cdisc

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.*

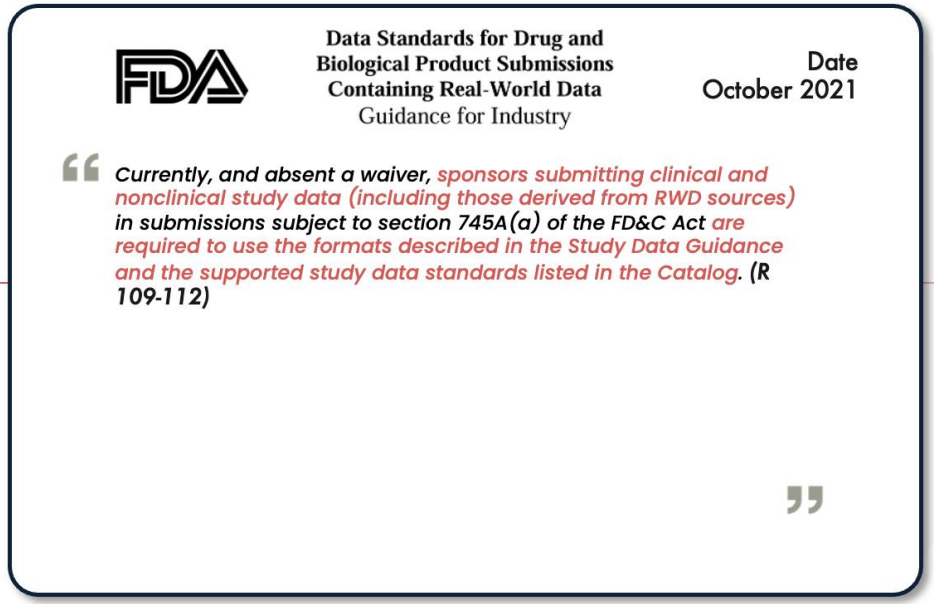
- *Mayur Saxena is an employee of Droice Labs, a company that processes and prepares RWD for clinical and research applications.*

Requirements for regulatory RWD

cdisc

Requirements for regulatory RWD

RWD



FDA

Requirements for regulatory RWD

RWD

(EHR, Claims, etc..)

FDA

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)**

FDA

cdisc

Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: **10 sites**

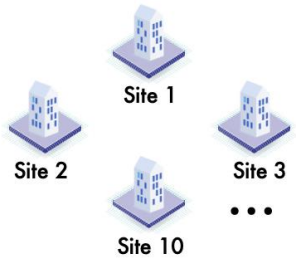
PATIENTS: **1,000**

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

DATA SOURCES: EHR - **20 hospitals**

PATIENTS: **50,000**

Intervention arm data sources



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)

