

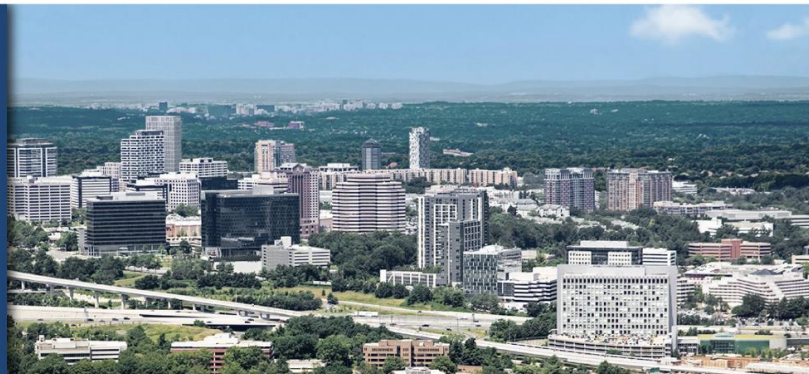


2023

US

INTERCHANGE

FALLS CHURCH, VA | 18-19 OCTOBER



Transforming RWD for Regulatory Submissions: How to Use SDTM for RWD

Presented by Mayur Saxena | Droice Labs



Meet the Speaker

Mayur Saxena, PhD

Title: Cofounder & CEO

Organization: Droice Labs

As an entrepreneur and scientist, Mayur has concentrated on advancing medicine with high-noise, big data analysis. Before founding Droice, he played key roles in several startups, including co-founding a biotechnology firm in the diabetes space. He earned his BTech at IIT Kanpur and his MS and PhD at Columbia University, focusing on the computational physics of disease.



Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*
- *Mayur Saxena is an employee of Droice Labs, a company that processes and prepares RWD for clinical and research applications.*

Requirements for regulatory RWD



Requirements for regulatory RWD

RWD



Data Standards for Drug and
Biological Product Submissions
Containing Real-World Data
Guidance for Industry

Date
October 2021

“ Currently, and absent a waiver, sponsors submitting clinical and nonclinical study data (including those derived from RWD sources) in submissions subject to section 745A(a) of the FD&C Act are required to use the formats described in the Study Data Guidance and the supported study data standards listed in the Catalog. (R 109-112)

”



Requirements for regulatory RWD

RWD

(EHR, Claims, etc..)



Data Standards for Drug and
Biological Product Submissions
Containing Real-World Data
Guidance for Industry

Date
October 2021

“ Currently, and absent a waiver, sponsors submitting clinical and nonclinical study data (including those derived from RWD sources) in submissions subject to section 745A(a) of the FD&C Act are required to use the formats described in the Study Data Guidance and the supported study data standards listed in the Catalog. (R 109-112)

FDA recognizes that a range of approaches may be used to apply currently supported data standards (e.g., Clinical Data Interchange Standards Consortium's (CDISC's) Study Data Tabulation Model (SDTM)) to RWD sources such as EHR or claims data. (R 125-127)

With adequate documentation of the conformance methods used and their rationale, study data derived from RWD can be transformed to SDTM datasets and submitted to FDA in an applicable drug submission. (R 129-131)

”





Scientist

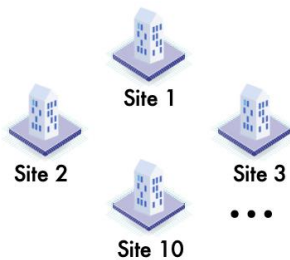
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

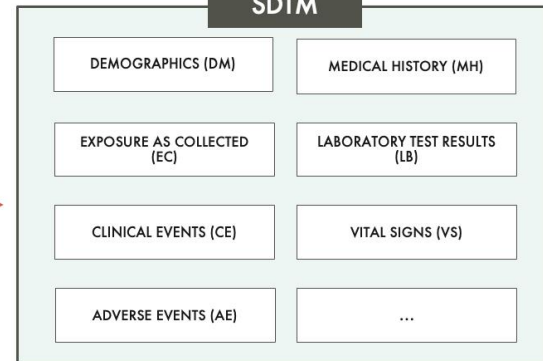
ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



Intervention arm data sources



SDTM



Scientist

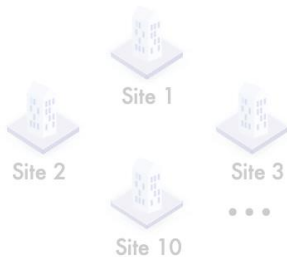
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

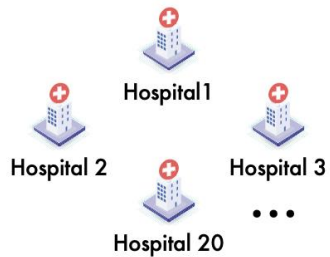
ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



Intervention arm data sources

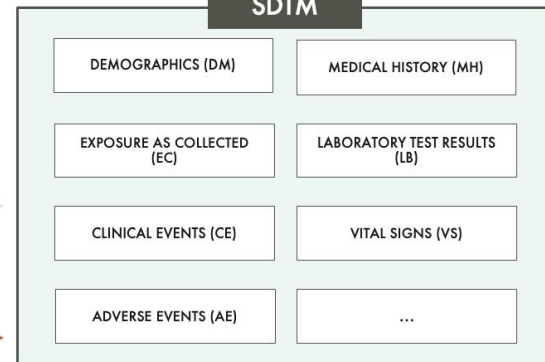


RWD Data sources



CDISC

SDTM



Scientist

Indication: Stroke prevention in atrial fibrillation

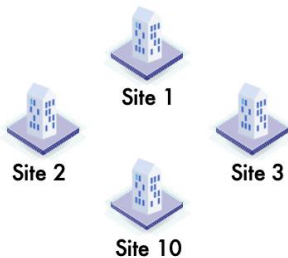
ARM 1: Intervention arm (study drug) DATA SOURCES: 10 sites
PATIENTS: 1,000

ARM 2: External control arm (standard of care; direct oral anticoagulant) DATA SOURCES: EHR - 20 hospitals
PATIENTS: 50,000



RCT Data Journey

Intervention arm data sources



CRF
(From 10 sites)

SDTM

DEMOGRAPHICS (DM)

MEDICAL HISTORY (MH)

EXPOSURE AS COLLECTED
(EC)

LABORATORY TEST RESULTS
(LB)

CLINICAL EVENTS (CE)

VITAL SIGNS (VS)

ADVERSE EVENTS (AE)

...



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: 10 sites
PATIENTS: 1,000

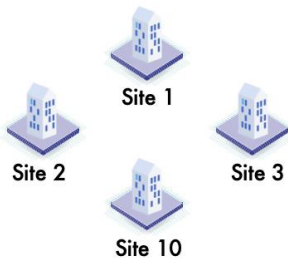
ARM 2:
External control arm (standard of care; direct oral anticoagulant)

DATA SOURCES: EHR - 20 hospitals
PATIENTS: 50,000



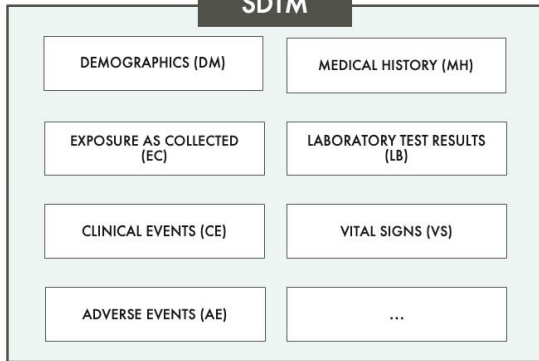
RCT Data Journey

Intervention arm data sources



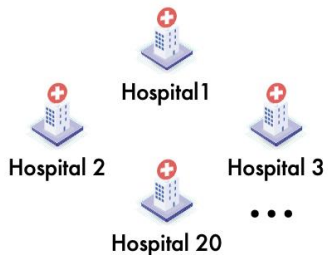
CRF
(From 10 sites)

SDTM



RWD Data Journey

RWD Data sources



?



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look

Data sources



Hospital 1



Hospital 2



Hospital 3



Hospital 4

⋮



Hospital 20



Scientist

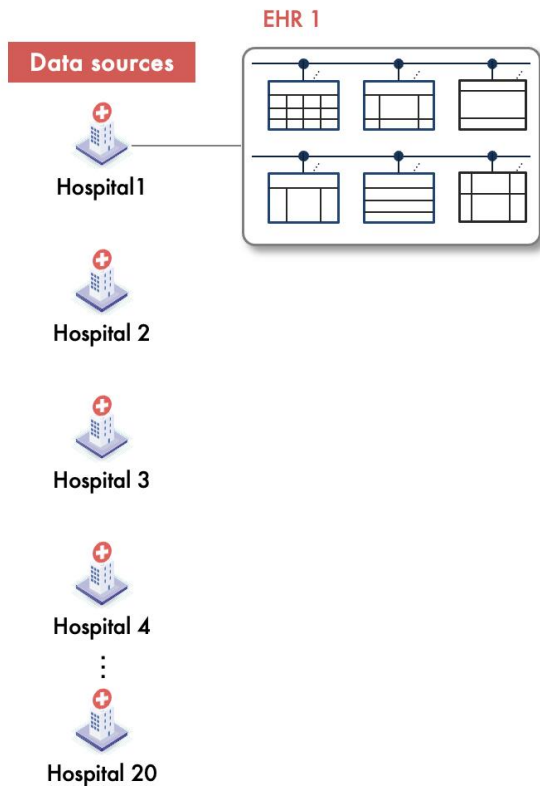
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look



Scientist

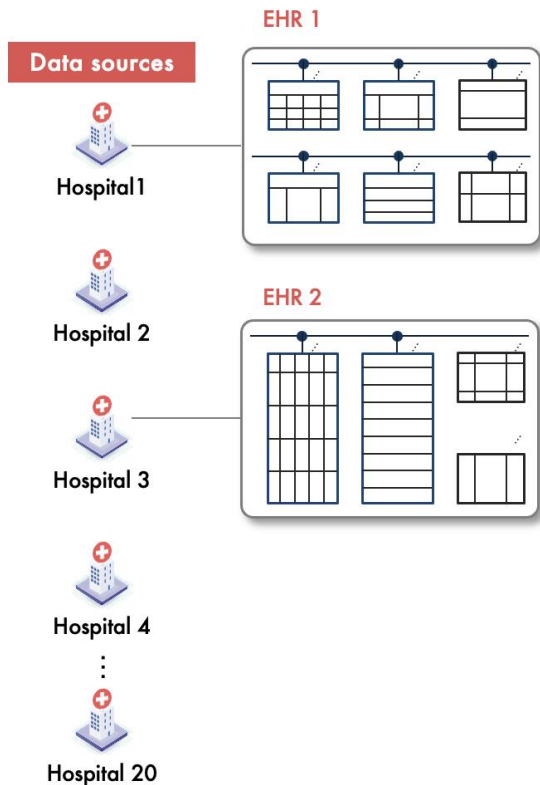
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look



Scientist

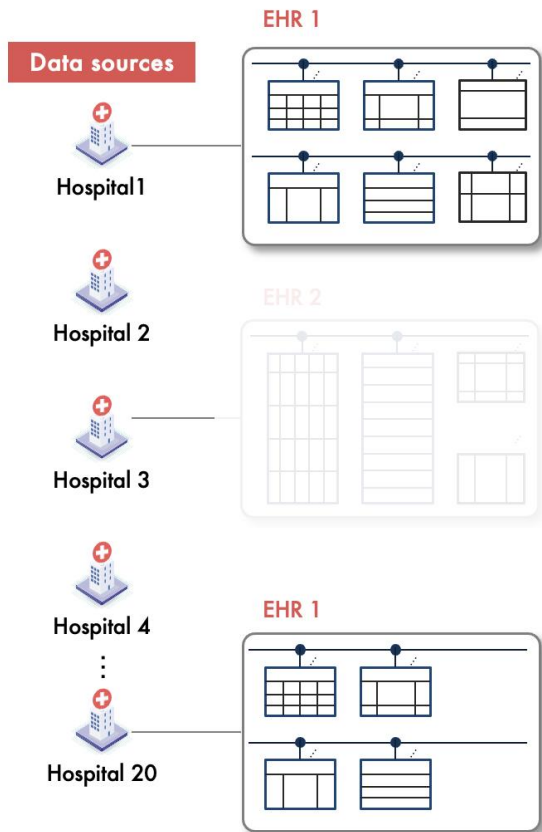
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look



Scientist

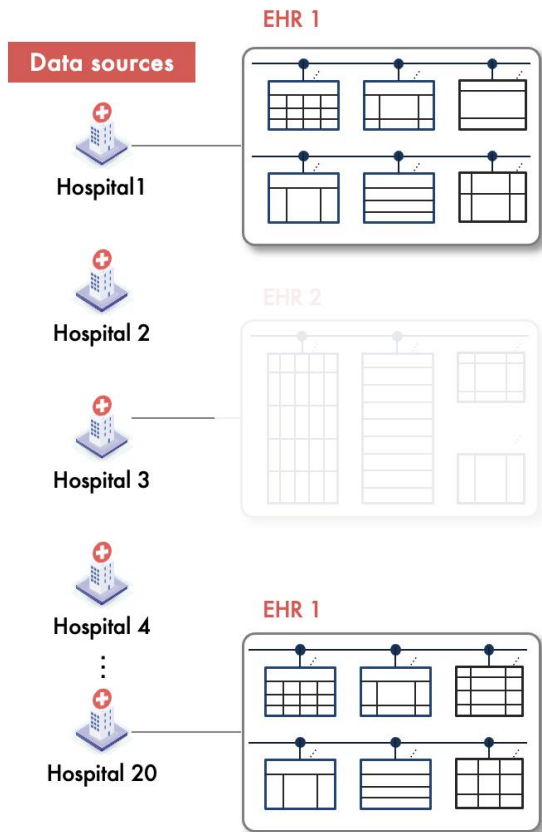
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look



Scientist

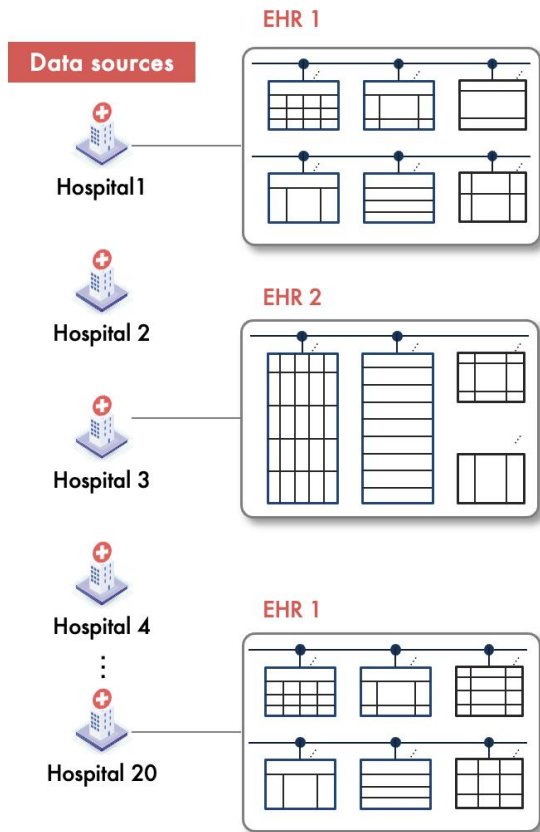
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



RWD to SDTM: A deeper look



Scientist

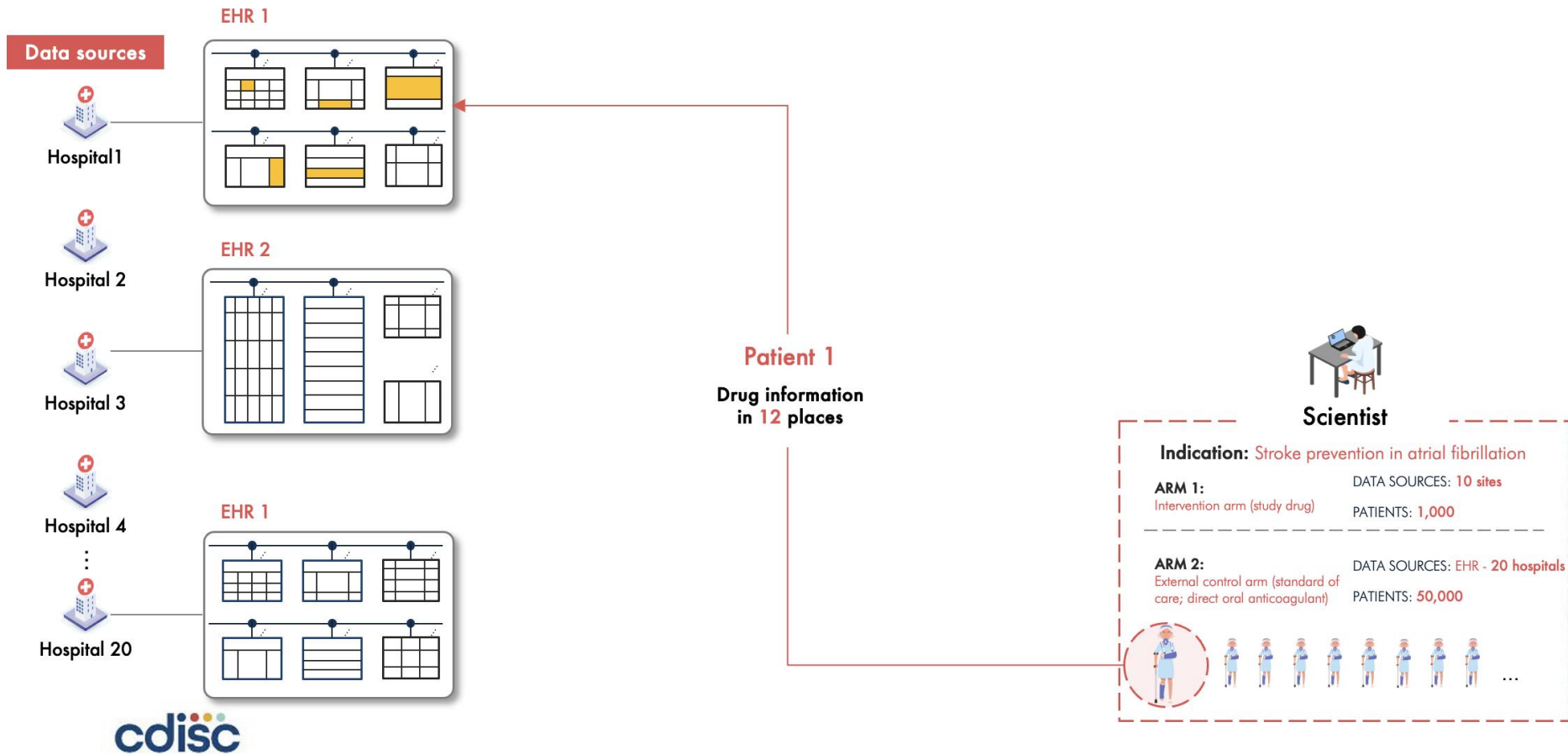
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

