6):1479-1508. doi:10.1007/s43440-020-00155-6
- Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18(1):41-58. doi:10.1038/nrd.2018.168
- Oprea TI, Bauman JE, Bologa CG, et al. Drug repurposing from an academic perspective. *Drug Discovery Today: Therapeutic Strategies*. 2011;8(3):61-69. doi:10.1016/j.ddstr.2011.10.002

Edge Tool Screen Grabs

CDRC
CURE DRUG REPURPOSING COLLABORATORY

CURE ID
Challenging cases... New approaches

Concept Mapping Decision Support



https://github.com/SoftwareCountry/Perseus

https://www.youtube.com/watch?v=JBgIiVKDOyc

Documentation Engine

CDRC
CRITICAL PATH INSTITUTE

OHDSI

The JHM Epic to OMOP ETL Guideline

Search this book...

Epic to OMOP ETL Project

Contents

Goals

- Publish a guide that can be computable (.ipynb) as well as easily searchable in html and pdf formats.
- Have the guide be compiled by the scripts so as new scripts are added they are automatically added to the guide
- Encode the scripts with markdown to enable navigation and readability
- Have sufficient meta data in the scripts to help users of the guide index the data for common questions.
- Whenever possible documentation should not be separated from the code.

Assumptions

- Epic data model is proprietary, distribution should be through Epic customer portal (galaxy)
- Epic uses SQL Server so we will base our scripts on MS T-SQL
- There is setup necessary to configure an OMOP CDM database with tables, vocabulary, and constraints.
- There are 60+ SQL stored procedures to transform Epic Clarity data into the OMOP CDM 5.3.1
- There is a sequence of operations to orchestrate the transformation
- The SQL scripts use an Ansi UTF text file with basic documentation capabilities (/* Comments */) for inline comments
- There are resources online available that are key to doing a transformation. Ie CDM github conventions and Book OHDSI
- Sites will need to change ETL scripts for site specific localizations where data is coming from flowsheets and smartdata forms.
- ETL scripts should be human readable

Users of the guide

- Database administrators implementing this ETL in their organization.
- Researchers wanting to see the clear provenance of data transformations from the clinical information system.

Questions the guide should facilitate

- Show the epic tables involved in the ETL process
- Show the source Epic clarity tables that feed into an OMOP Domain (condition, measurements, etc)
- Show the domains that are fed from a particular epic table
- Show the order of operations of the scripts

```
--for Logging
DECLARE @RC     INT = 0 , @procName VARCHAR(50) ='OMOP_Fill_Prep', @tablename VARCHAR(50)='ETL_Pre
DECLARE @elapsed_seconds INT = 0 , @comment VARCHAR(255), @start_datetime Datetime = getdate(),
        @end_datetime Datetime, @block_number INT;

-- Flush the tables
        --EXEC stage.OMOP_DEL_Flush;

-- Load SrcIDMaps
        EXEC stage.OMOP_INS_SrcIDMaps_Provider; --from Clarity_SER
        EXEC stage.OMOP_INS_SrcIDMaps_Person;  -- from derived_inpatient_encounters, derived_outpa
        EXEC stage.OMOP_INS_SrcIDMaps_CareSite;  -- from Clarity_DEP, Clarity_POS
        EXEC stage.OMOP_INS_SrcIDMaps_Location;  --from edw_geocode_patient
        EXEC stage.OMOP_INS_SrcIDMaps__Visit; --from from derived_inpatient_encounters, derived_ou

--refresh Source_to_Concept_map
        --new proc loading from Clarity 2C tables and refreshing STCM
        EXEC stage.OMOP_INS_SrcToConcept_JHMRace;
        EXEC stage.OMOP_INS_SrcToConcept_JHMEthnicity;
        EXEC stage.OMOP_INS_SrcToConcept_JHMGender;

        EXEC stage.OMOP_INS_SrcToConcept_JHMPatClass;
        EXEC stage.OMOP_INS_SrcToConcept_JHMOpVisitTypes;
        EXEC stage.OMOP_INS_SrcToConcept_JHMAdmitSource;
        EXEC stage.OMOP_INS_SrcToConcept_JHMDischDisp;
        --EXEC stage.OMOP_INS_SrcToConcept__JHMVisit; replaced by JHMPatClass and JHMOPVisitTypes
        --EXEC stage.OMOP_INS_SrcToConcept__JHMAdmDisch;  replaced by JHMAdmitSource and JHMDischD


-- Fill core block
        EXEC stage.OMOP_ETL_edw_geocode_patient;                                    --creates
        EXEC stage.OMOP_ETL_care_site;                                              --
        EXEC stage.OMOP_ETL_epic_provider;                                          --

-- person, visit
        --EXEC STAGE.OMOP_ETL_derived_inpatient_encounters__expired;
        EXEC stage.OMOP_ETL_derived_epic_patient;                                   --
        EXEC stage.OMOP_ETL_derived_ip_encounters @block_total = 10; --creates visit_occurrence
        EXEC stage.OMOP_ETL_derived_op_encounters @block_total = 10; --creates visit_occurrence

set @end_datetime = getdate();
set @elapsed_seconds = DATEDIFF(second,@start_datetime,@end_datetime);
```

Data Quality Dashboard

Examples of DQ Checks from Kahn et al (2016)

Atemporal Plausibility	• 48% of labs outside of normal range
Temporal Plausibility	• Unexpected change in number of records from month to month
Completeness	• 62% of *route_concept_id* is missing
Value Conformance	• ICD9 codes in *condition_concept_id*
Relational Conformance	• *visit_date* and *visit_datetime* inconsistency in