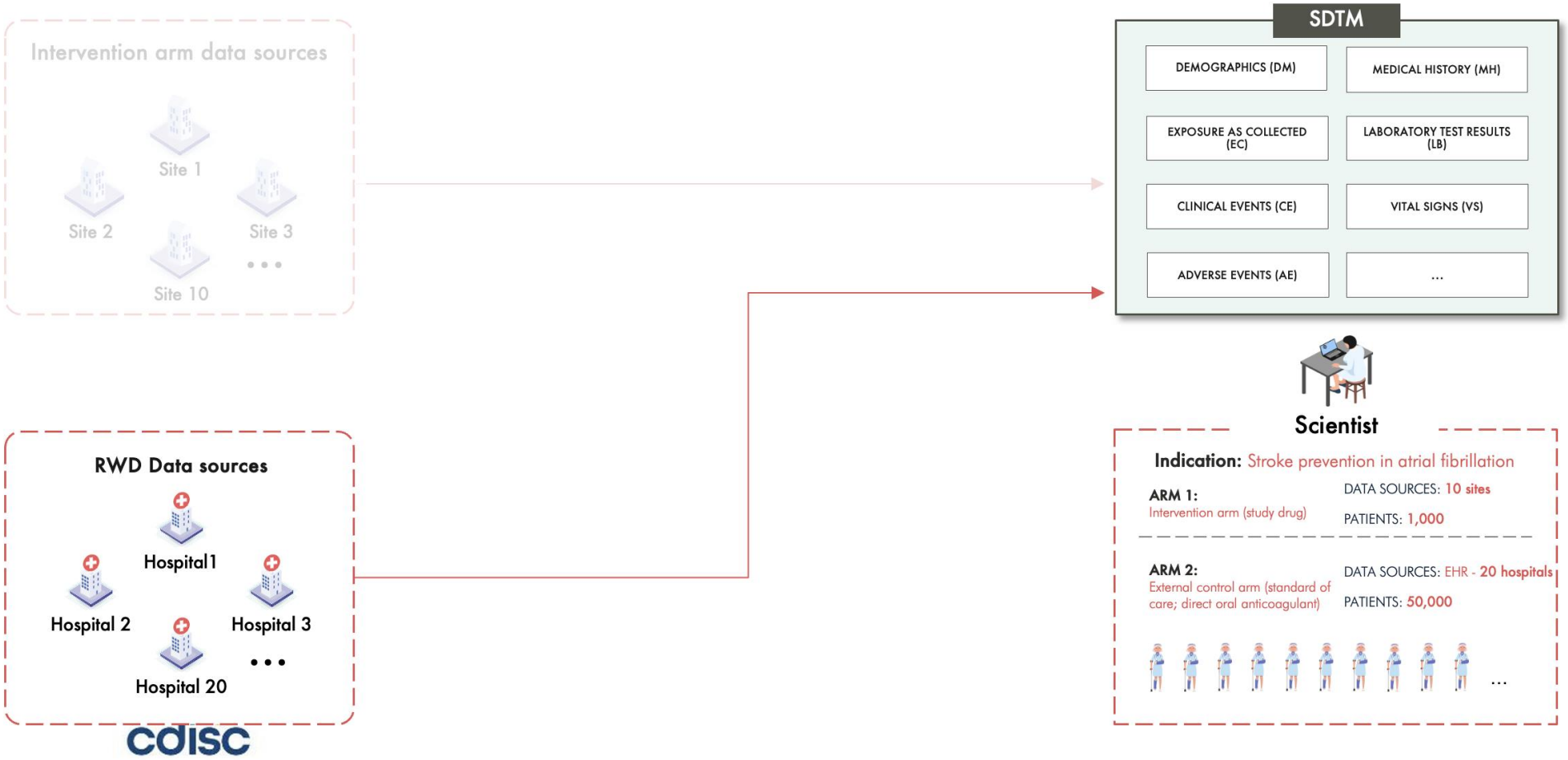
DATA SOURCES: **10 sites**

PATIENTS: **1,000**

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

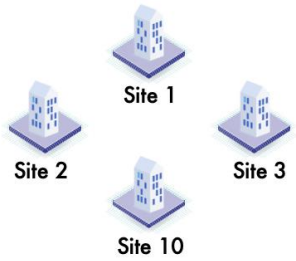
DATA SOURCES: EHR - **20 hospitals**

PATIENTS: **50,000**


Intervention arm data sources
Site 1
Site 2
Site 3
Site 10
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
...
RWD Data sources
Hospital1
Hospital 2
Hospital 3
Hospital 20
...
cdisc

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**

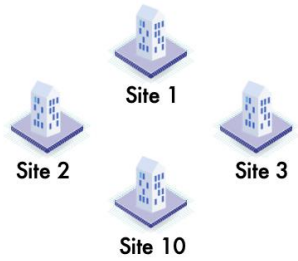
ARM 2:
External control arm (standard of care; direct oral anticoagulant)

DATA SOURCES: EHR - **20 hospitals**

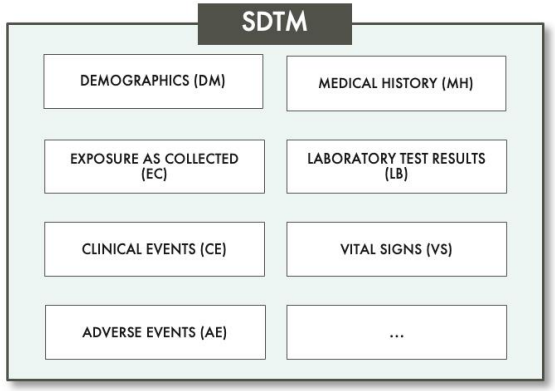
PATIENTS: **50,000**

RCT Data Journey

Intervention arm data sources



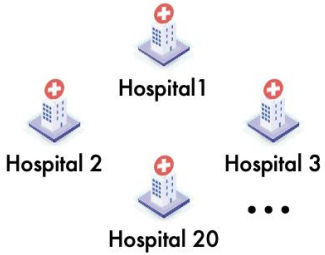
CRF
(From 10 sites)



Scientist

RWD Data Journey

RWD Data sources



?

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**

CDISC

RWD to SDTM: A deeper look

Data sources

Hospital1

Hospital 2

Hospital 3

Hospital 4

⋮

Hospital 20

Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: **10 sites**

PATIENTS: **1,000**

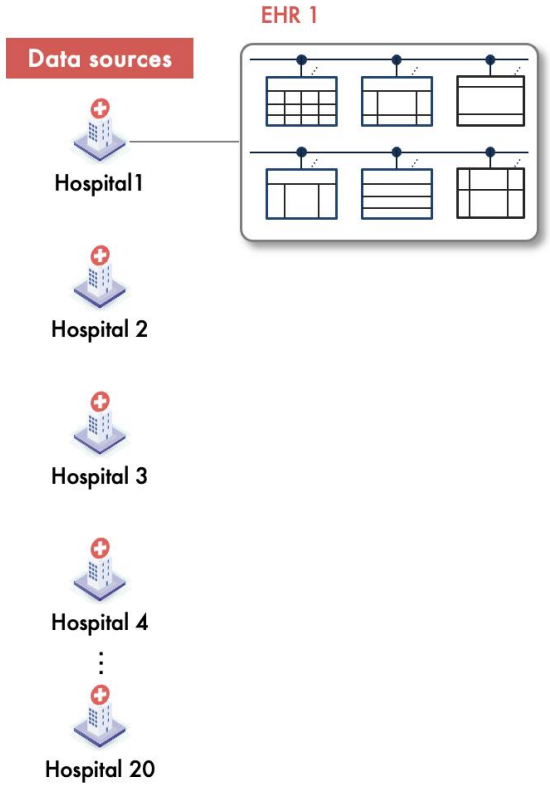
ARM 2:
External control arm (standard of care; direct oral anticoagulant)