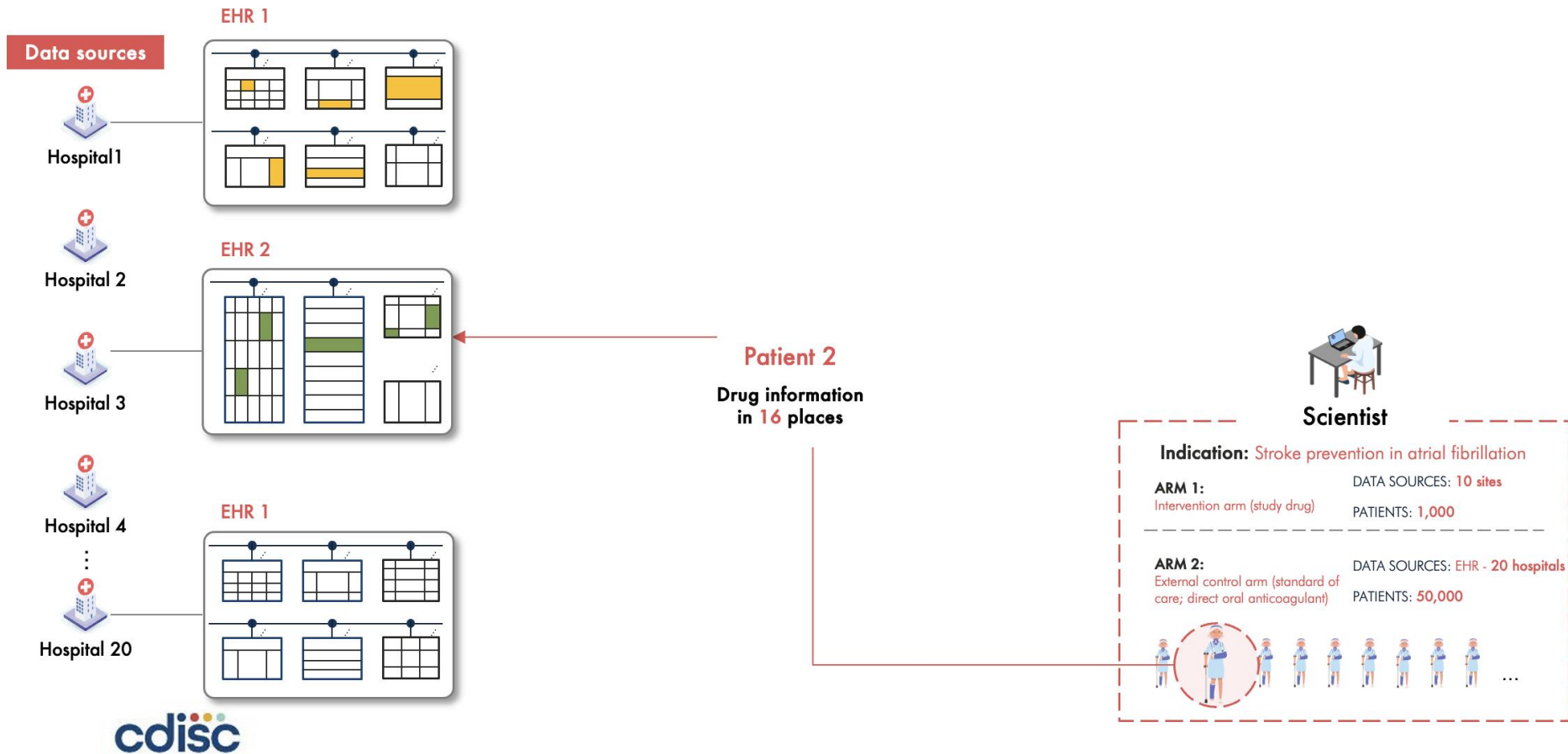
ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



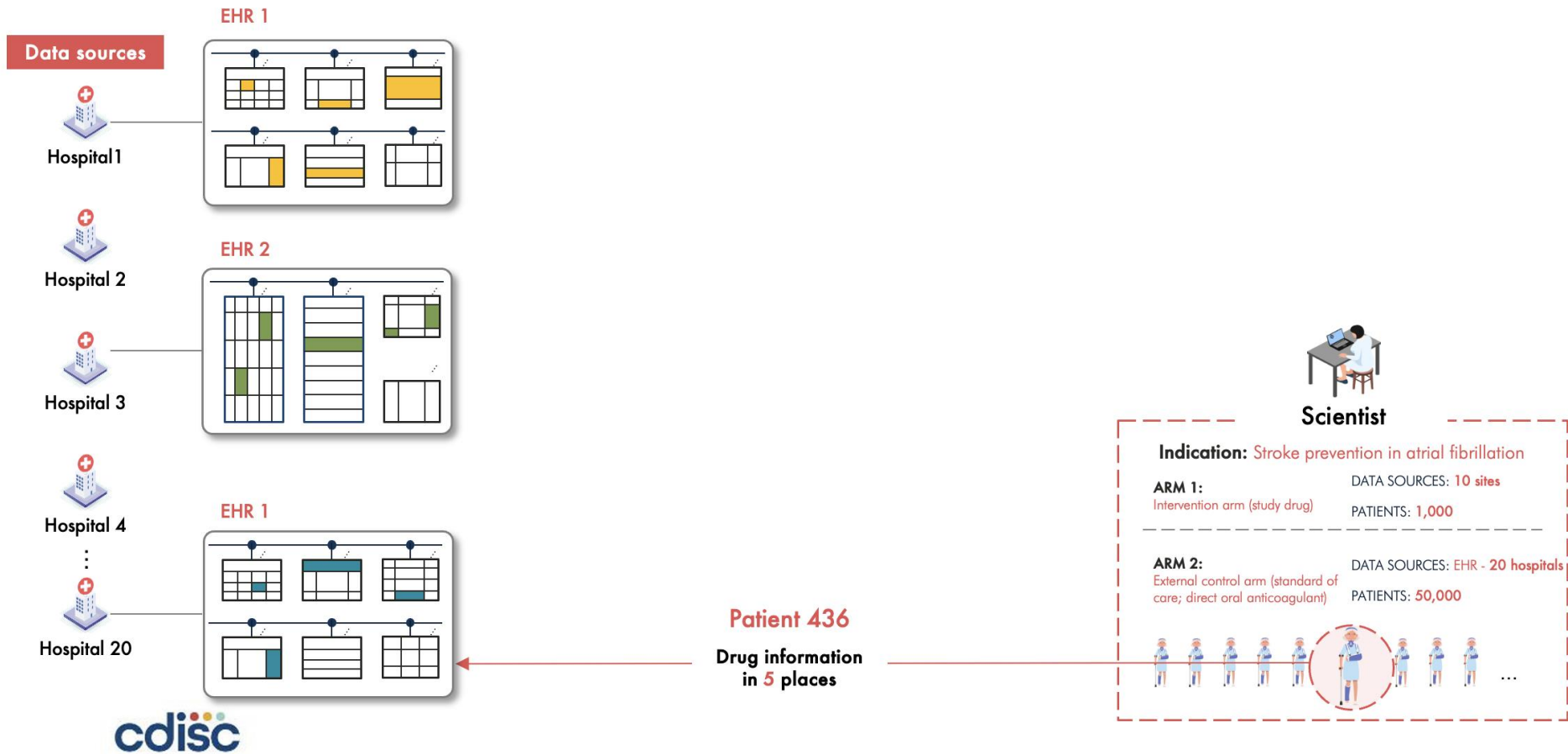
RWD to SDTM: A deeper look



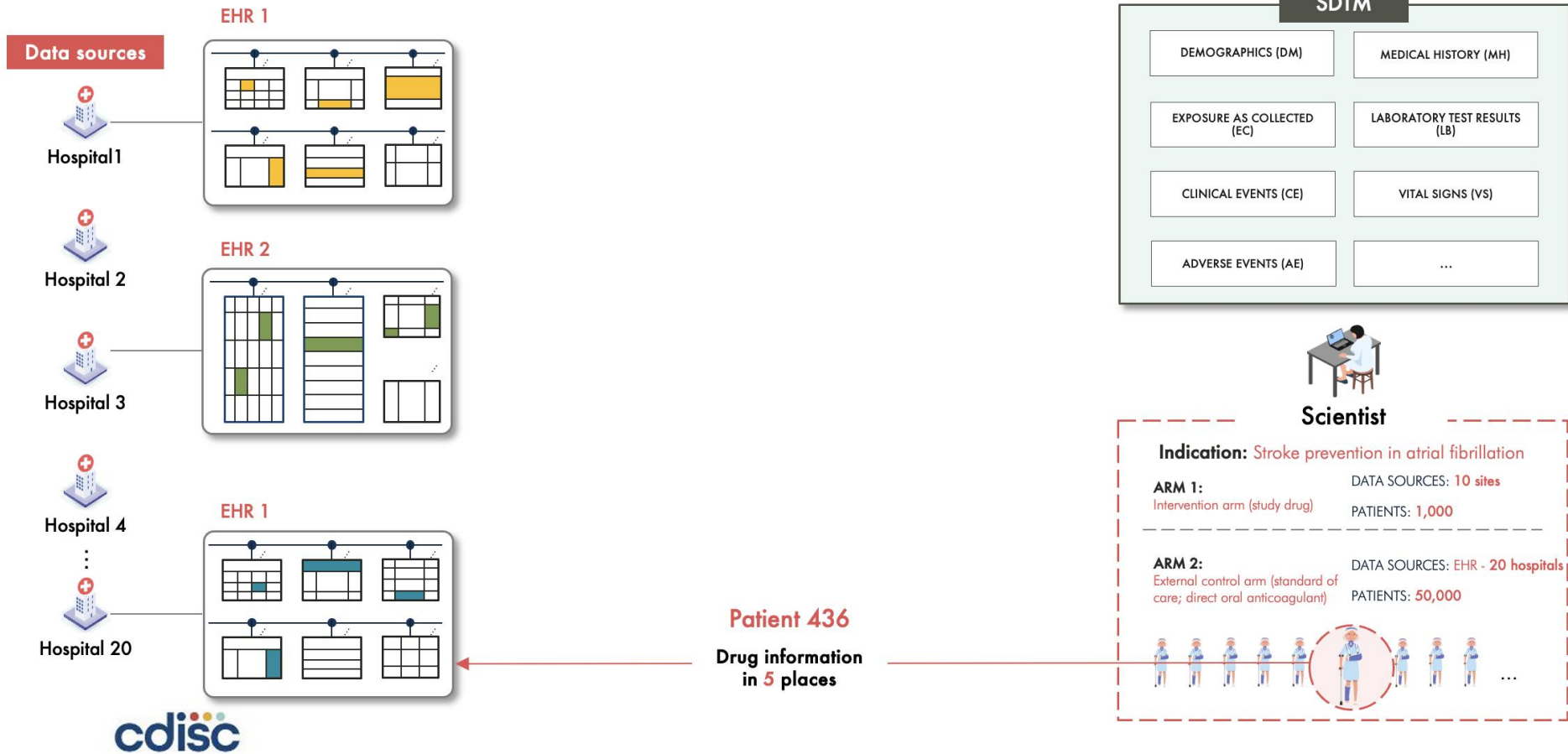
RWD to SDTM: A deeper look



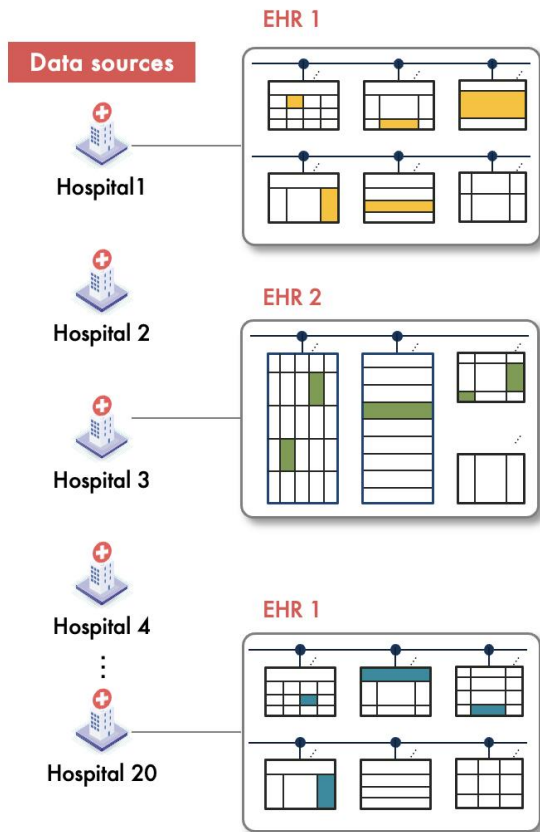
RWD to SDTM: A deeper look



RWD to SDTM: A deeper look



Common data models (CDM)



Scientist

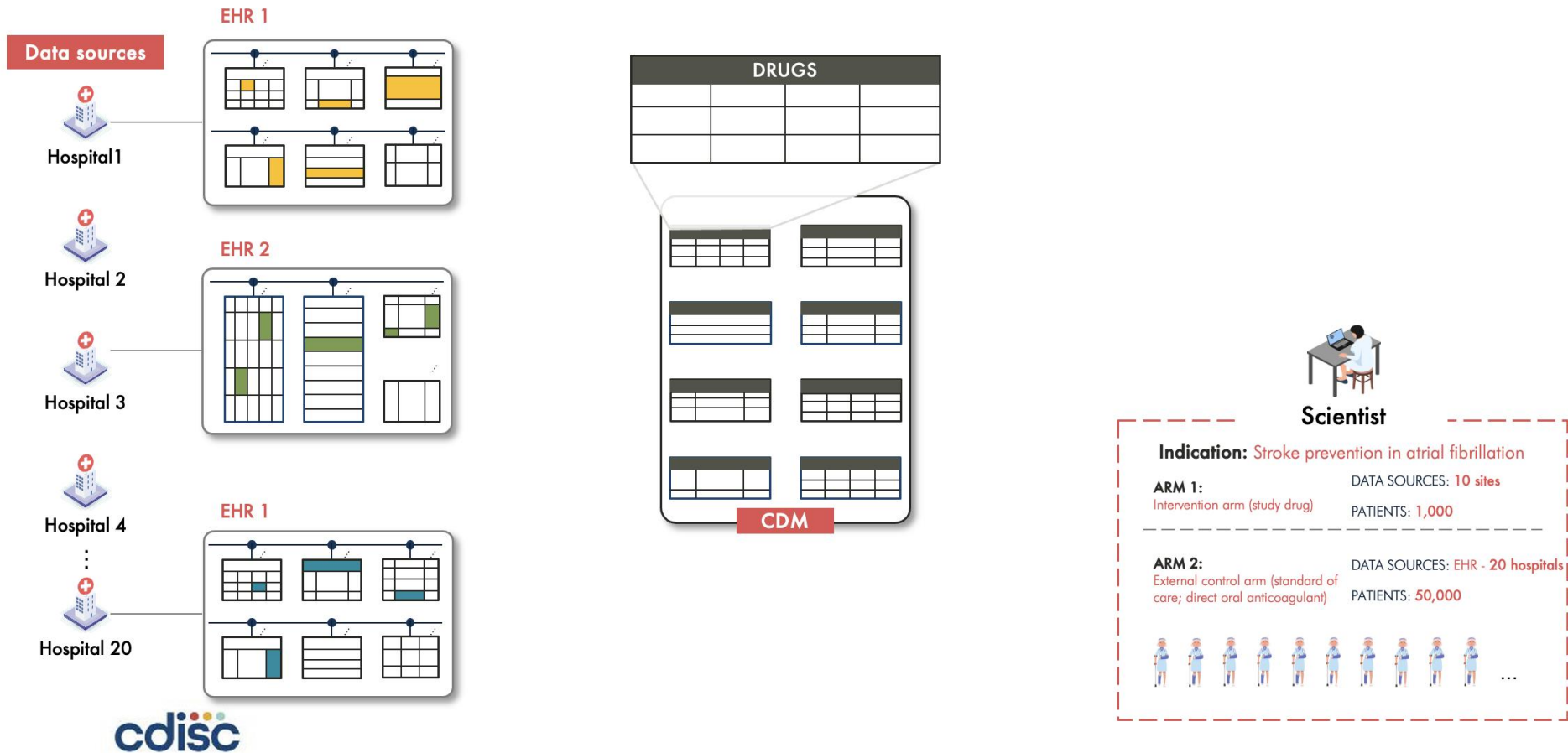
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

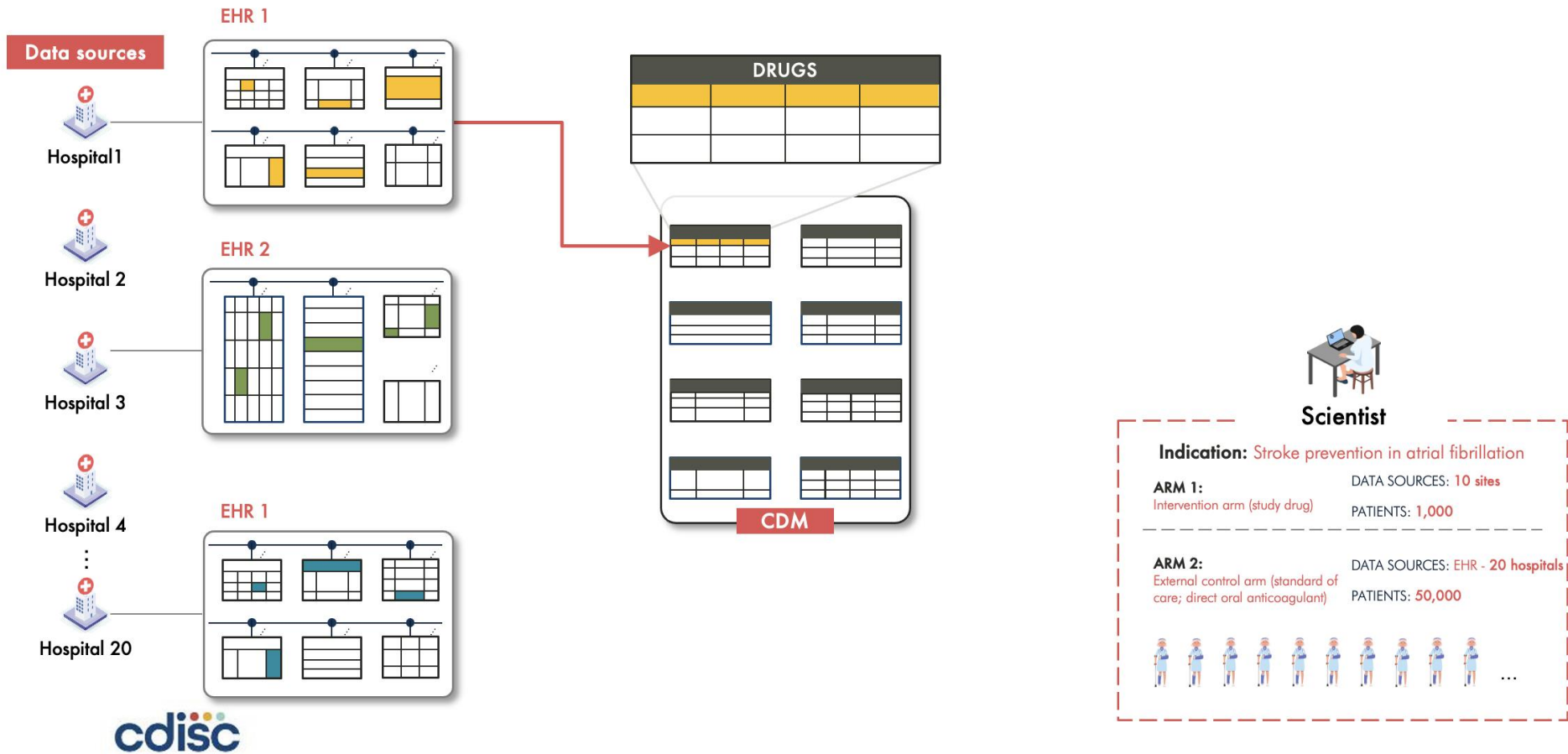
ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