DATA QUALITY ASSESSMENT

JOHNS HOPKINS MEDICINE ENTERPRISE

DataQualityDashboard Version: 1.0.0

Results generated at 2021-08-28 09:20:06 in 7 hours

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1938	75	2013	96%	283	4	287	99%	2221	79	2300	97%
Conformance	643	3	646	100%	100	0	100	100%	743	3	746	100%
Completeness	356	7	363	98%	15	0	15	100%	371	7	378	98%
Total	2937	85	3022	97%	398	4	402	99%	3335	89	3424	**97%**

Data Quality Dashboard

Examples of DQ Checks from Kahn et al (2016)

Atemporal Plausibility	• 48% of labs outside of normal range
Temporal Plausibility	• Unexpected change in number of records from month to month
Completeness	• 62% of *route_concept_id* is missing
Value Conformance	• ICD9 codes in *condition_concept_id*
Relational Conformance	• *visit_date* and *visit_datetime* inconsistency in

DATA QUALITY ASSESSMENT

JOHNS HOPKINS MEDICINE ENTERPRISE

DataQualityDashboard Version: 1.0.0

Results generated at 2021-08-28 09:20:06 in 7 hours

	Verification		Validation	Total
	Pass	Fail		
Plausibility	1938	75		
Conformance	643	3		
Completeness	356	7		
Total	2937	85		

STATUS	CONTEXT	CATEGORY	SUBCATEGORY	LEVEL	DESCRIPTION	% RECORDS
FAIL	Verification	Completeness	None	FIELD	The number and percent of records with a value of 0 in the standard concept field UNIT_CONCEPT_ID in the SPECIMEN table. (Threshold=5%).	0%
FAIL	Verification	Completeness	None	FIELD	The number and percent of records with a value of 0 in the standard concept field UNIT_CONCEPT_ID in the DOSE_ERA table. (Threshold=0%).	0%

Showing 6 to 7 of 7 entries (filtered from 3,424 total entries)

Previous 1 2 Next

ATLAS

- Home
- Data Sources
- Search
- Concept Sets
- Cohort Definitions
- Characterizations
- Cohort Pathways
- Incidence Rates
- Profiles
- Estimation
- Prediction
- Jobs
- Configuration
- Feedback

Cohort Entry Events

Events having any of the following criteria:

a condition occurrence of Type 2 Diabetes Mellitus Diagn...

a drug exposure of Type 2 Diabetes Mellitus Prescri...

with continuous observation of at least 0 days before and 0 days after event index date

Limit initial events to: all events per person.

Restrict intial events to:

having all of the following criteria:

with at most 0 using all occurrences of:

a condition occurrence of Type 1 Diabetes Mellitus Diagn...

where event starts between All days Before and All days After index start date add additional constraint

restrict to the same visit occurrence

allow events from outside observation period

And having any of the following criteria:

having all of the following criteria:

with at most 0 using all occurrences of:

a condition occurrence of Type 2 Diabetes Mellitus Diagn...

where event starts between All days Before and All days After index start date add additional constraint

restrict to the same visit occurrence

allow events from outside observation period

and with at least 1 using all occurrences of:

a drug exposure of Type 2 Diabetes Mellitus Prescri...

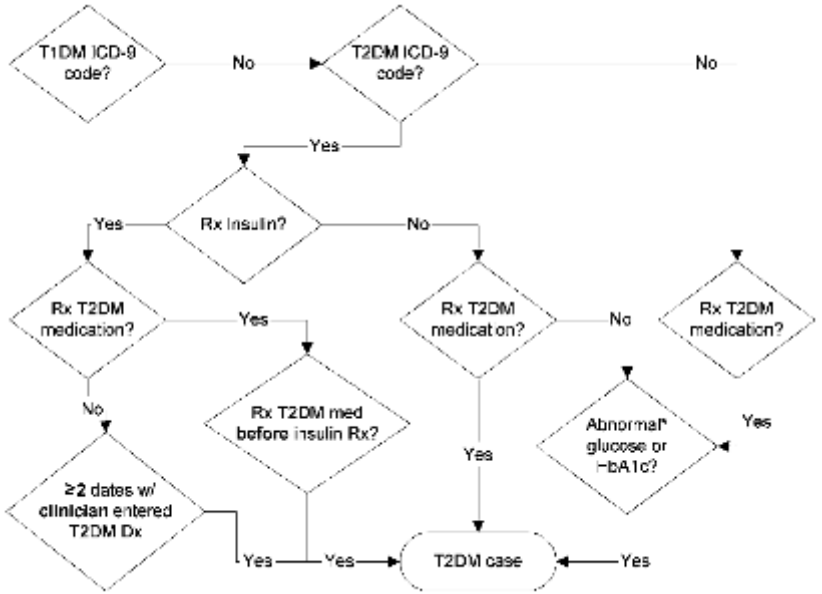
where event starts between All days Before and All days After index start date add additional constr

restrict to the same visit occurrence

allow events from outside observation period

Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study

Abel N Kho,[1] M Geoffrey Hayes,[1] Laura Rasmussen-Torvik,[1] Jennifer A Pacheco,[1] William K Thompson,[1] Loren L Armstrong,[1] Joshua C Denny,[2] Peggy L Peissig,[3] Aaron W Miller,[3] Wei-Qi Wei,[4] Suzette J Bielinski,[4] Christopher G Chute,[4] Cynthia L Leibson,[4] Gail P Jarvik,[5] David R Crosslin,[5] Christopher S Carlson,[6] Katherine M Newton,[7] Wendy A Wolf,[8] Rex L Chisholm,[1] William L Lowe[1]



PheKB >1,000 lines of code

J Am Med Inform Assoc 2012;**19**:212–218. doi:10.1136/amiajnl-2011-000439