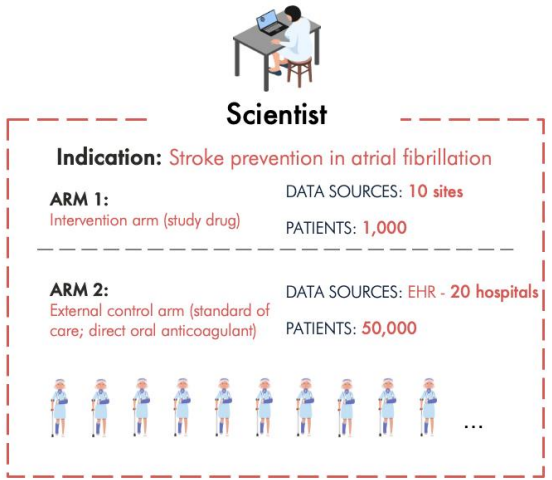
DATA SOURCES: EHR - **20 hospitals**

PATIENTS: **50,000**

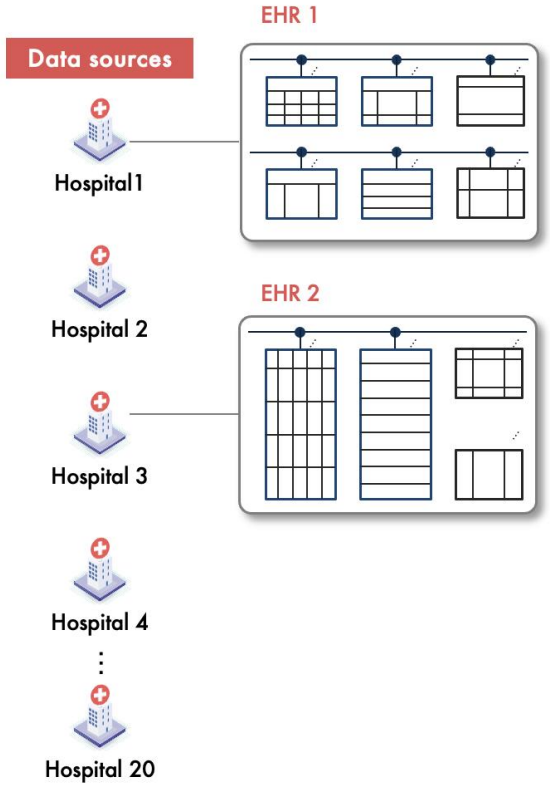
...