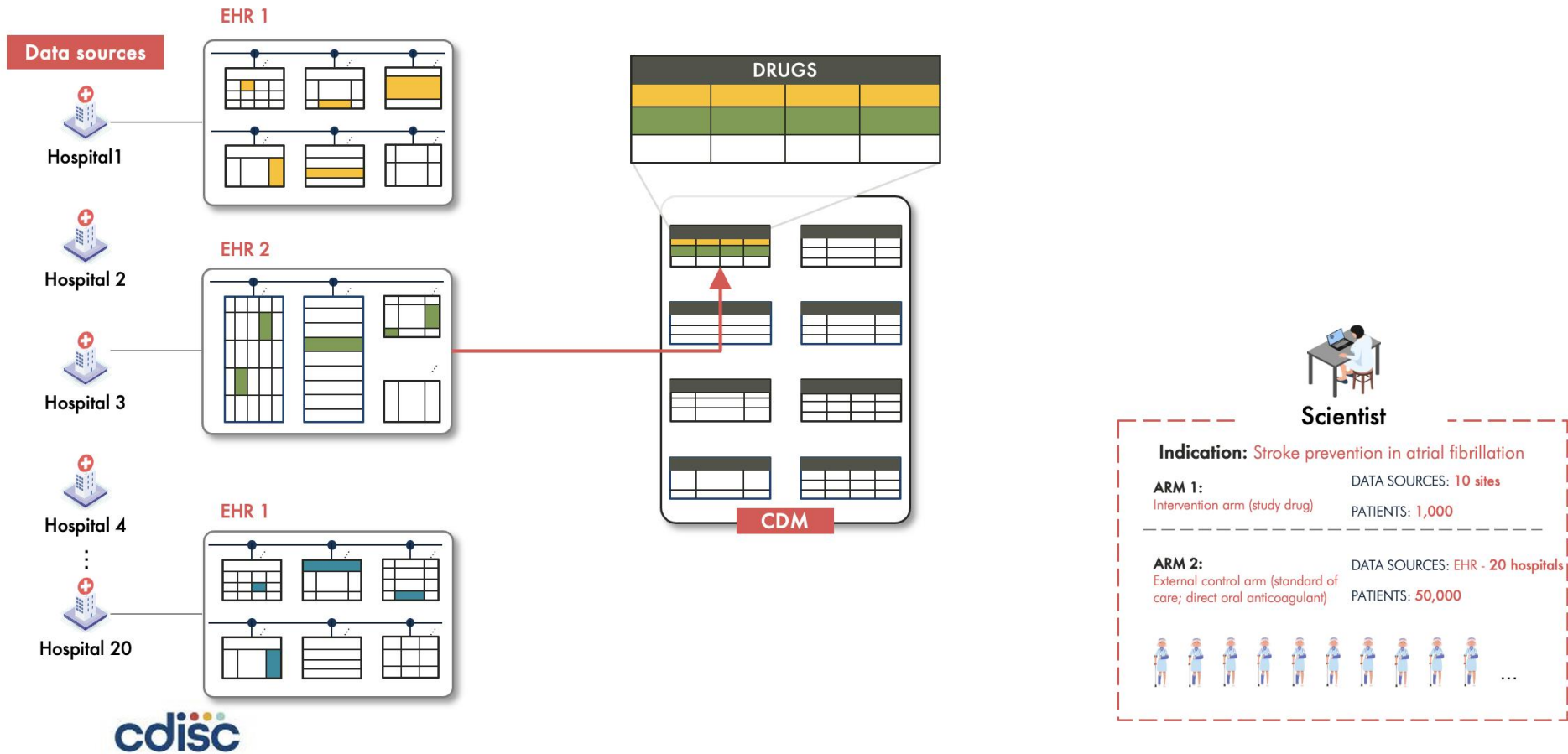
Common data models (CDM)



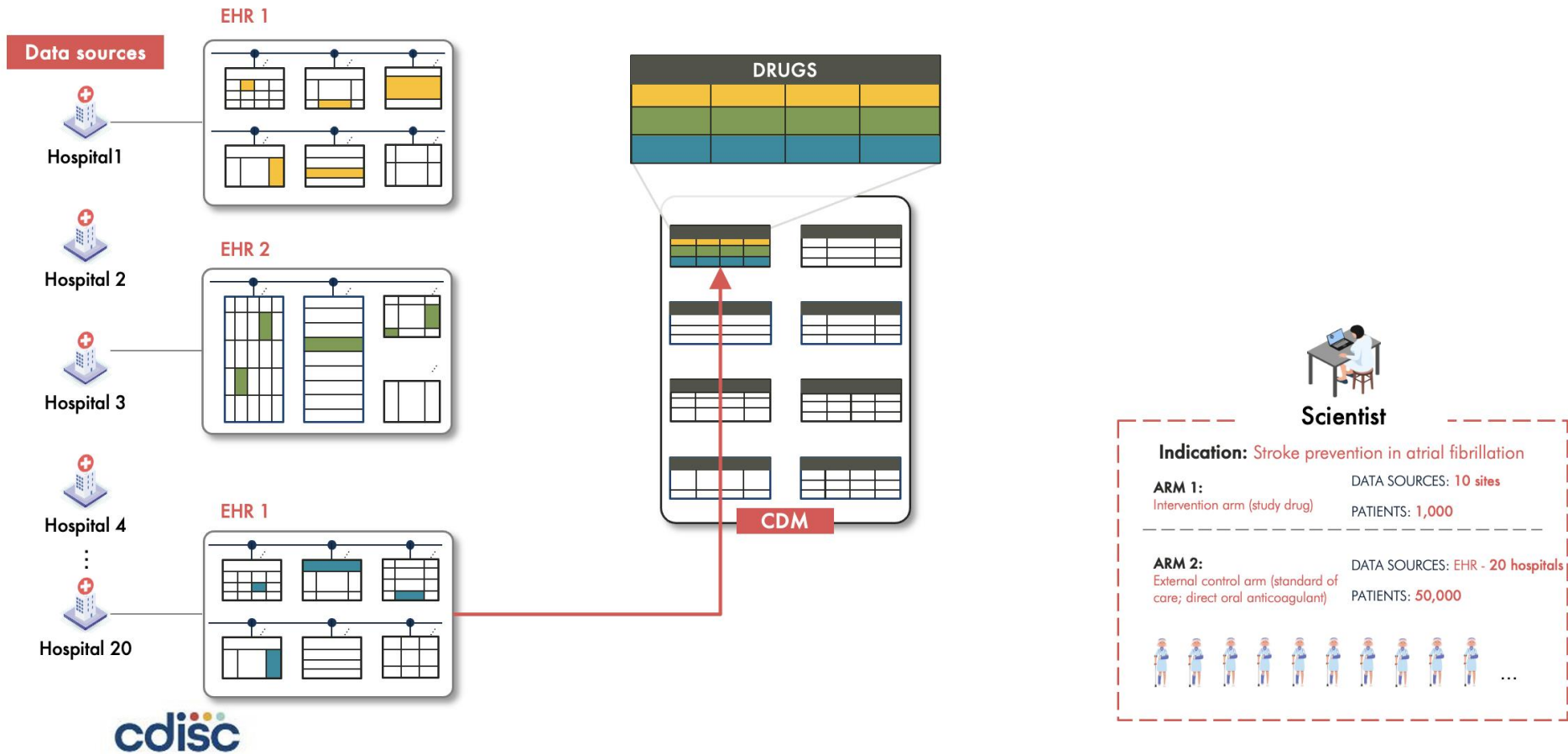
Common data models (CDM)



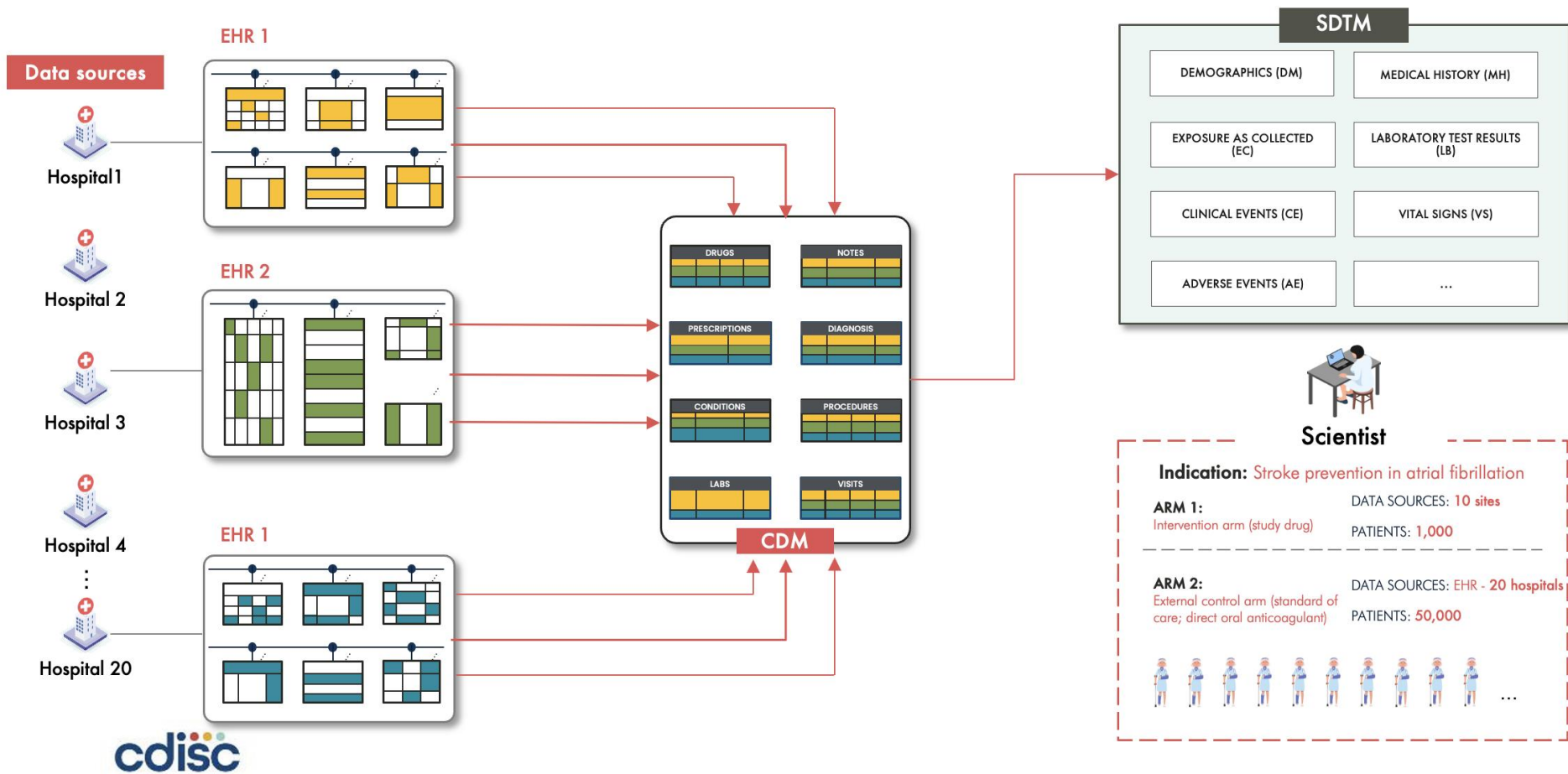
Common data models (CDM)



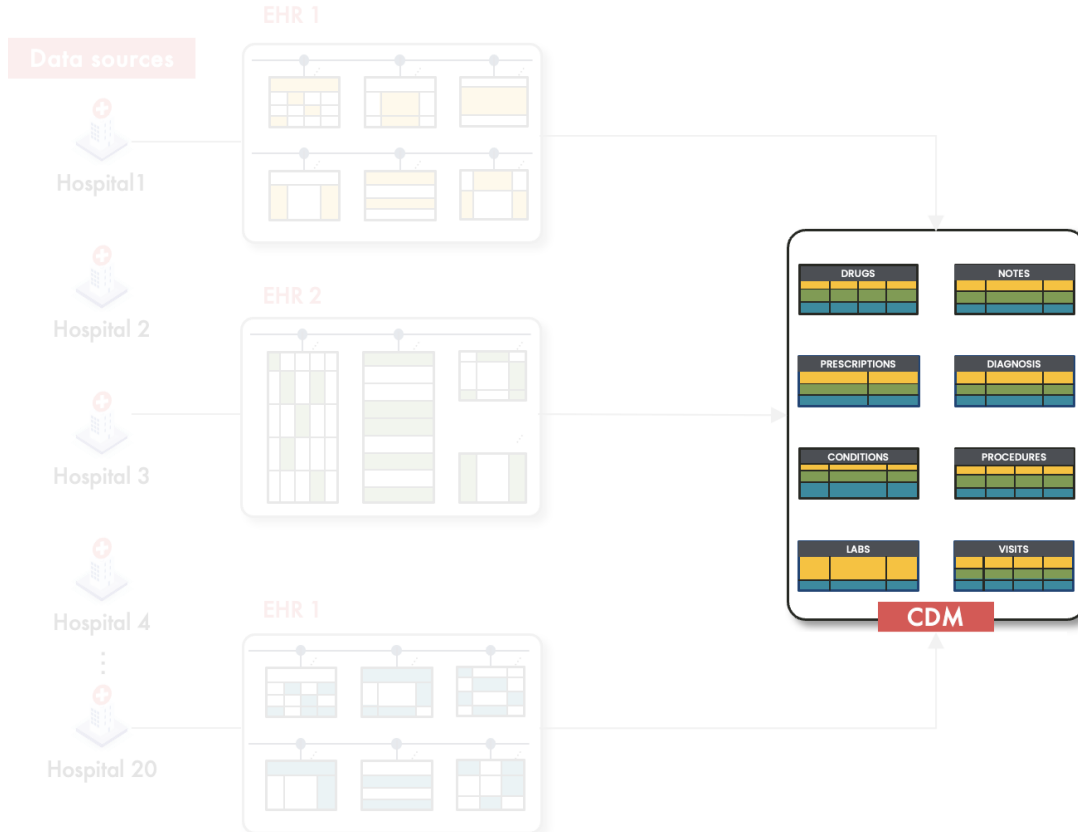
Common data models (CDM)



Common data models (CDM)



Common data models (CDM)



“

...CDMs can introduce additional **challenges** for researchers to consider... (R 331)

...combining many data sources, especially with the addition of data transformation into a CDM, adds a layer of **complexity** that should be considered... (R 324-325)

...

”



Scientist

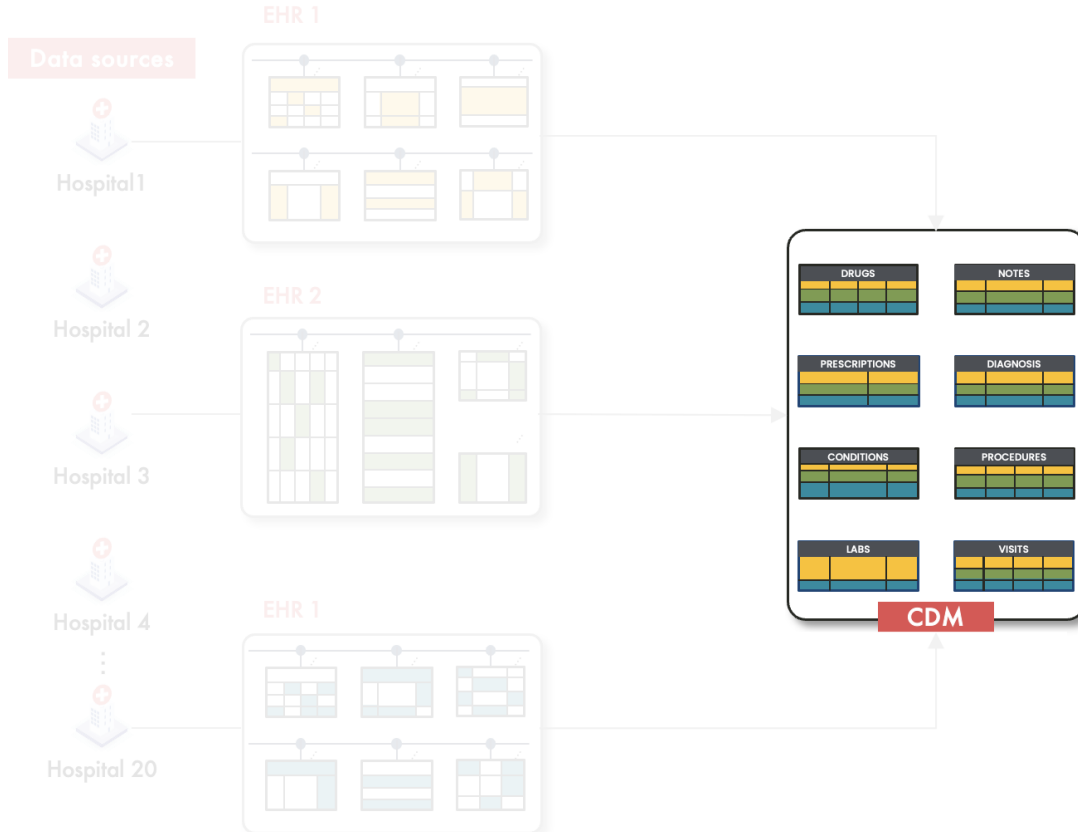
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



Common data models (CDM)



“

...CDMs can introduce additional **challenges** for researchers to consider... (R 331)