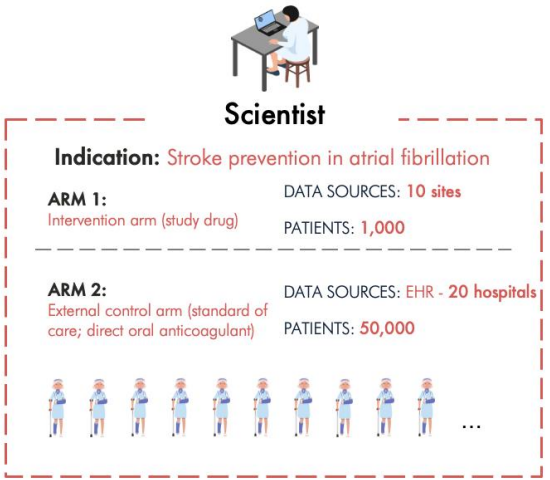
RWD to SDTM: A deeper look



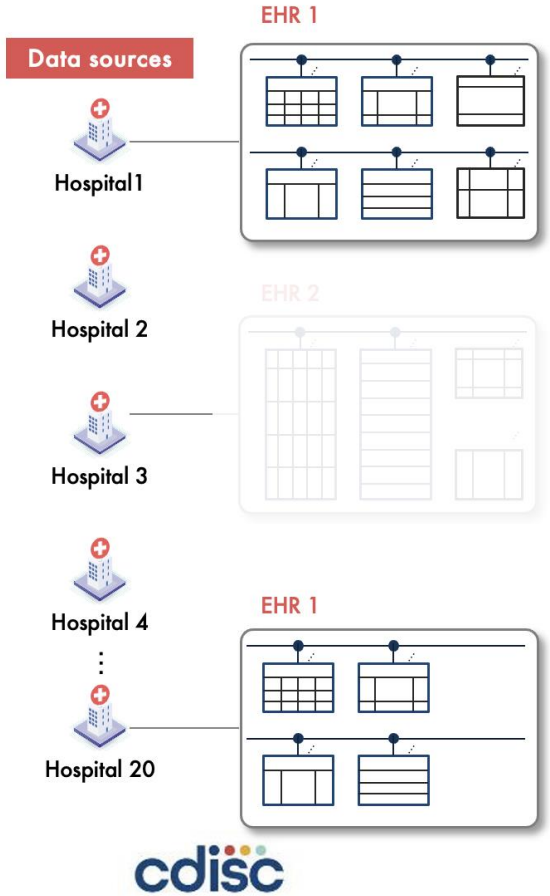


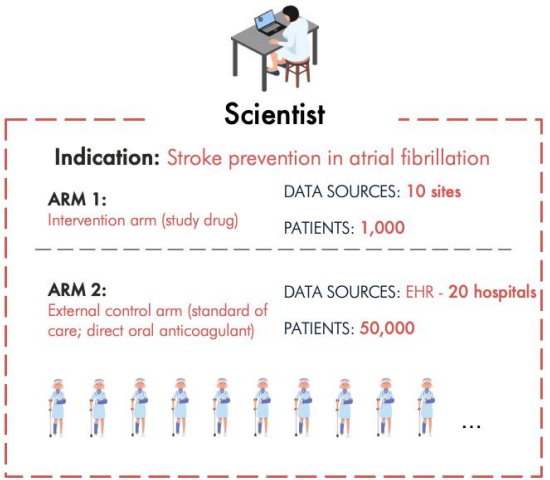
RWD to SDTM: A deeper look



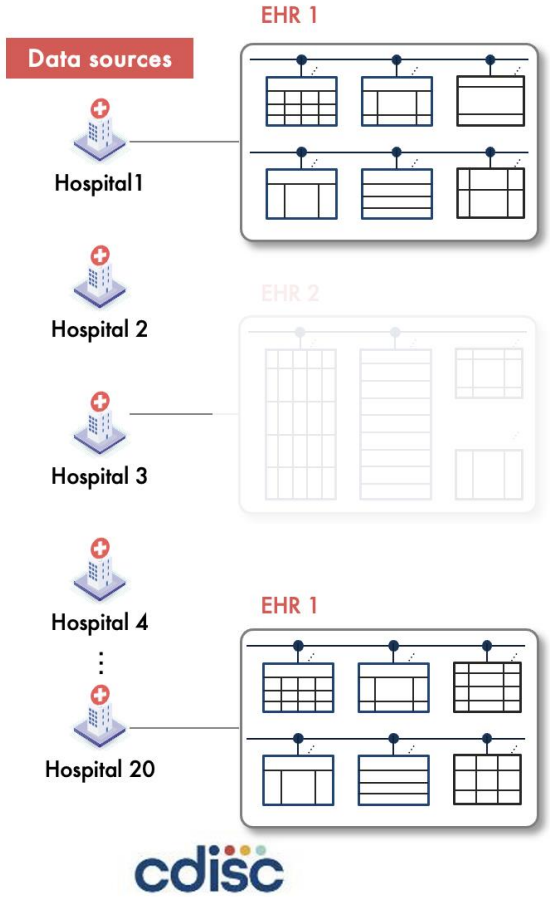


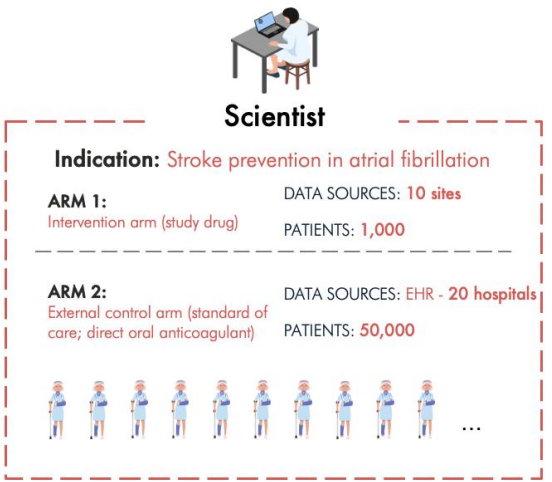
RWD to SDTM: A deeper look



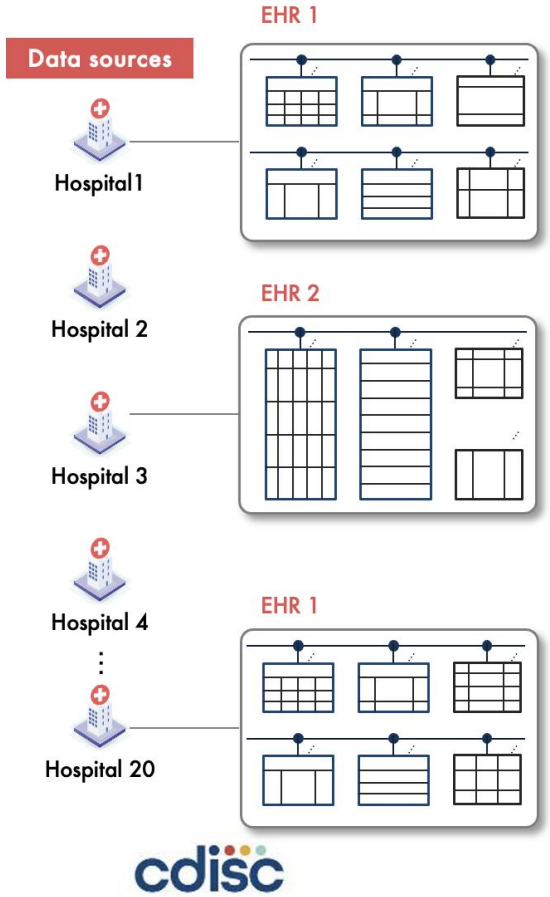


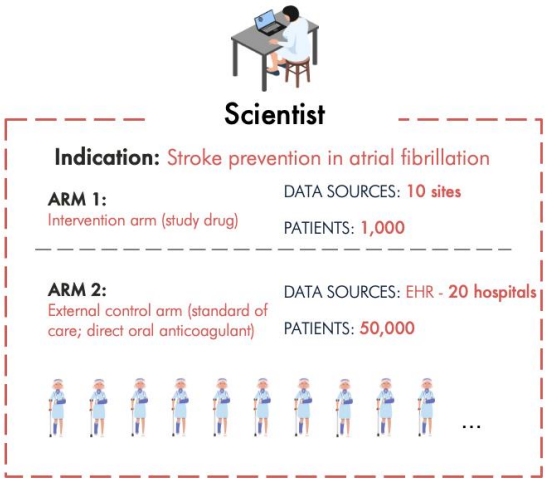
RWD to SDTM: A deeper look



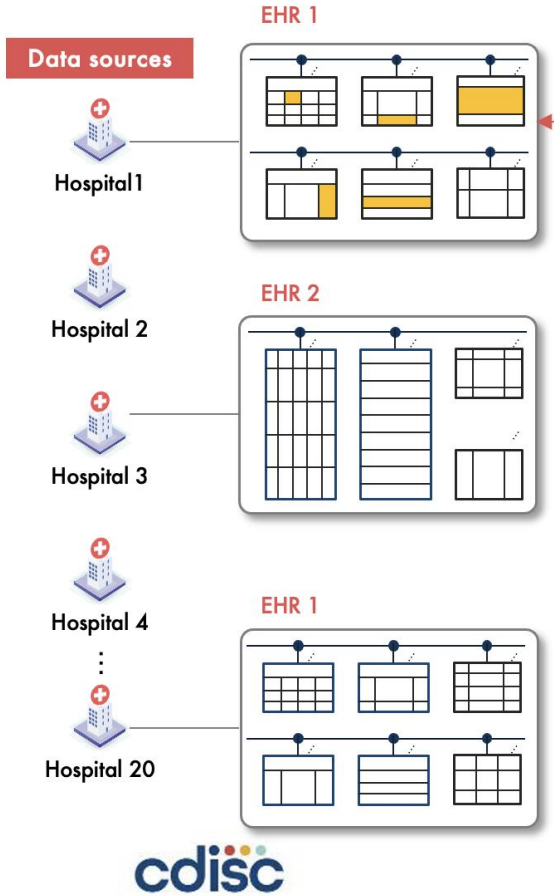


RWD to SDTM: A deeper look



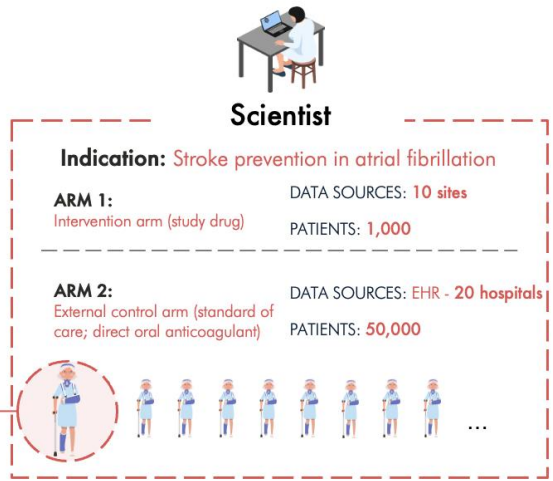


RWD to SDTM: A deeper look

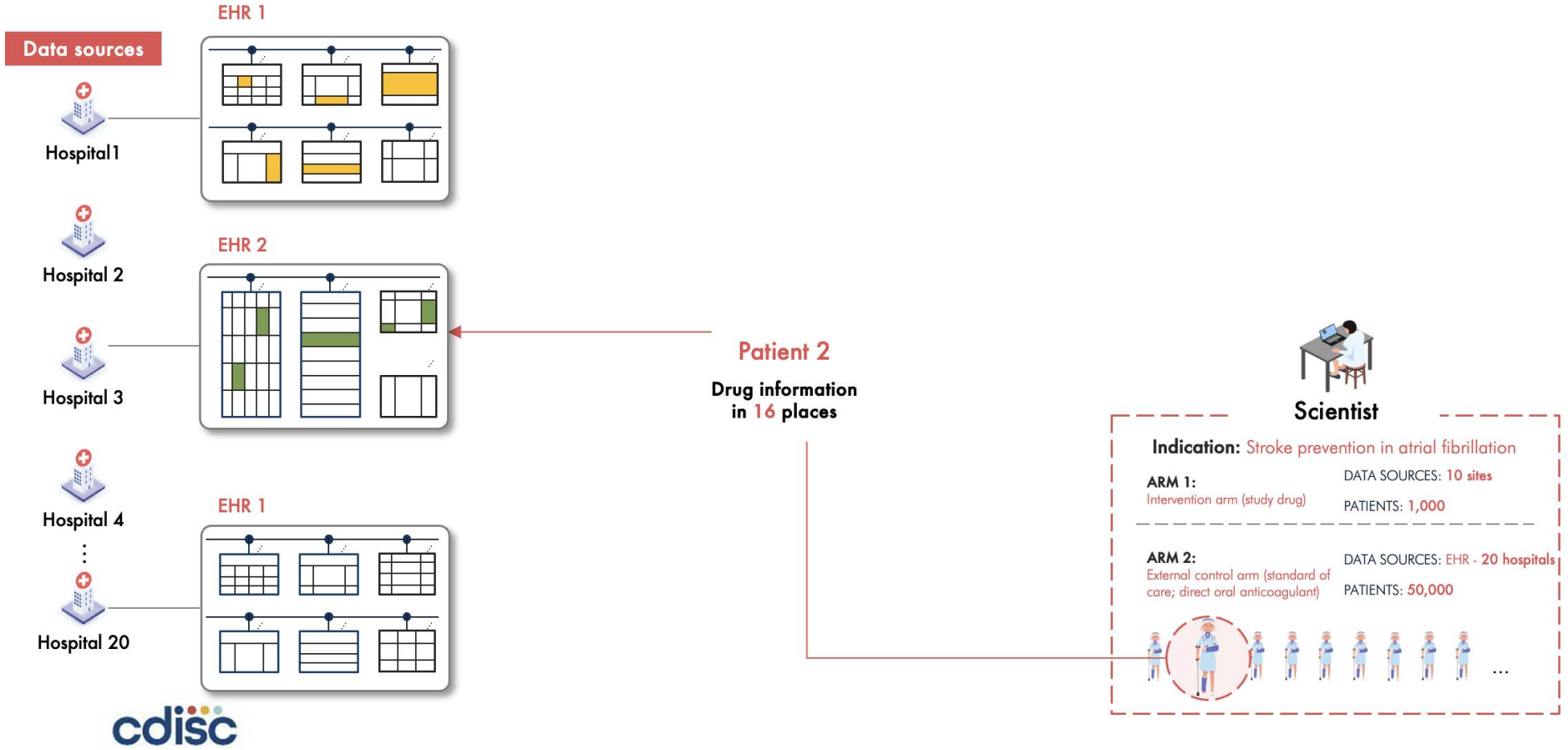


Patient 1

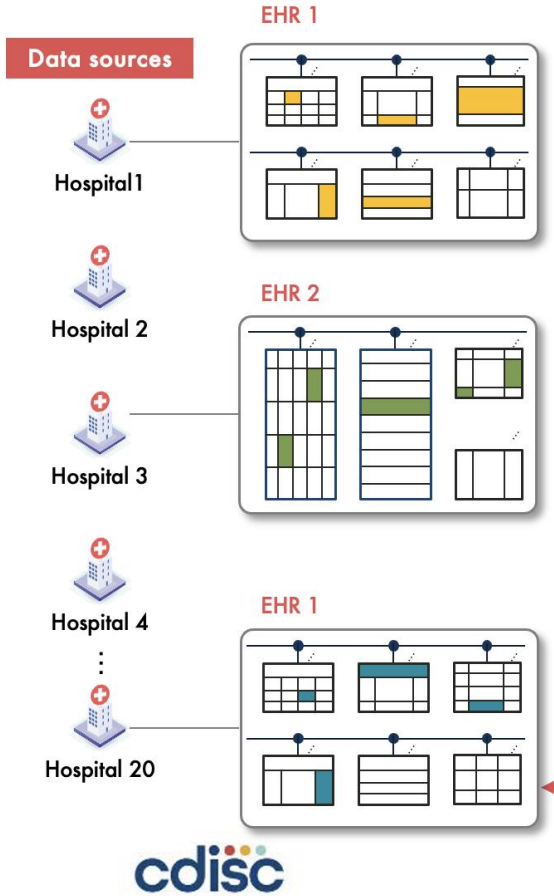