...combining many data sources, especially with the addition of data transformation into a CDM, adds a layer of **complexity** that should be considered.... (R 324-325)

“

...data in CDM-driven networks rarely contain all of the source information present at the individual health care sites, and the data elements chosen for a given CDM network may not be sufficient for all research purposes or questions... (R 337-340)



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



Common data models (CDM)

EHR 1

CDM 1

Comparing automated vs. manual data collection for COVID-specific medications from electronic health records
Andrew L. Yin^{1,2,3,4,5}, Winston L. Guo^{4,1}, Egan T. Sholle^{4,1}, Maragala Baij
Parag Govai^{4,6}, Assem Jabri⁴, Han A. Li^{4,8}, Laura C. Prothero⁵, Graham
COVID-19 Data Abstraction Consortium, ⁶Monika M. Safford⁴, Thom

Choosing Among Common Data Models for Real-World Data Analyses Fit for Making Decisions About the Effectiveness of Medical Products
Sebastian Schreyer^{1,2}, Jeff S. Brown², Andrew Bate³, Gianluca Trifiro⁴, Dorothee B Bartels³

A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data
Michael G. Kahn, MD, PhD¹, Tiffany J. Callahan, MPH¹, Juliana Barnard, MA¹, Alan E. Bauck¹, Jeff Brown, PhD^{1,2}
Bruce N. Davidson, PhD¹, Hossein Estiri, PhD¹, Garsten Goerg, PhD¹, Erin
Stewart G. Johnson, MS^{1,2}, Siaw-Teng Liaw, MBBS, PhD, FRACGP, FACH¹,
Daniella Meeker, PhD¹, Tsan C. Ong, PhD¹, Patrick Byers, PhD¹, Ning Shen
Chunhua Wang, PhD, FACMI^{1,2}, Meredith N. Zozus, PhD^{1,2} and Lisa Schill

Converting to a common data model: what is lost in translation? : Commentary on "fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model"
Spreek¹

An evaluation of the THIN database in the OMOP Common Data Model for active drug safety surveillance
Xiaofeng Zhou¹, Sundaresan Murugesan, Harshvinder Bhullar, Qing Liu, Bing Cai,
Chuck Wentworth, Andrew Bate

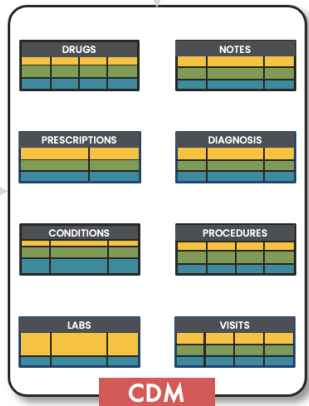
Fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model
Amy Matcho¹, Patrick Ryan, Daniel File, Christian Reich

Extract, transform, load framework for the conversion of health databases to OMOP
Juan C. Quiroz, Conceptualization, Methodology, Writing - original draft, Writing - review & editing, ¹Timothy
Conceptualization, Methodology, Software, Writing - re
Writing - review & editing, ¹ Angus Ritchie, Conceptual
review & editing, ¹ and Blanca Gallego, Conceptualizat
Validation, Writing - review & editing¹

Improving Data Quality in Clinical Research Informatics Tools
Ahmed AbuHalimeh¹

A Framework for Classification of Electronic Health Data Extraction-Transformation-Loading Challenges in Data Network Participation
Toan Ong, PhD,¹ Rosina Pradhananga,¹

A Semantic Transformation Methodology for the Secondary Use of Observational Healthcare Data in Postmarketing Safety Studies
Anil Pacaci^{1,2}, Suat Gonul^{1,3}, Anil Sinaci¹, Mustafa Yuksek¹, Golice B Laleci Erturkmen¹



“

...CDMs can introduce additional **challenges** for researchers to consider...(R 331)

...combining many data sources, especially with the addition of data transformation into a CDM, adds a layer of **complexity** that should be considered.... (R 324-325)

“

...data in CDM-driven networks rarely contain all of the source information present at the individual health care sites, and the data elements chosen for a given CDM network may not be sufficient for all research purposes or questions...(R 337-340)



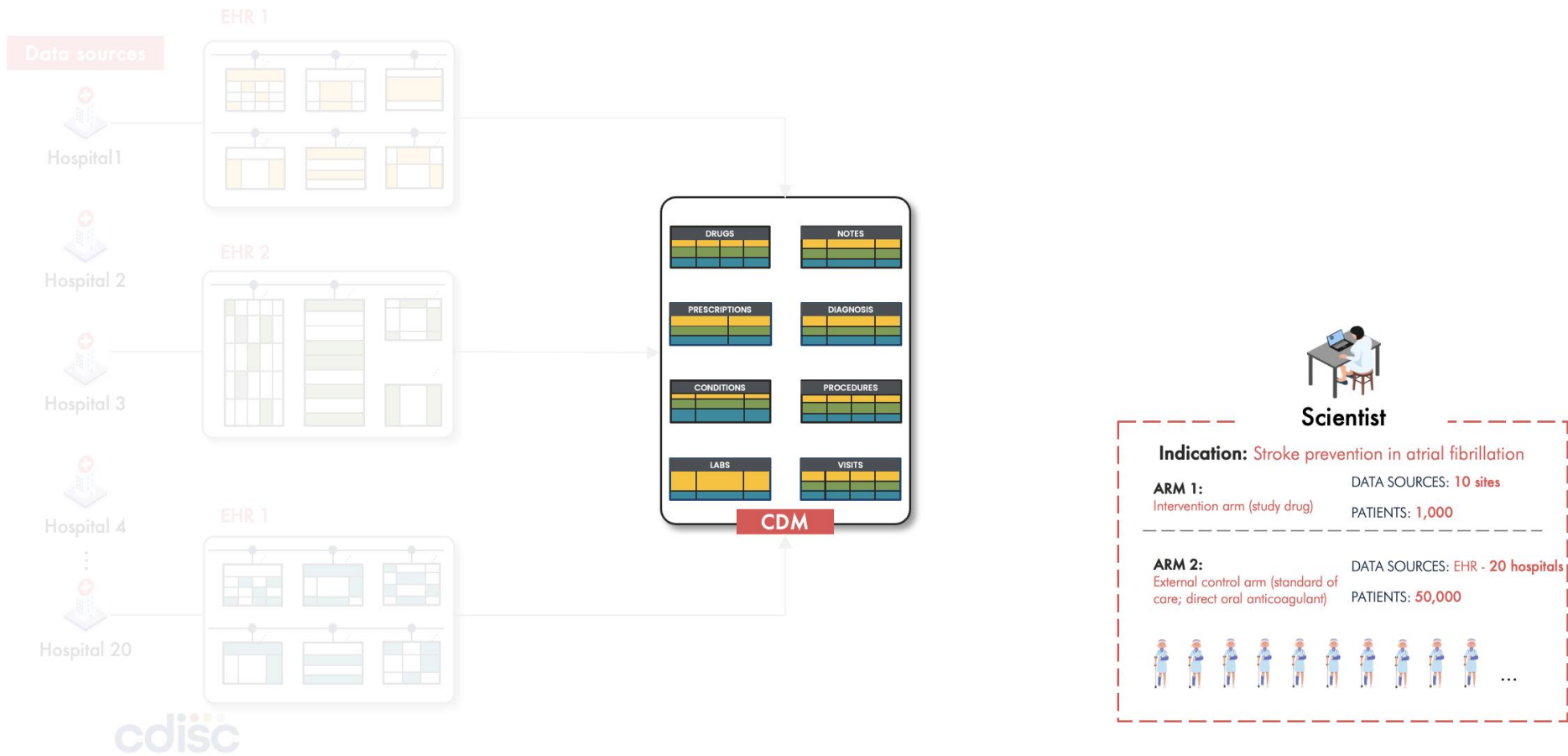
Scientist

Indication: Stroke prevention in atrial fibrillation

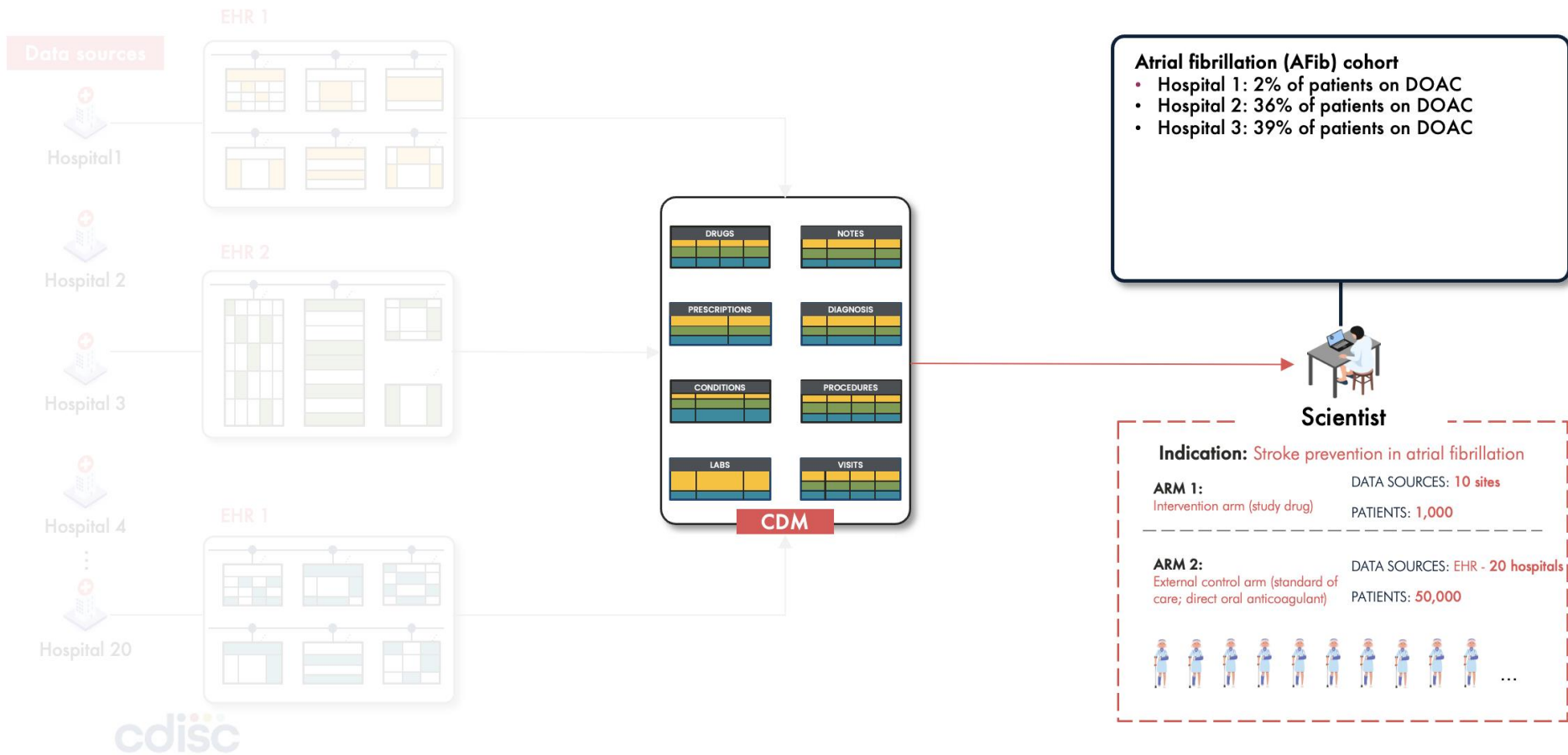
ARM 1: Intervention arm (study drug) DATA SOURCES: 10 sites
PATIENTS: 1,000

ARM 2: External control arm (standard of care; direct oral anticoagulant) DATA SOURCES: EHR - 20 hospitals
PATIENTS: 50,000