Drug information in 12 places



RWD to SDTM: A deeper look

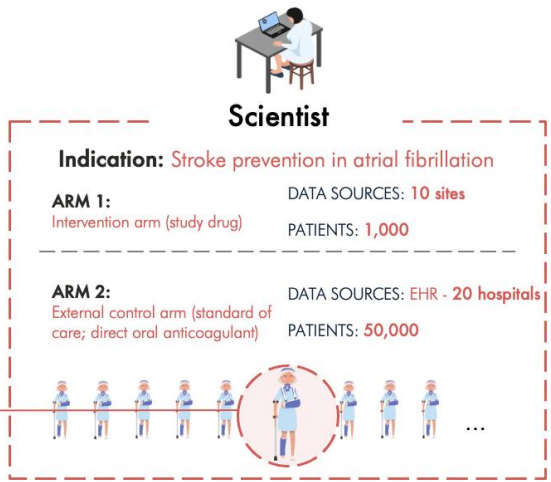


RWD to SDTM: A deeper look

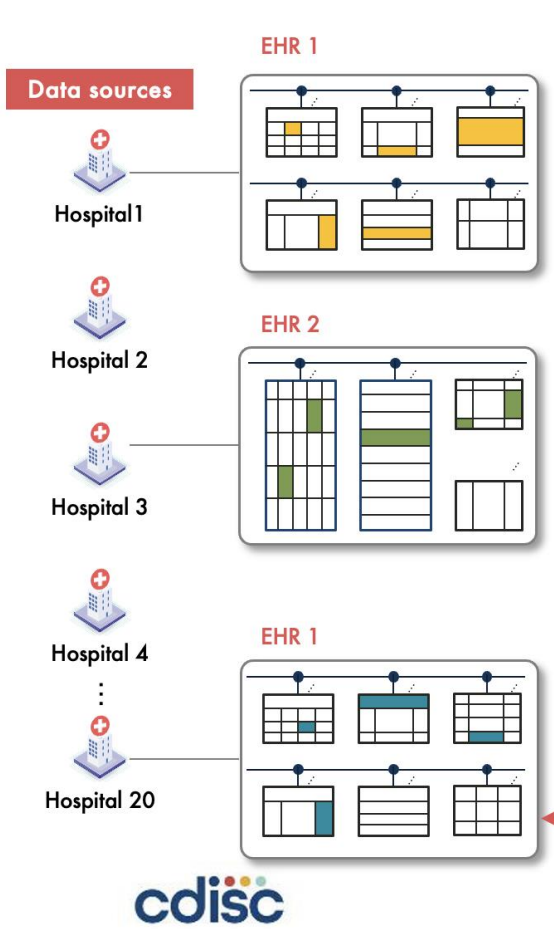


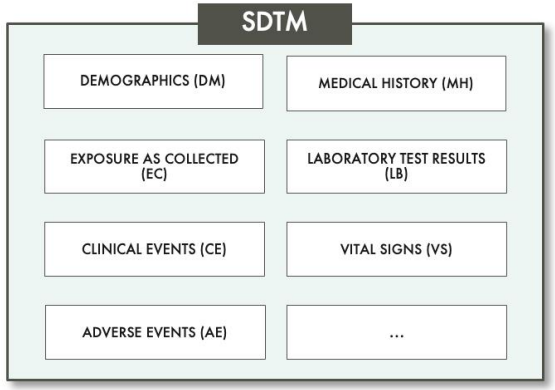
Patient 436

Drug information in 5 places



RWD to SDTM: A deeper look





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