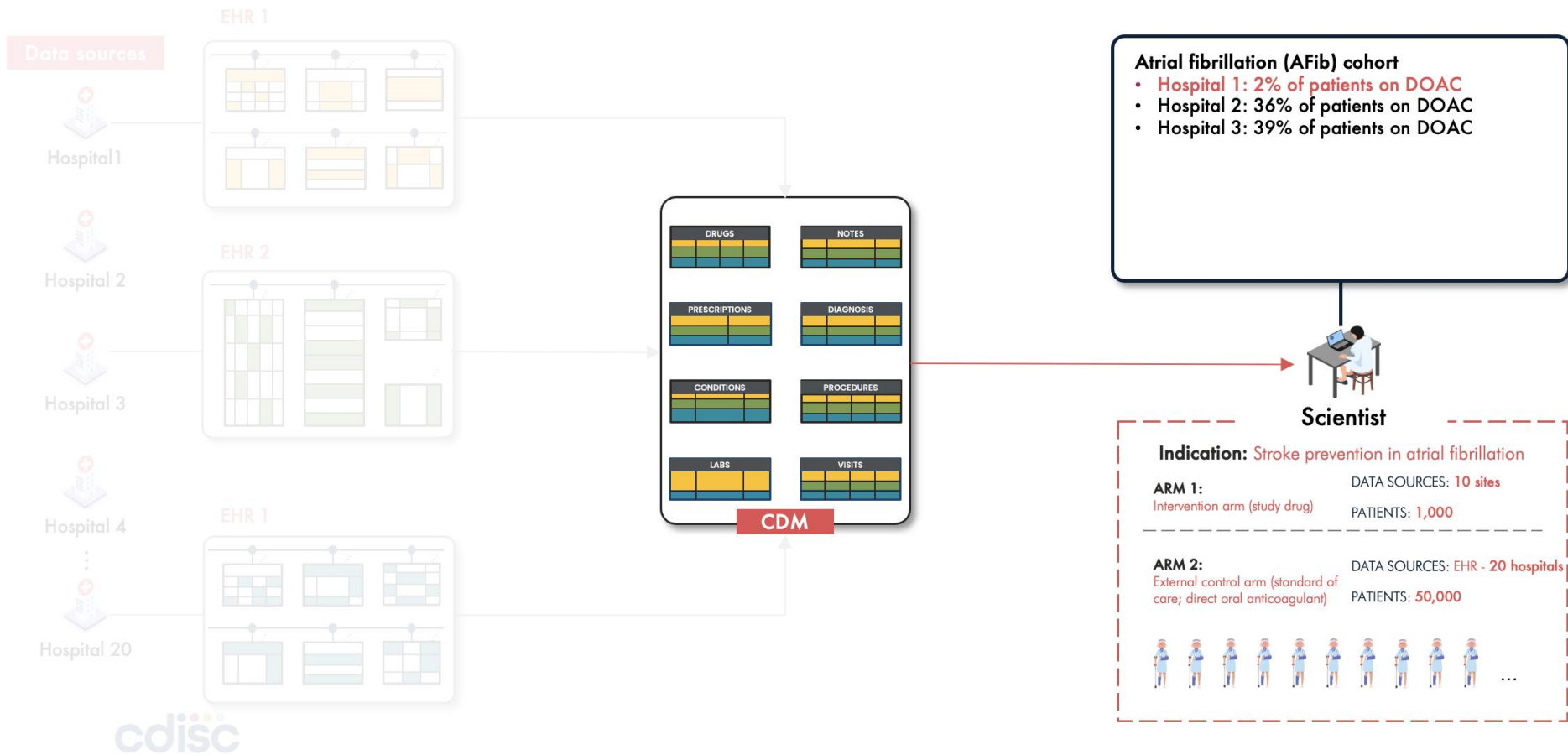
Data loss in common data models



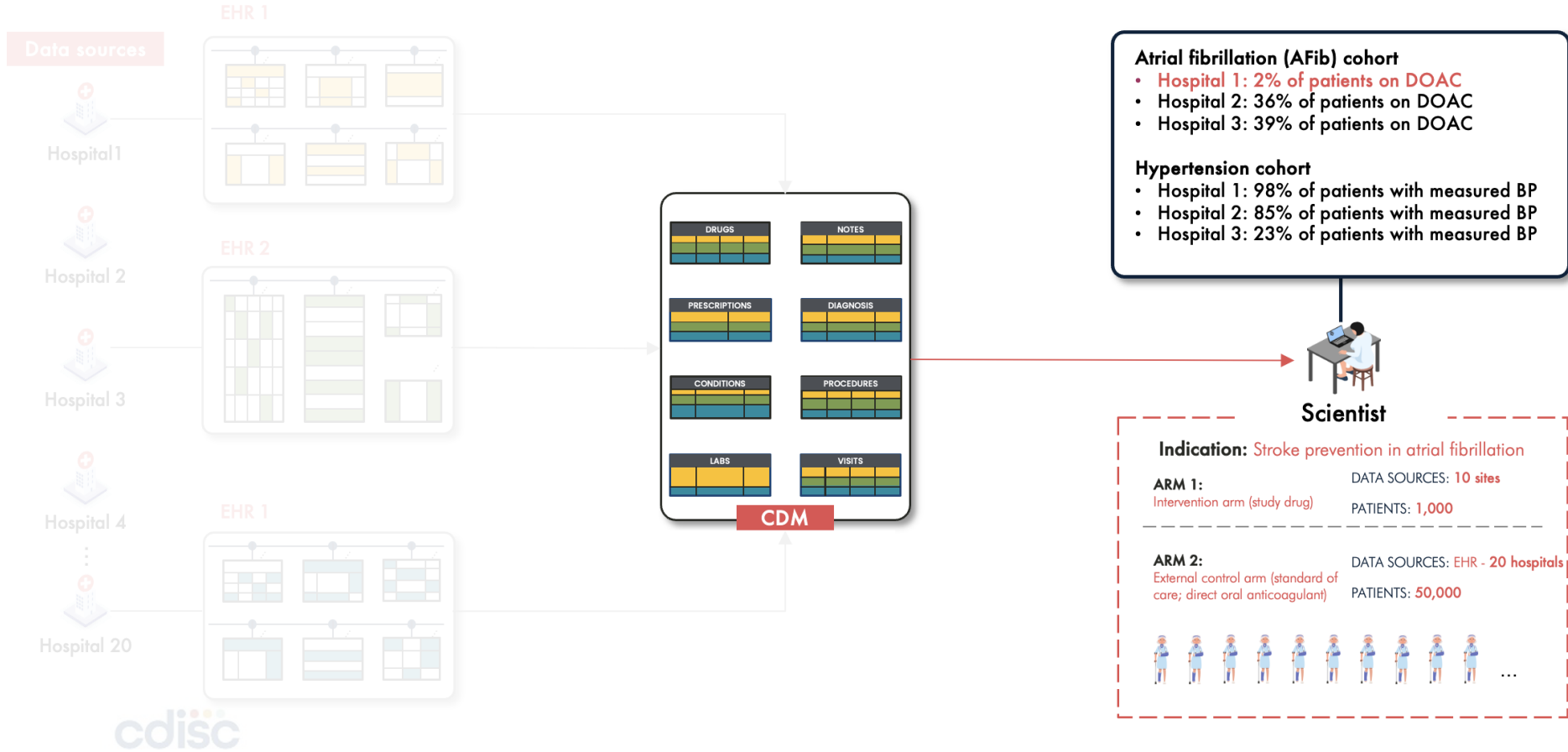
Data loss in common data models



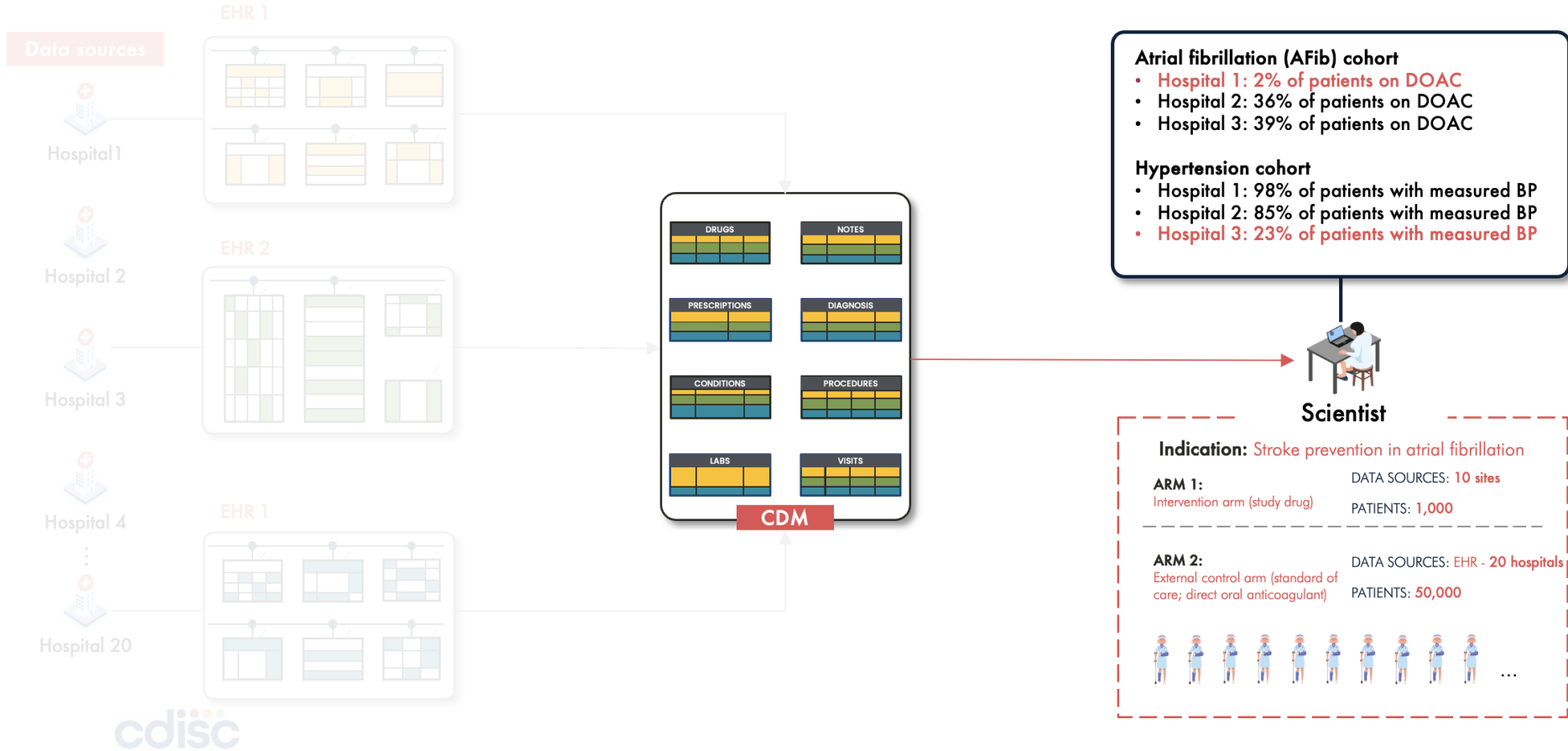
Data loss in common data models



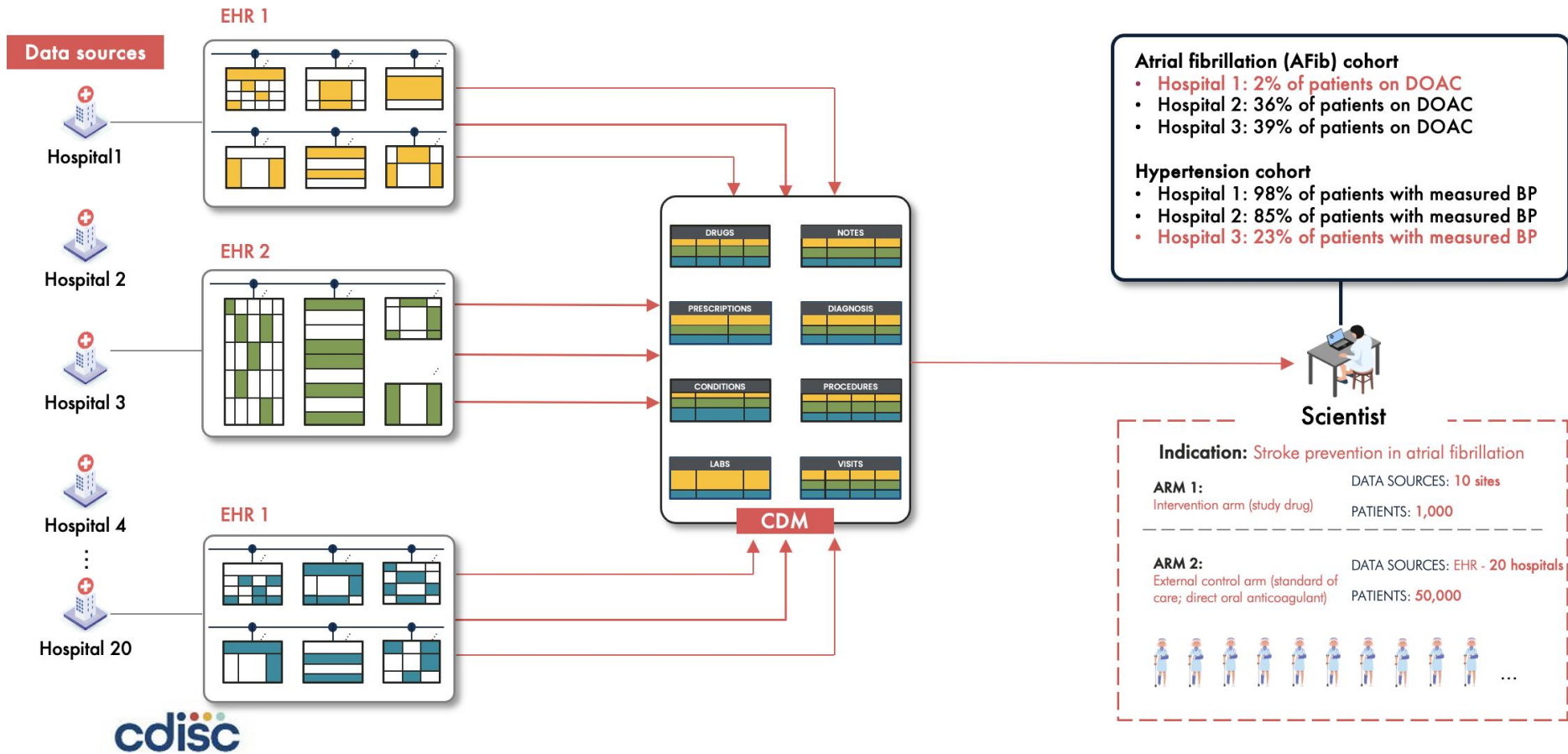
Data loss in common data models



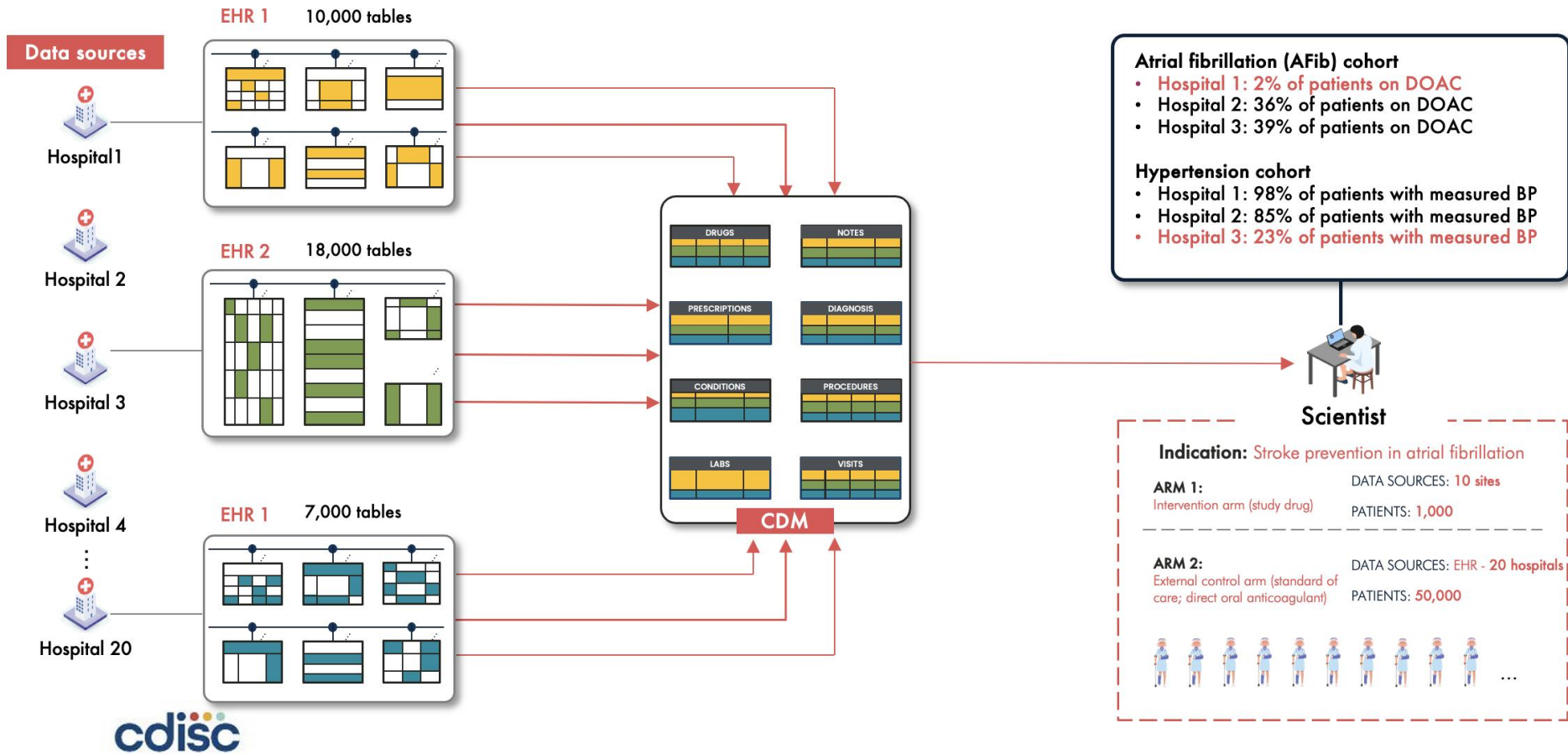
Data loss in common data models



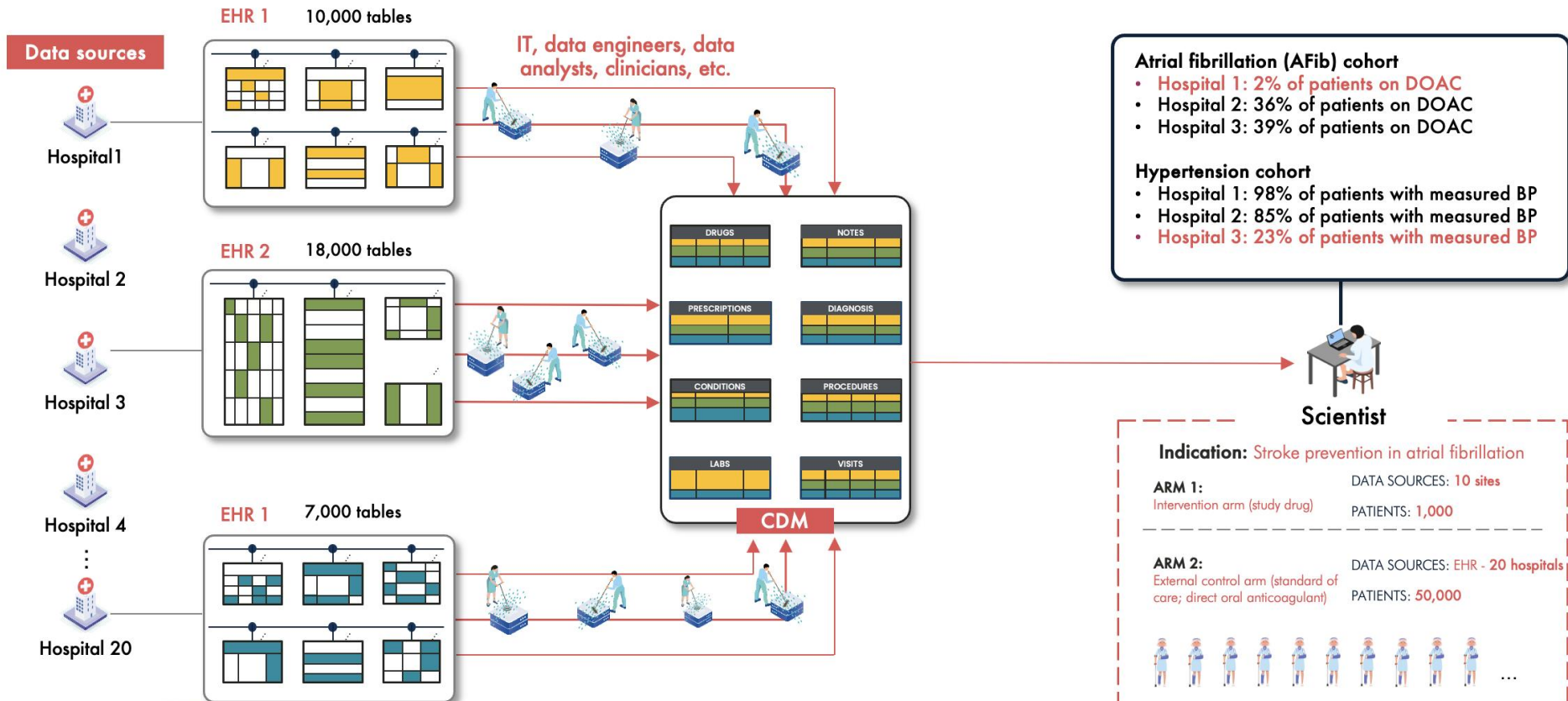
Data loss in common data models



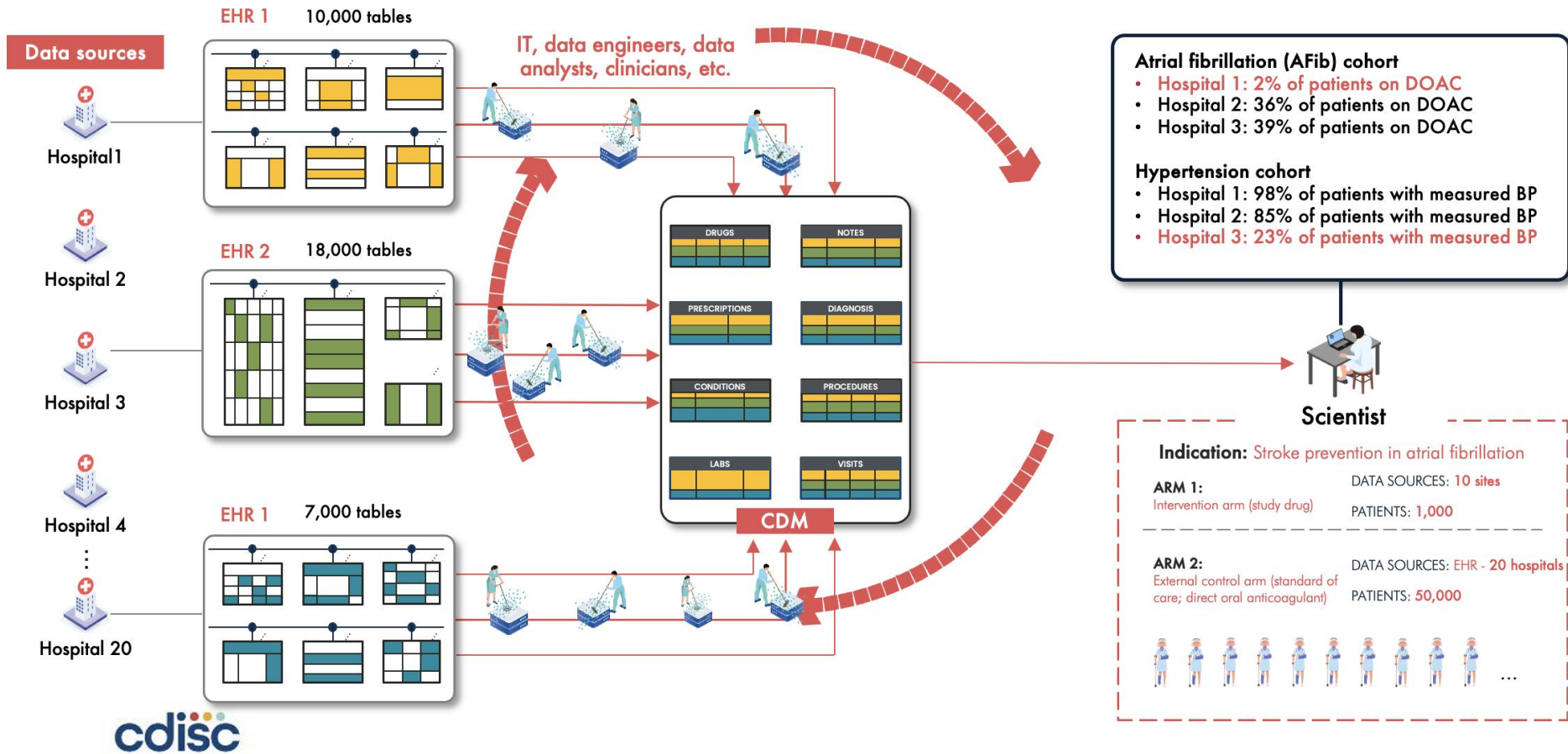
Data loss in common data models



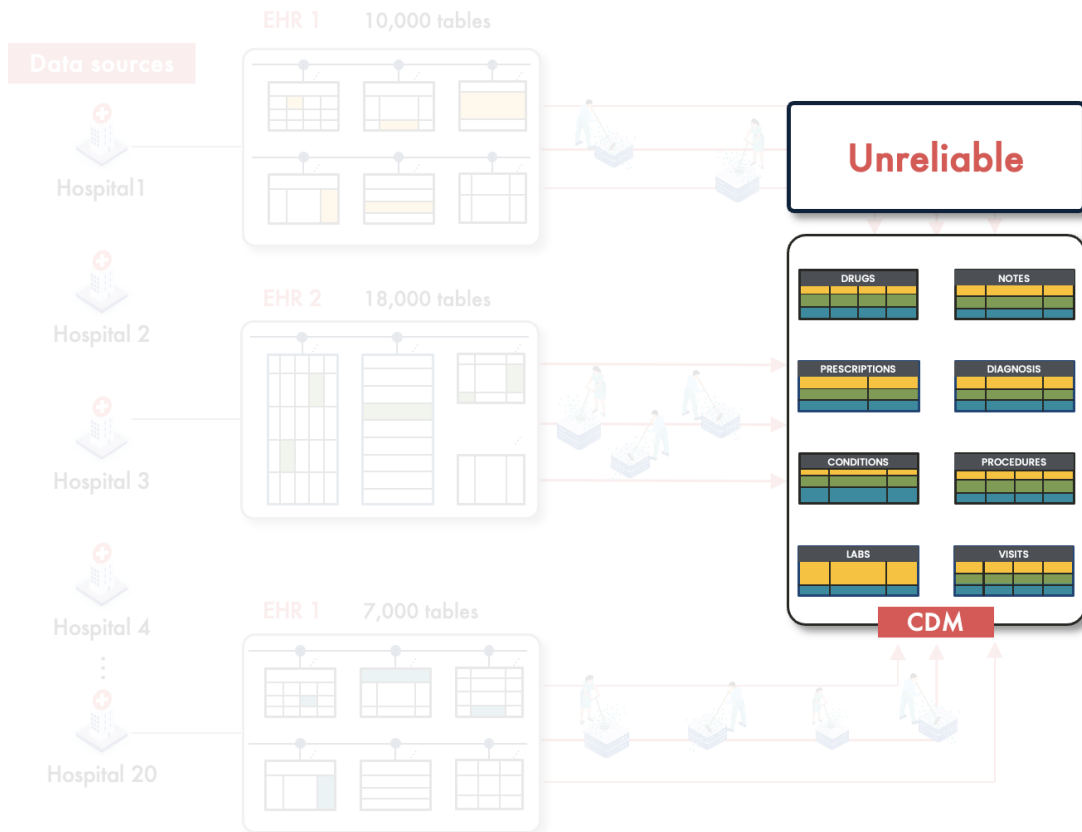
Data loss in common data models



Data loss in common data models



Data loss in common data models



Atrial fibrillation (AFib) cohort

- Hospital 1: 2% of patients on DOAC
- Hospital 2: 36% of patients on DOAC
- Hospital 3: 39% of patients on DOAC

Hypertension cohort

- Hospital 1: 98% of patients with measured BP
- Hospital 2: 85% of patients with measured BP
- Hospital 3: 23% of patients with measured BP



Scientist

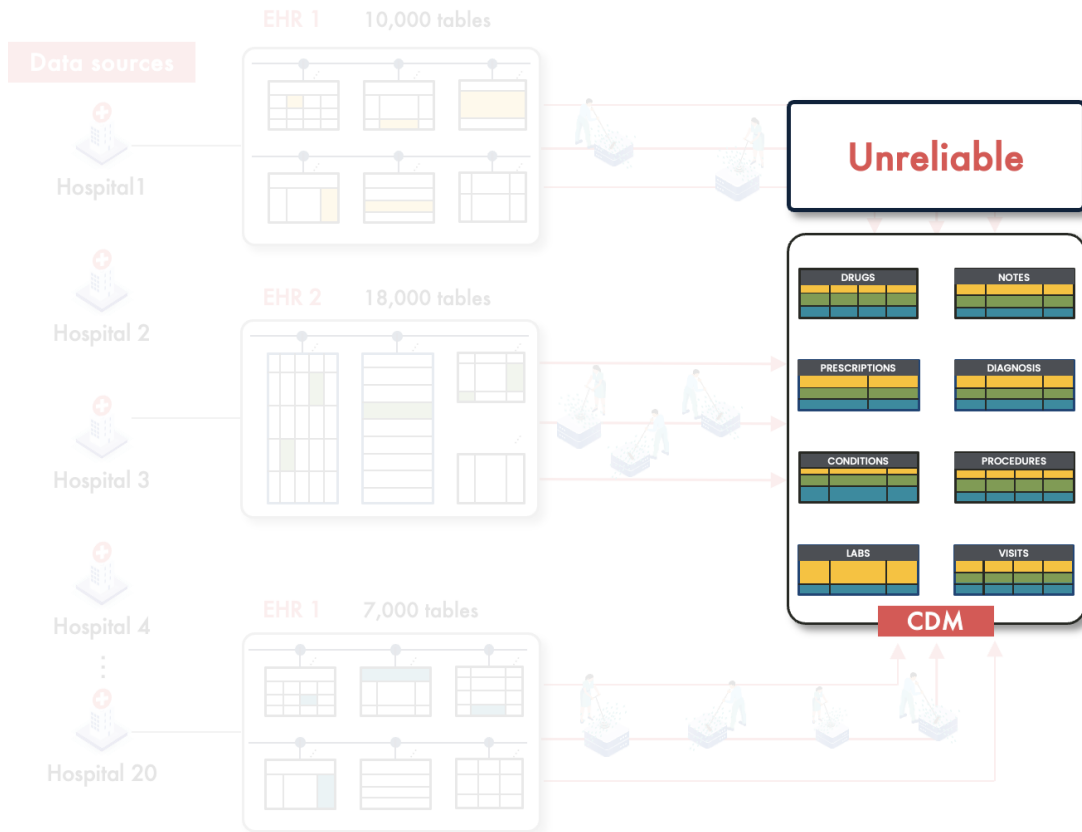
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



Data loss in common data models



Atrial fibrillation (AFib) cohort

Hospital 1: 2% of patients on DOAC

I don't know my data, and I cannot trust it

- Hospital 1: 98% of patients with measured BP
- Hospital 2: 85% of patients with measured BP
- Hospital 3: 23% of patients with measured BP



Scientist

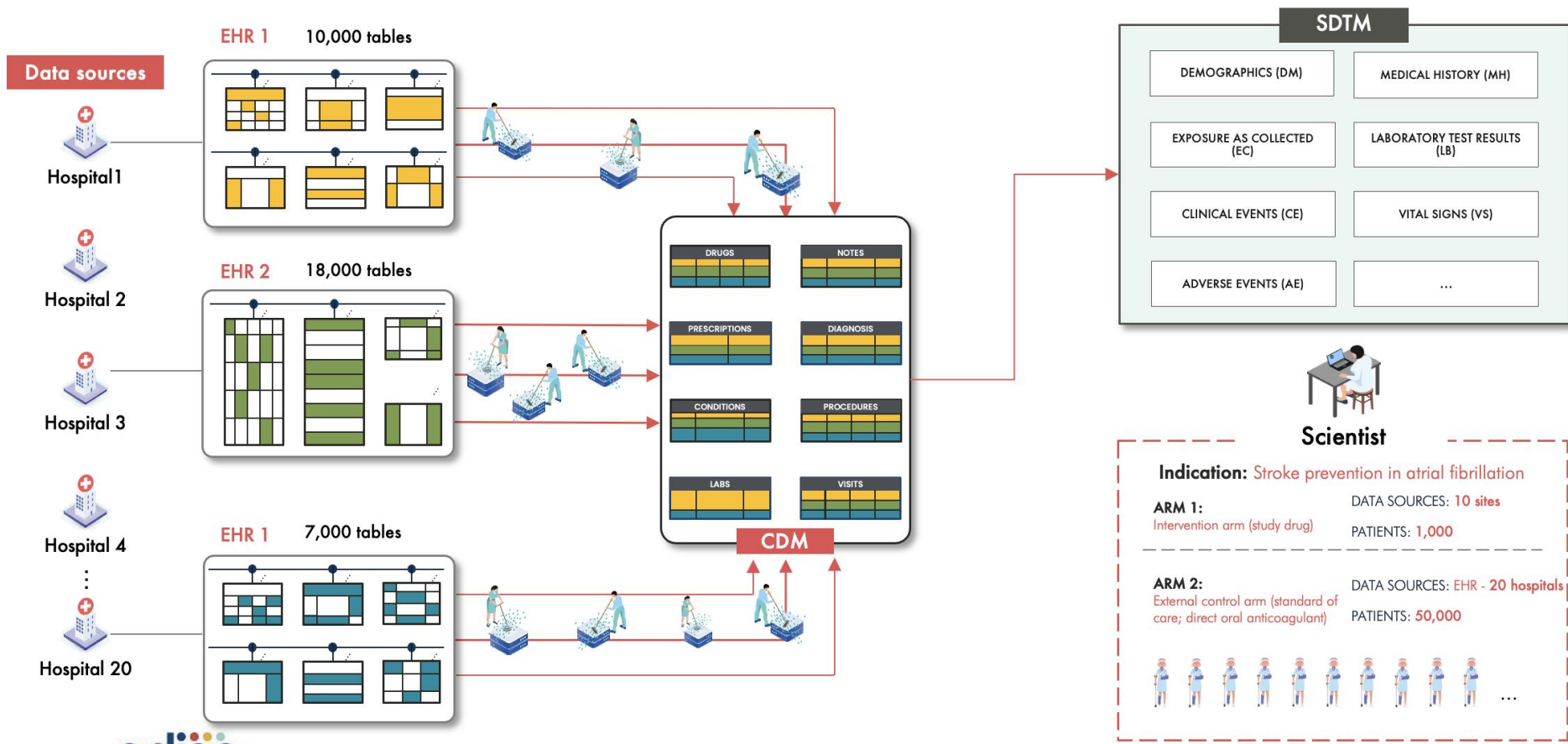
Indication: Stroke prevention in atrial fibrillation

ARM 1: DATA SOURCES: 10 sites
Intervention arm (study drug) PATIENTS: 1,000

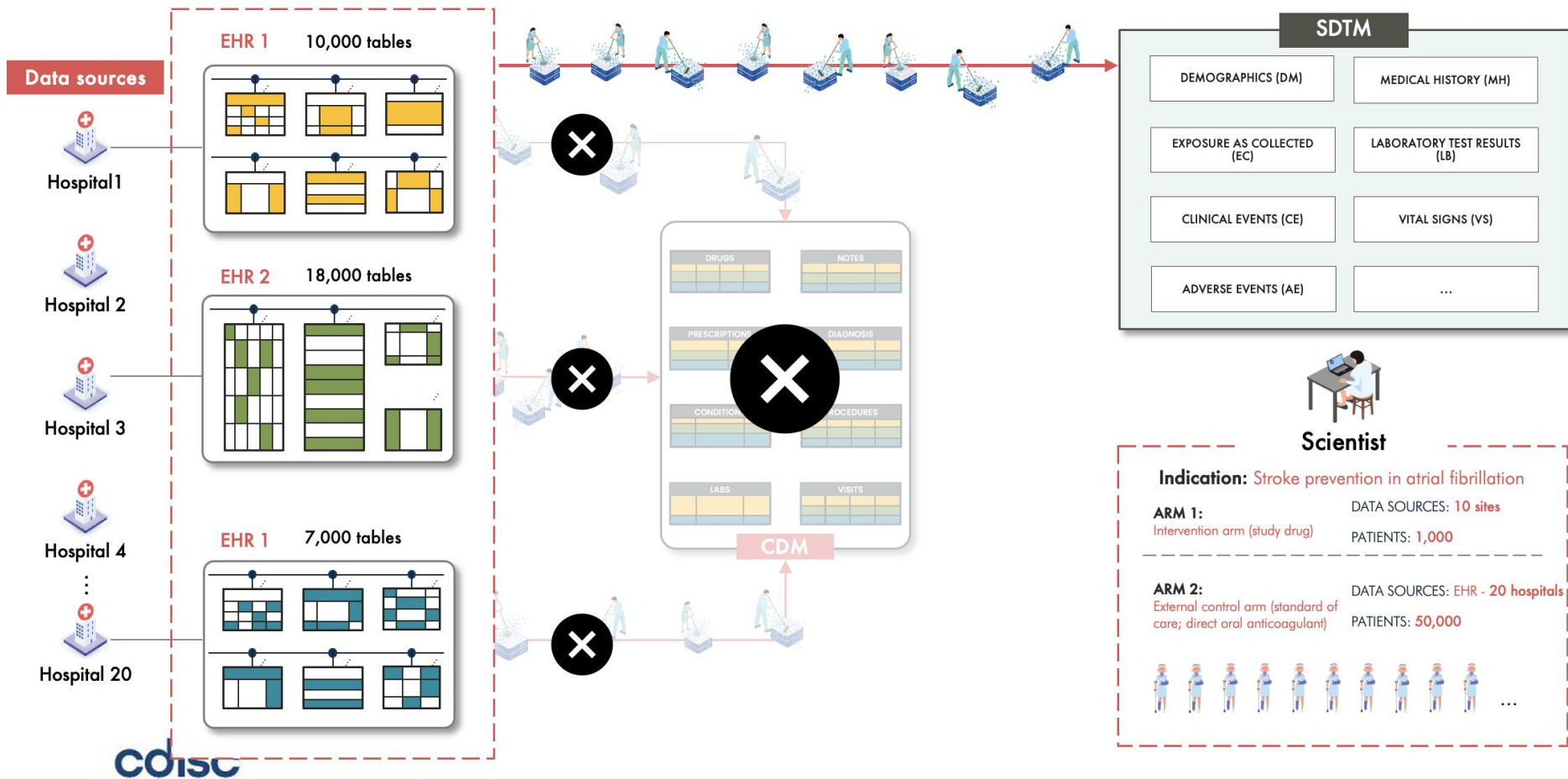
ARM 2: DATA SOURCES: EHR - 20 hospitals
External control arm (standard of care; direct oral anticoagulant) PATIENTS: 50,000



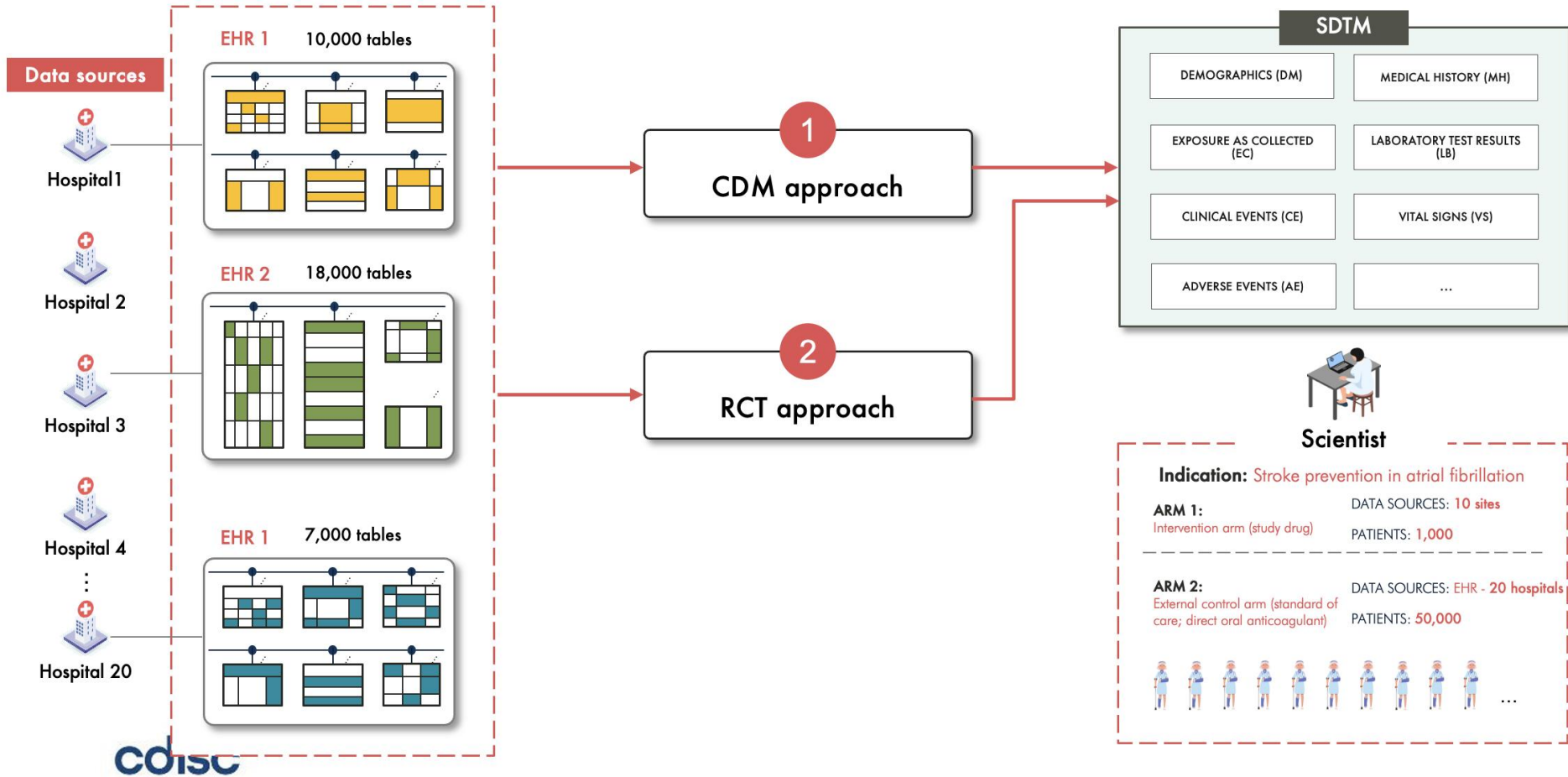
Data loss in common data models



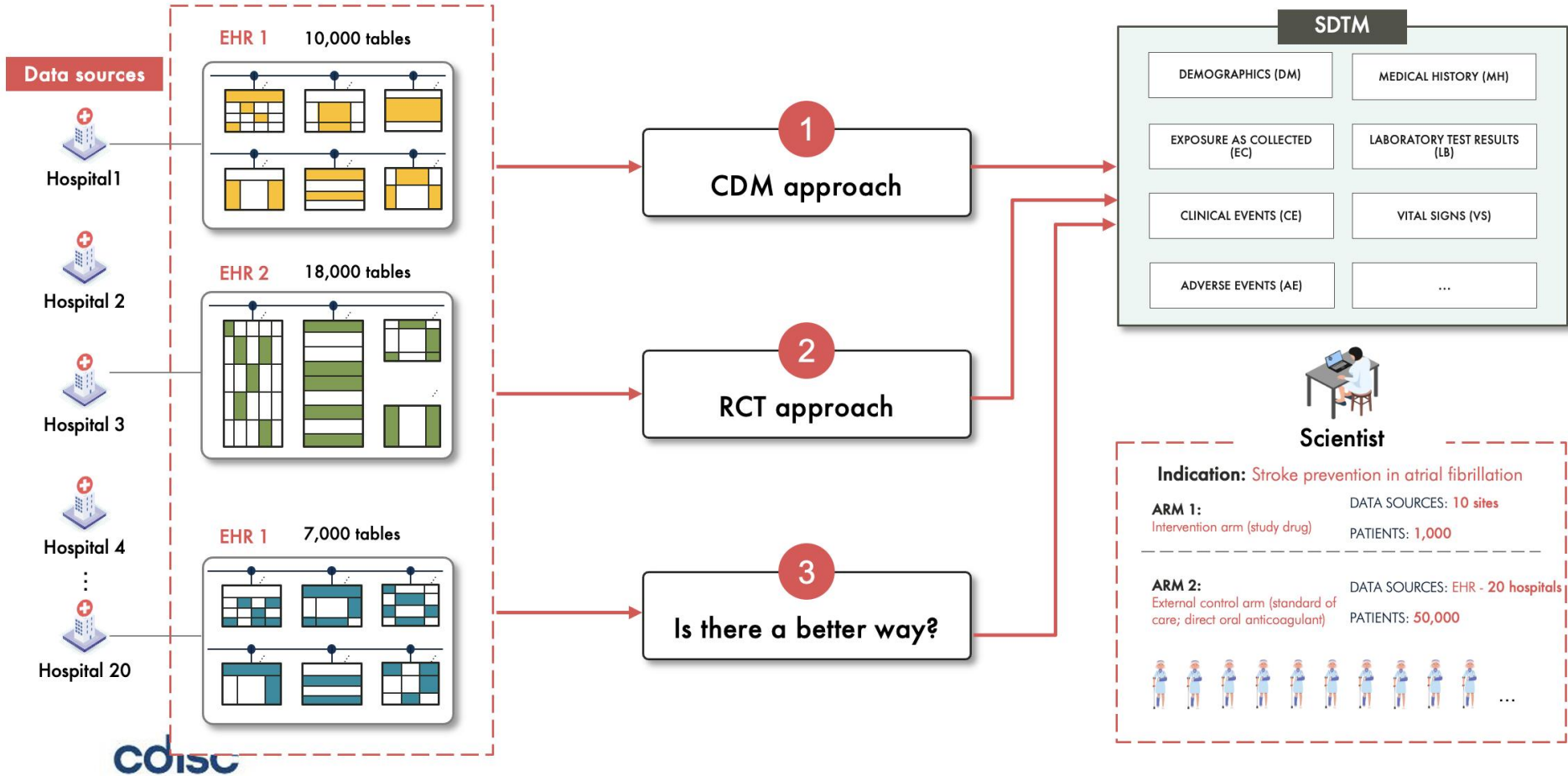
Traditional data preparation approach



Options



Options



RWD: Straight to SDTM project



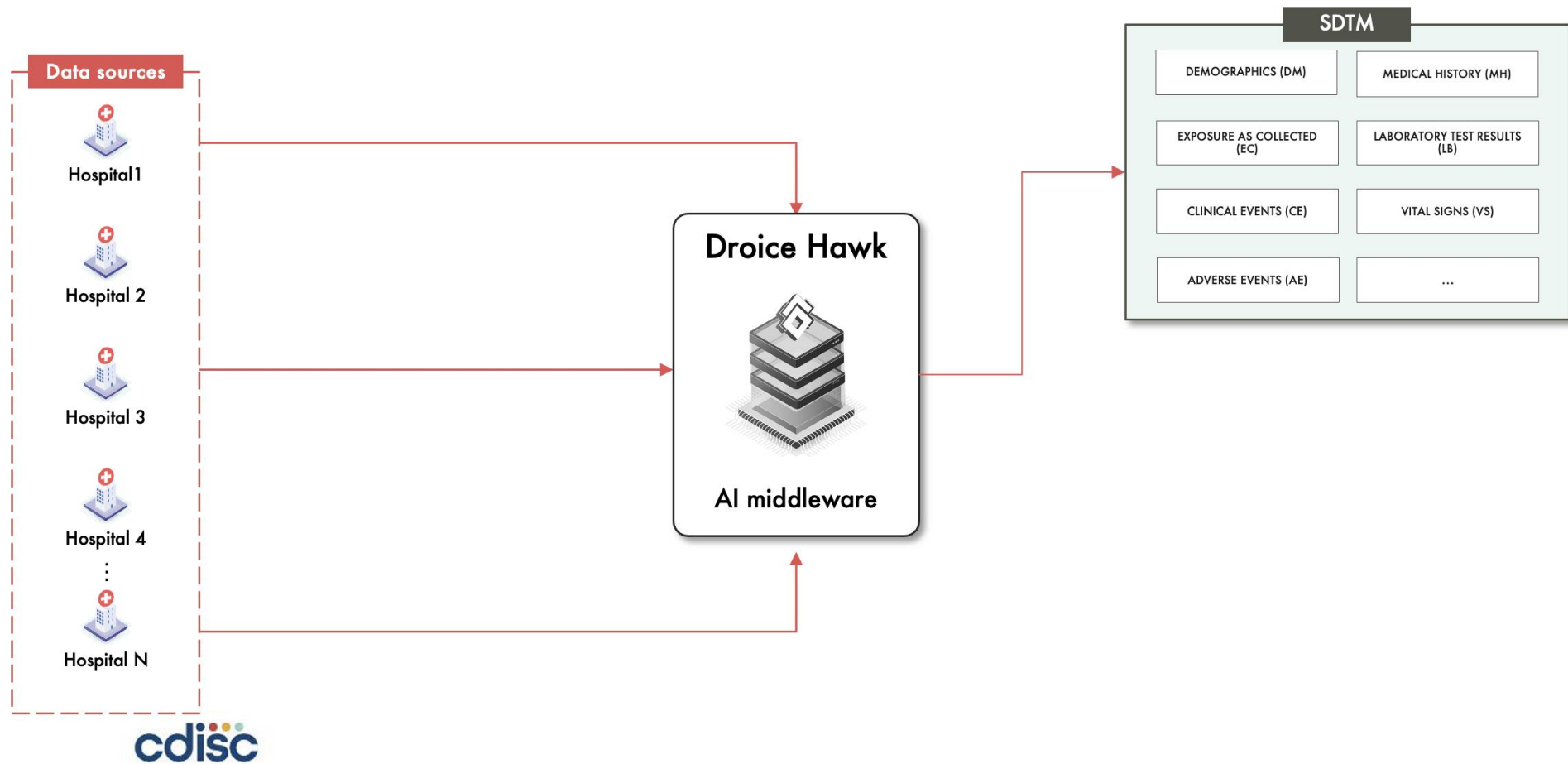
(standards expertise)



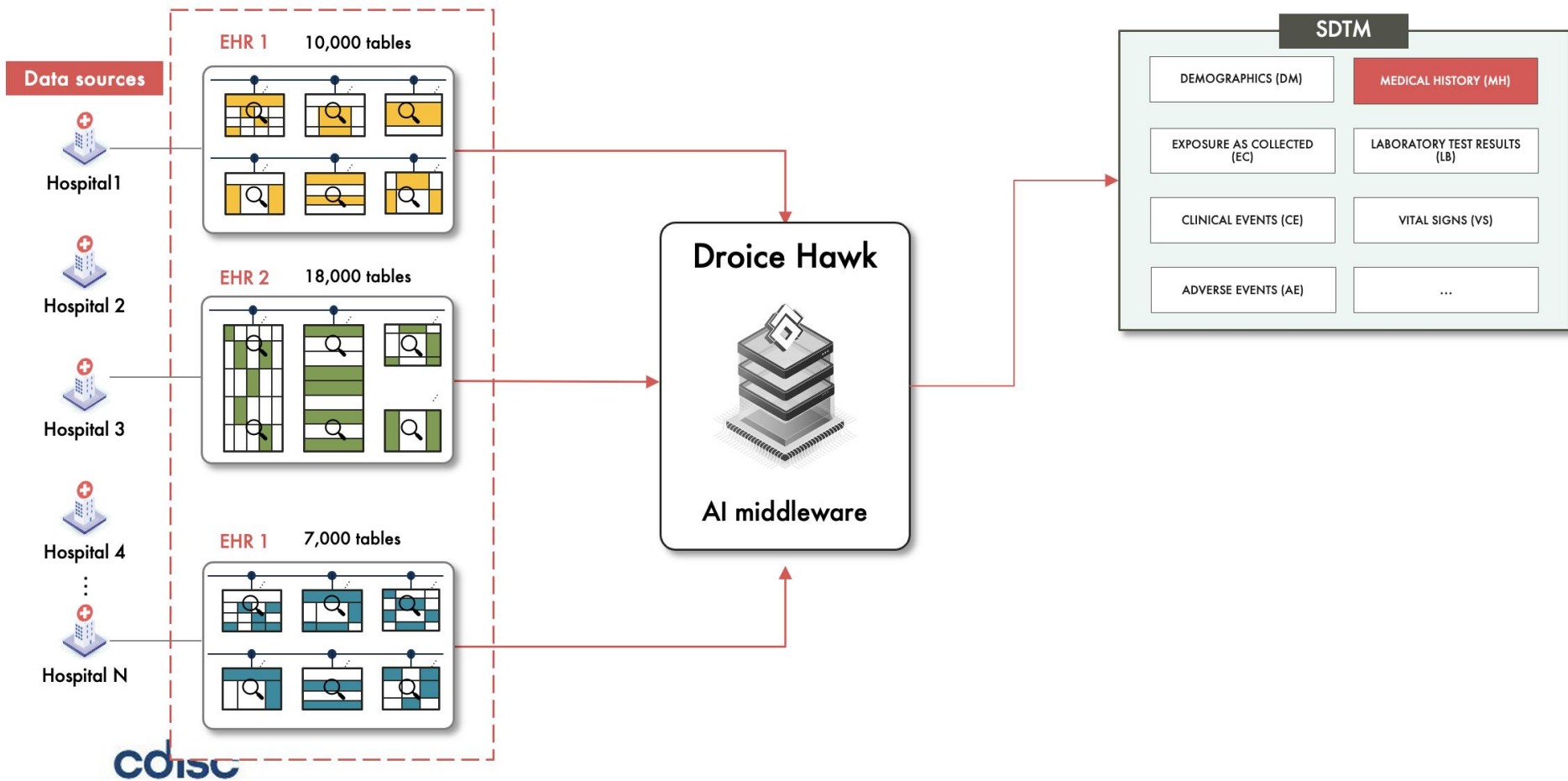
Droice Labs

(AI middleware)

RWD: Straight to SDTM project



RWD: Straight to SDTM project



RWD: Straight to SDTM project

Medical history (MH)

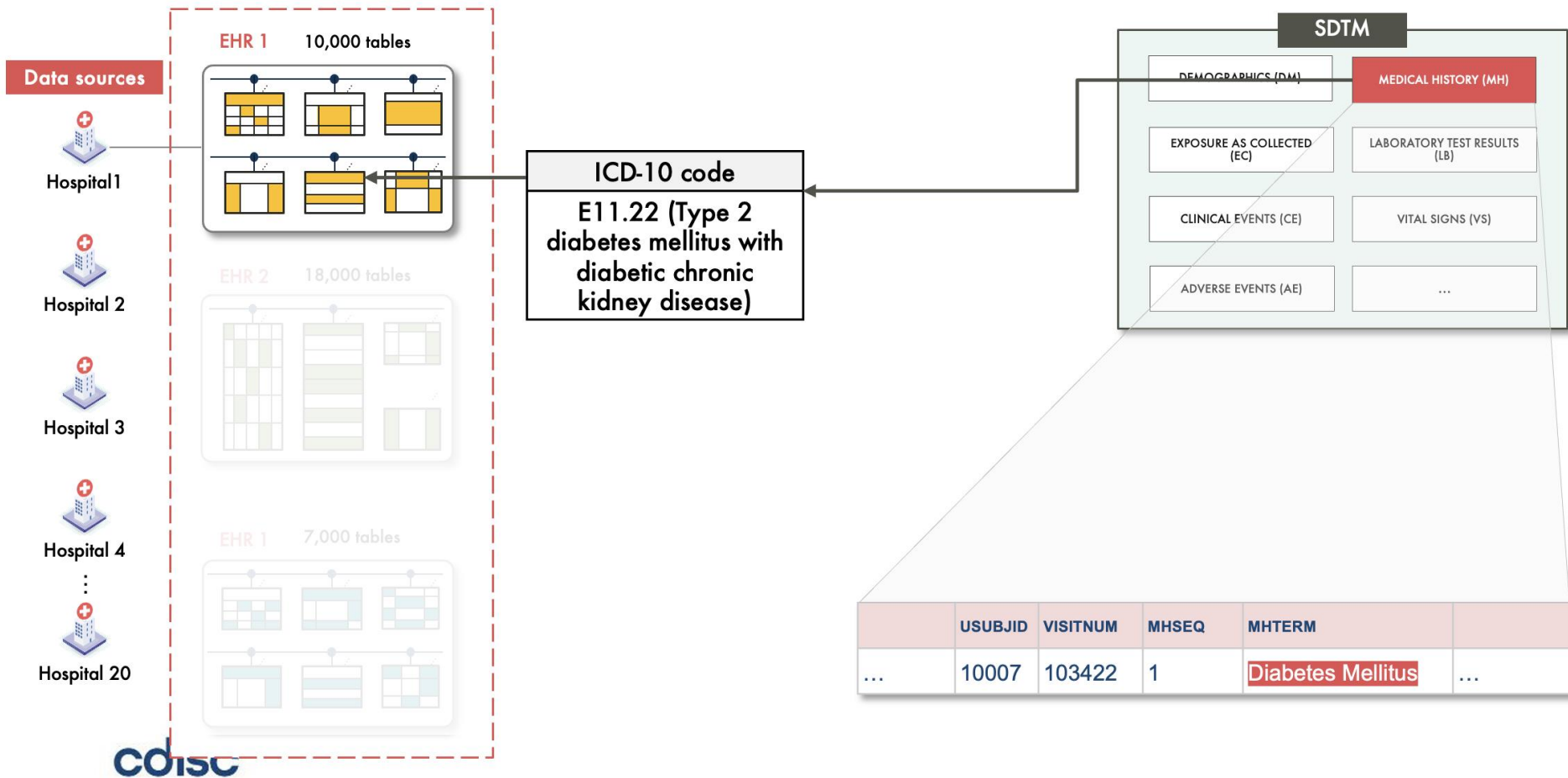
STUDYID	DOMAIN	USUBJID	VISITNUM	MHSEQ	MHTERM	MHCAT	MHDTC	MHSTDTC
EX-030	MH	10005	115407	8	Hypothyroidism	PAST MEDICAL HISTORY	2024-10-07	2024-10-07
EX-030	MH	10006	101357	1	Chronic Heart Failure	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	2	Hypothyroidism	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	3	History Of Myocardial Infarction	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	4	Peripheral Arterial Disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	5	Dementia	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	6	Hyperlipidemia	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	7	Chronic kidney disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	8	Atrial Fibrillation	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	9	Coronary Artery Disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	10	Hypertension	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	107449	11	Aortic Plaque	PAST MEDICAL HISTORY	2023-07-07	2023-07-07
EX-030	MH	10006	102846	12	Chronic Obstructive Pulmonary Disease	PAST MEDICAL HISTORY	2023-07-07	2023-07-07
EX-030	MH	10007	103422	1	Diabetes Mellitus	PAST MEDICAL HISTORY	2021-03-05	2021-03-05
EX-030	MH	10007	104819	2	Hyperlipidemia	PAST MEDICAL HISTORY	2021-08-12	2021-08-12
EX-030	MH	10007	104819	3	Chronic kidney disease	PAST MEDICAL HISTORY	2023-02-28	2023-02-28
EX-030	MH	10007	104819	4	Atrial Fibrillation	PAST MEDICAL HISTORY	2023-02-28	2023-02-28

RWD: Straight to SDTM project

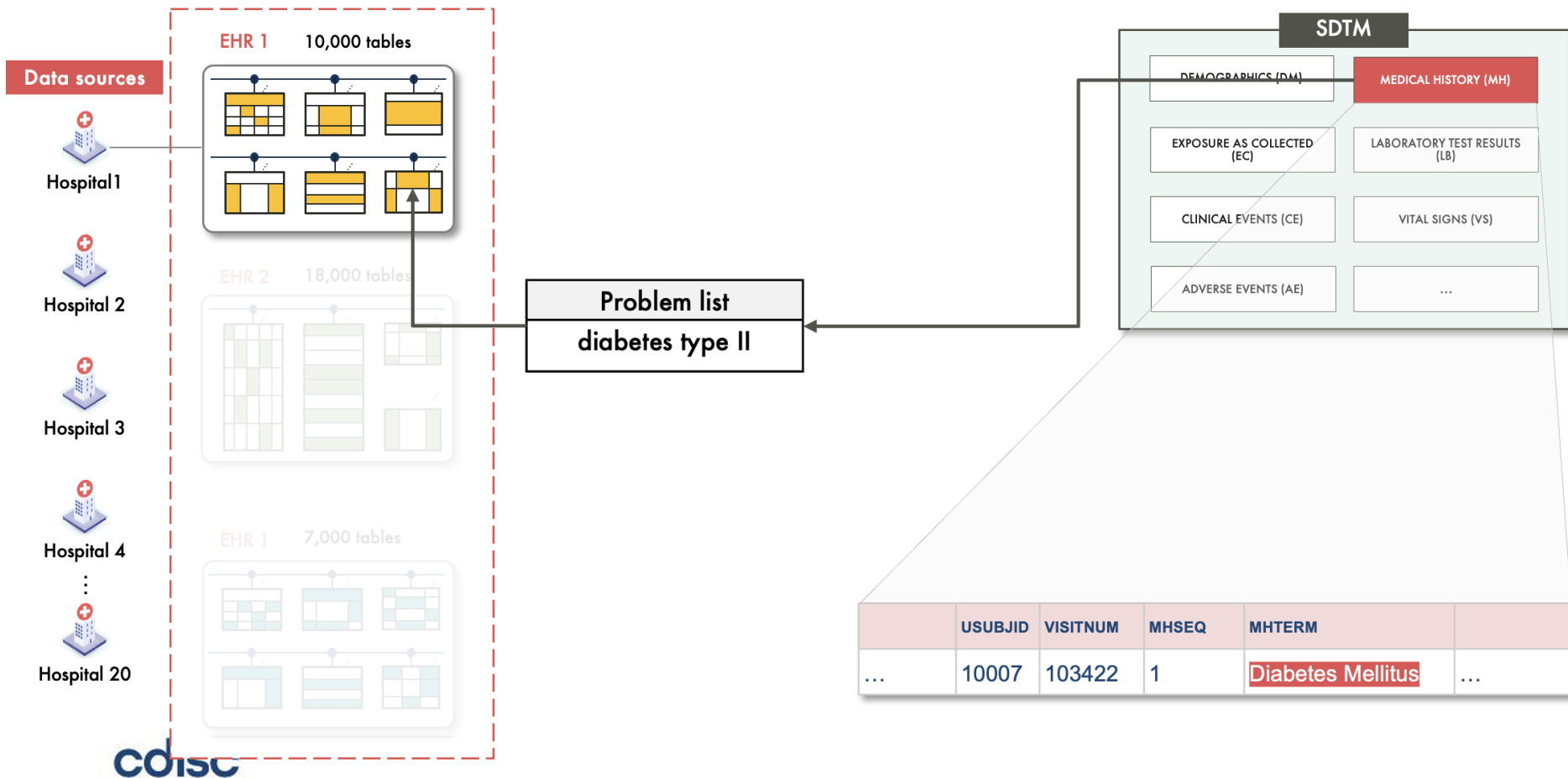
Medical history (MH)

STUDYID	DOMAIN	USUBJID	VISITNUM	MHSEQ	MHTERM	MHCAT	MHDTC	MHSTDTC
EX-030	MH	10005	115407	8	Hypothyroidism	PAST MEDICAL HISTORY	2024-10-07	2024-10-07
EX-030	MH	10006	101357	1	Chronic Heart Failure	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	2	Hypothyroidism	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	3	History Of Myocardial Infarction	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	4	Peripheral Arterial Disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	5	Dementia	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	6	Hyperlipidemia	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	7	Chronic kidney disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	8	Atrial Fibrillation	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	9	Coronary Artery Disease	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	10	Hypertension	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	107449	11	Aortic Plaque	PAST MEDICAL HISTORY	2023-07-07	2023-07-07
EX-030	MH	10006	102846	12	Chronic Obstructive Pulmonary Disease	PAST MEDICAL HISTORY	2023-07-07	2023-07-07
EX-030	MH	10007	103422	1	Diabetes Mellitus	PAST MEDICAL HISTORY	2021-03-05	2021-03-05
EX-030	MH	10007	104819	2	Hyperlipidemia	PAST MEDICAL HISTORY	2021-08-12	2021-08-12
EX-030	MH	10007	104819	3	Chronic kidney disease	PAST MEDICAL HISTORY	2023-02-28	2023-02-28
EX-030	MH	10007	104819	4	Atrial Fibrillation	PAST MEDICAL HISTORY	2023-02-28	2023-02-28

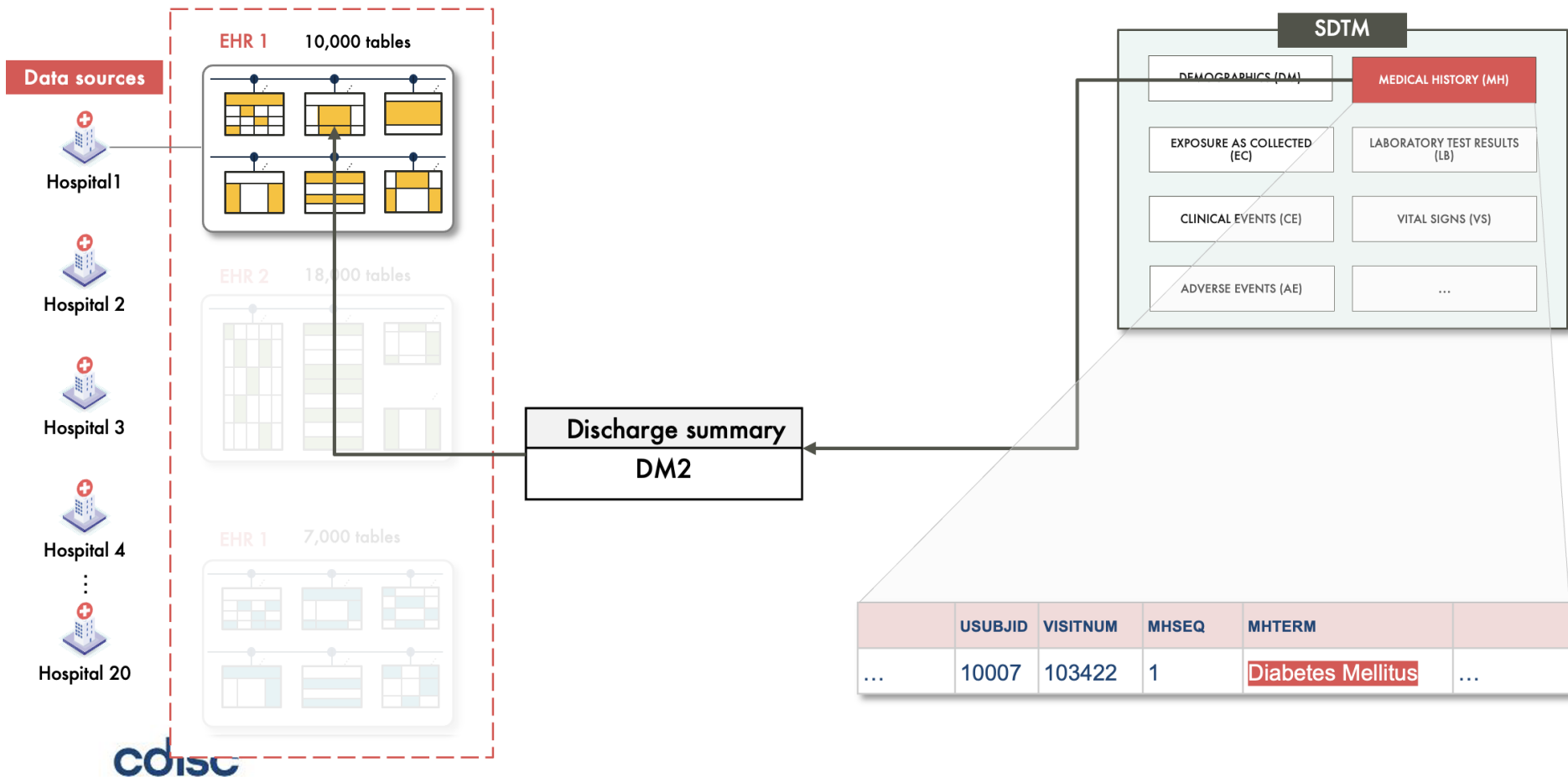
RWD: Straight to SDTM project



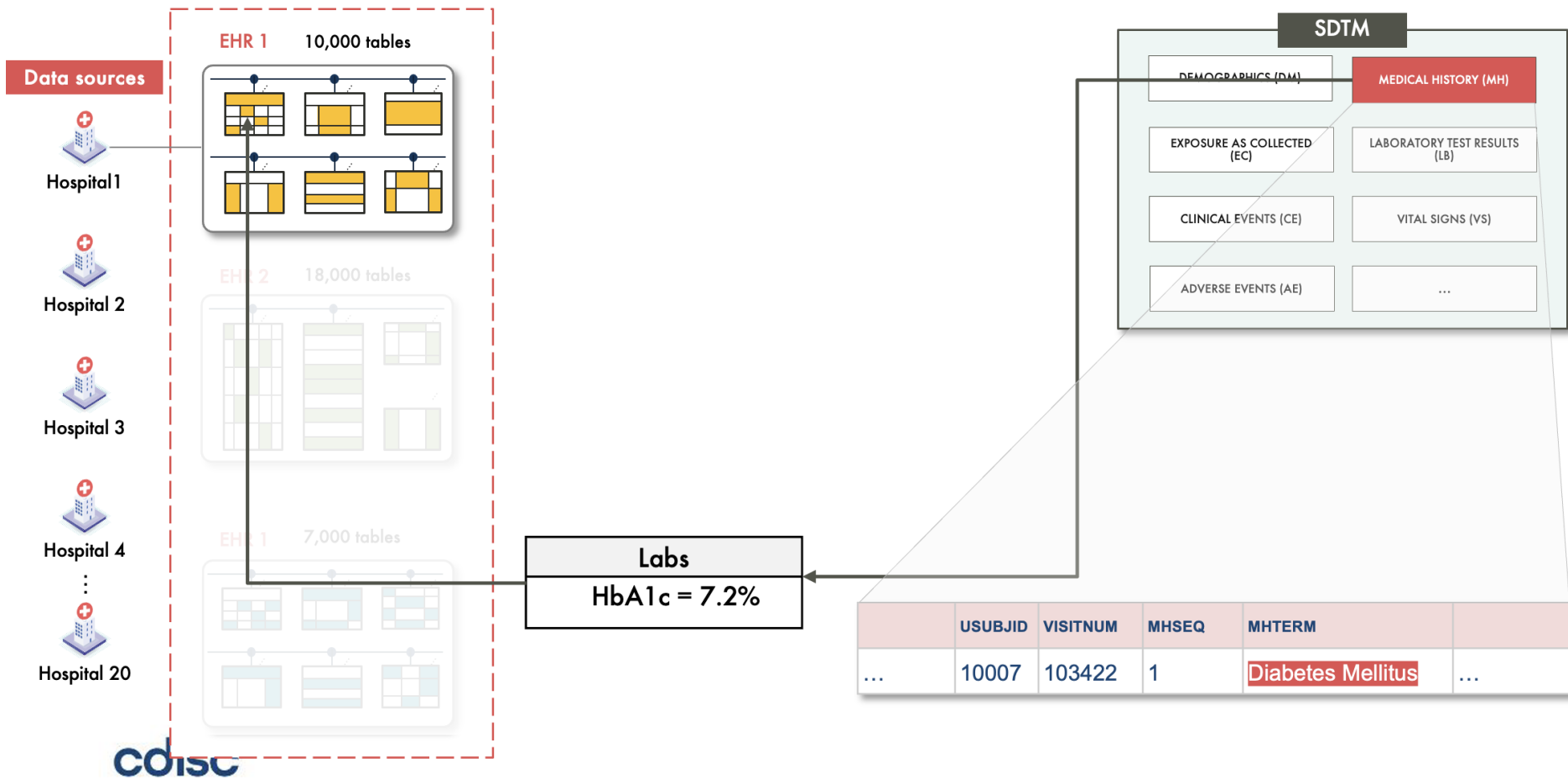
RWD: Straight to SDTM project



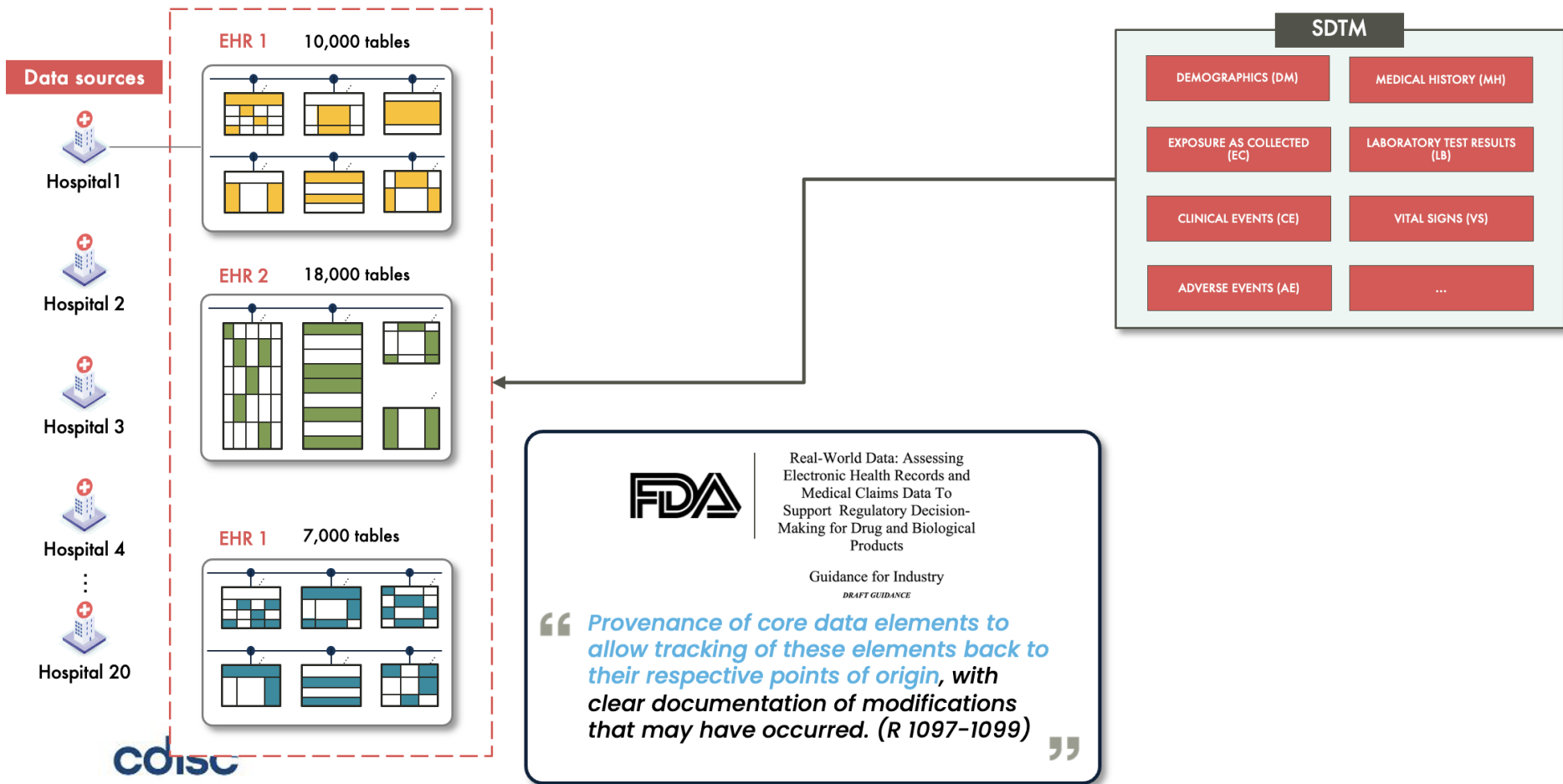
RWD: Straight to SDTM project



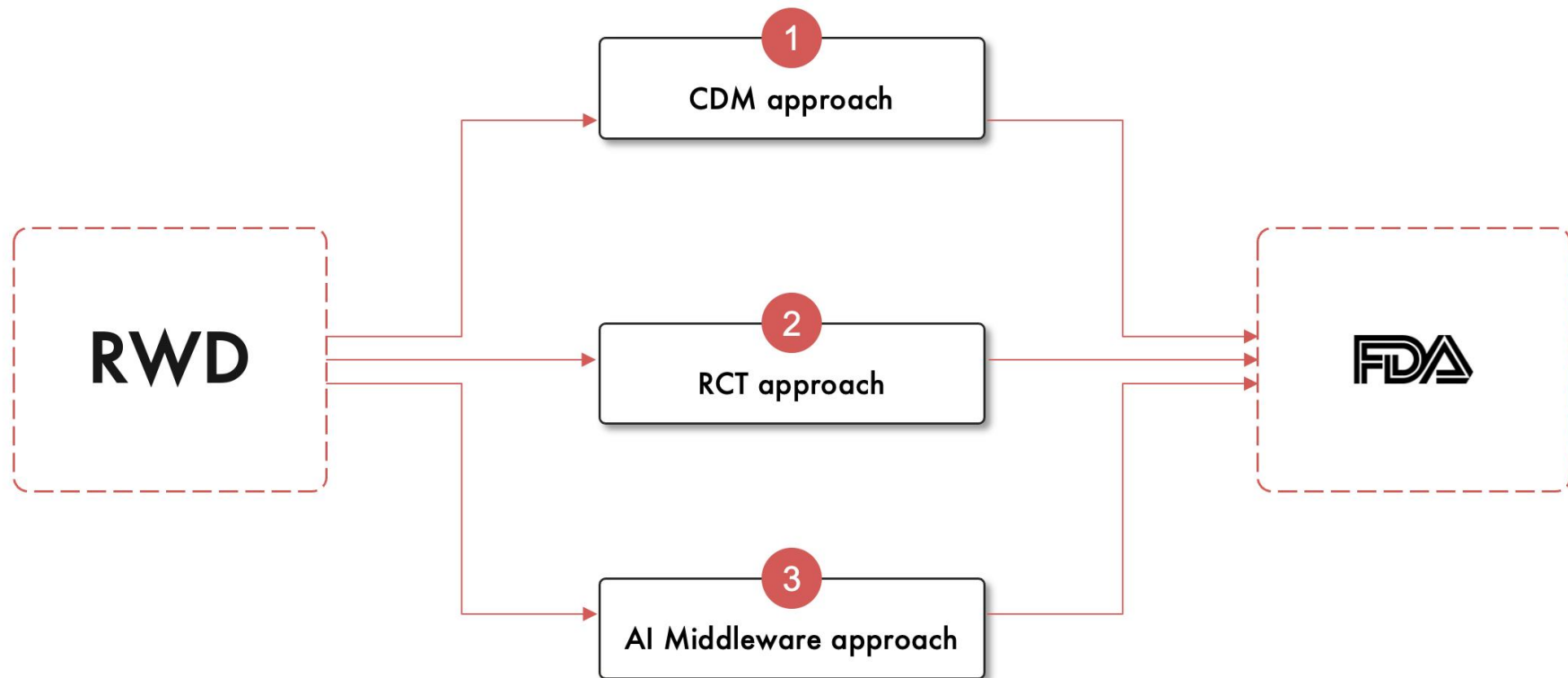
RWD: Straight to SDTM project

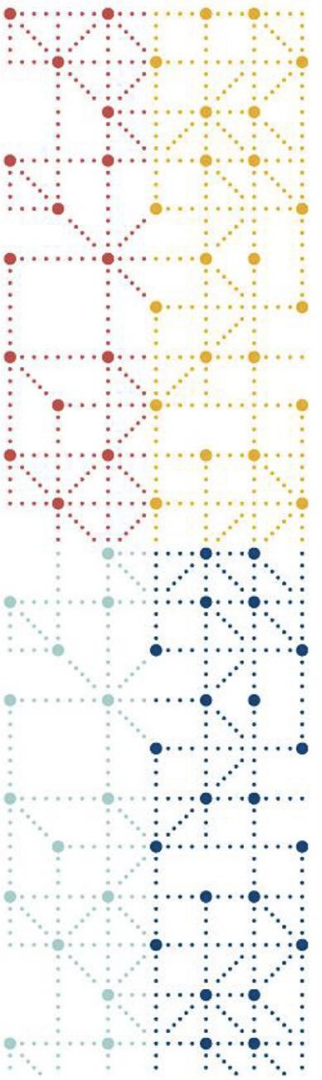


RWD: Straight to SDTM project



Takeaways





Thank You!