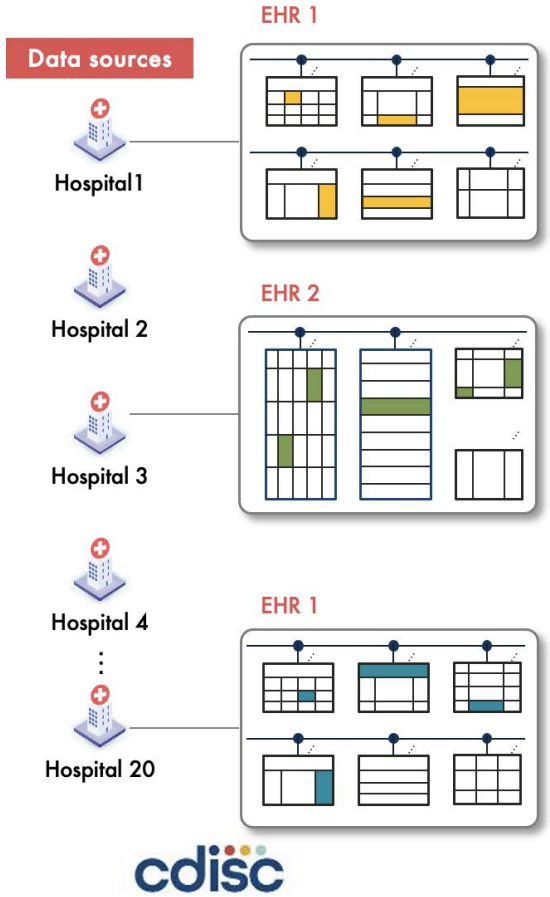
PATIENTS: 50,000

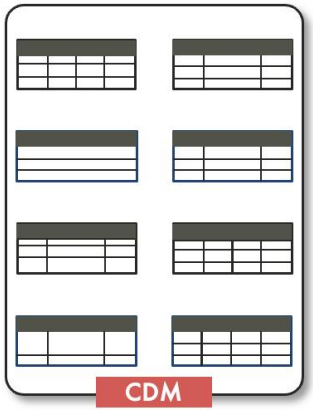
Patient 436

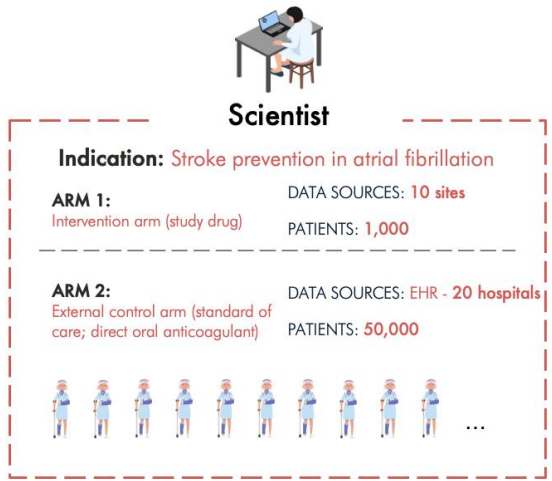
Drug information in 5 places

cdisc

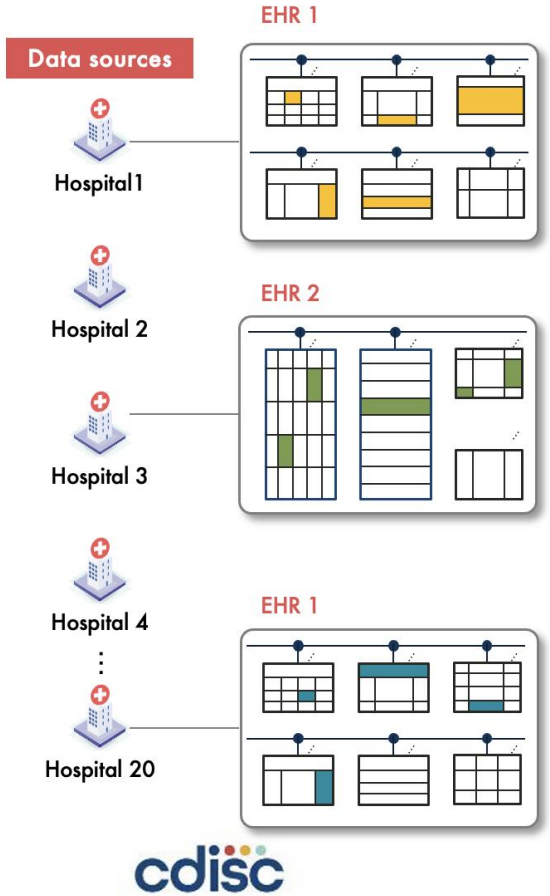
Common data models (CDM)

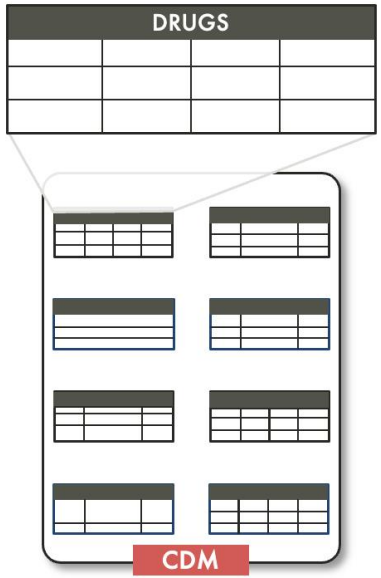






Common data models (CDM)





Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: 10 sites

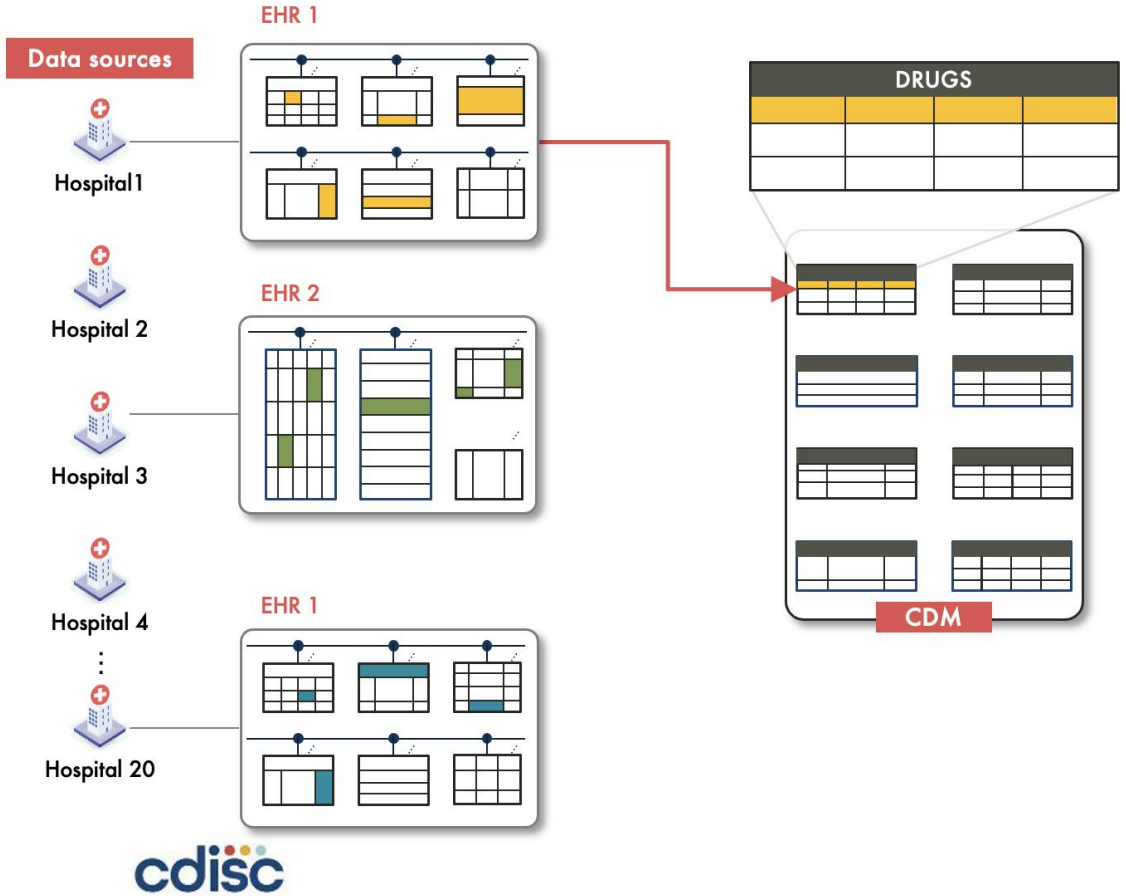
PATIENTS: 1,000

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

DATA SOURCES: EHR - 20 hospitals

PATIENTS: 50,000

Common data models (CDM)



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: 10 sites

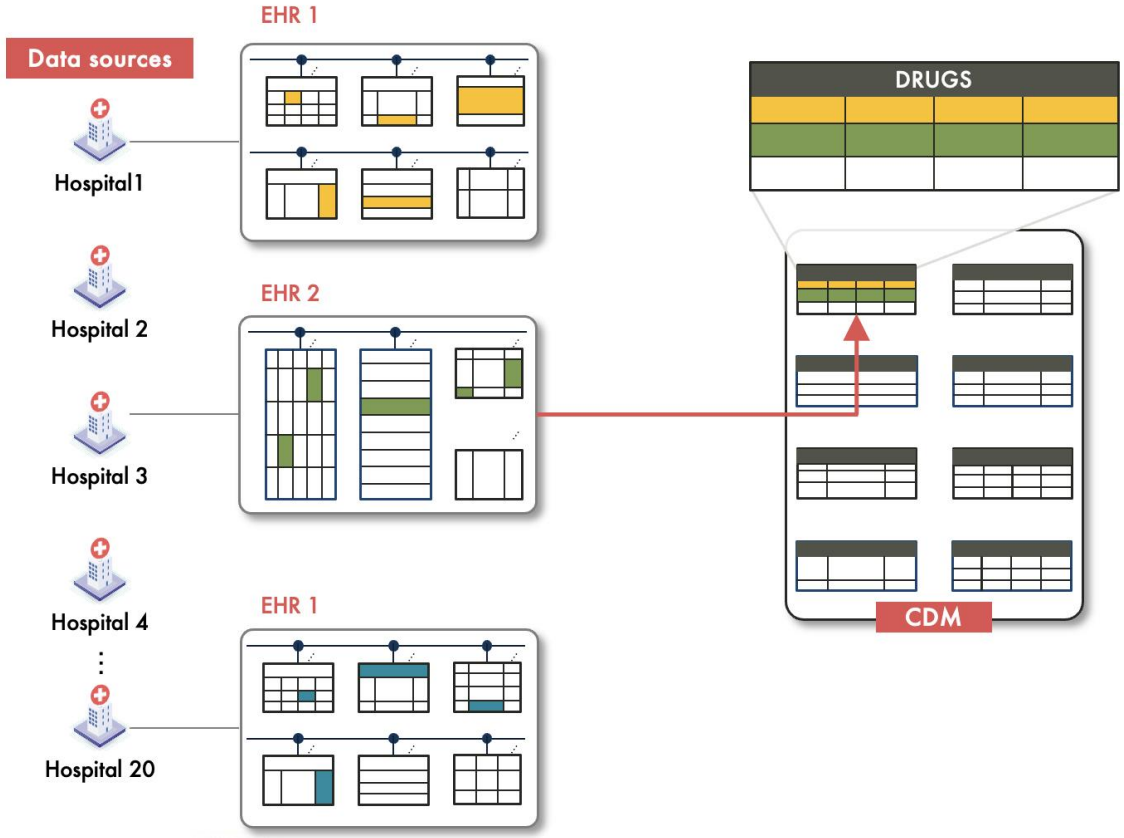
PATIENTS: 1,000

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

DATA SOURCES: EHR - 20 hospitals

PATIENTS: 50,000

Common data models (CDM)



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: 10 sites

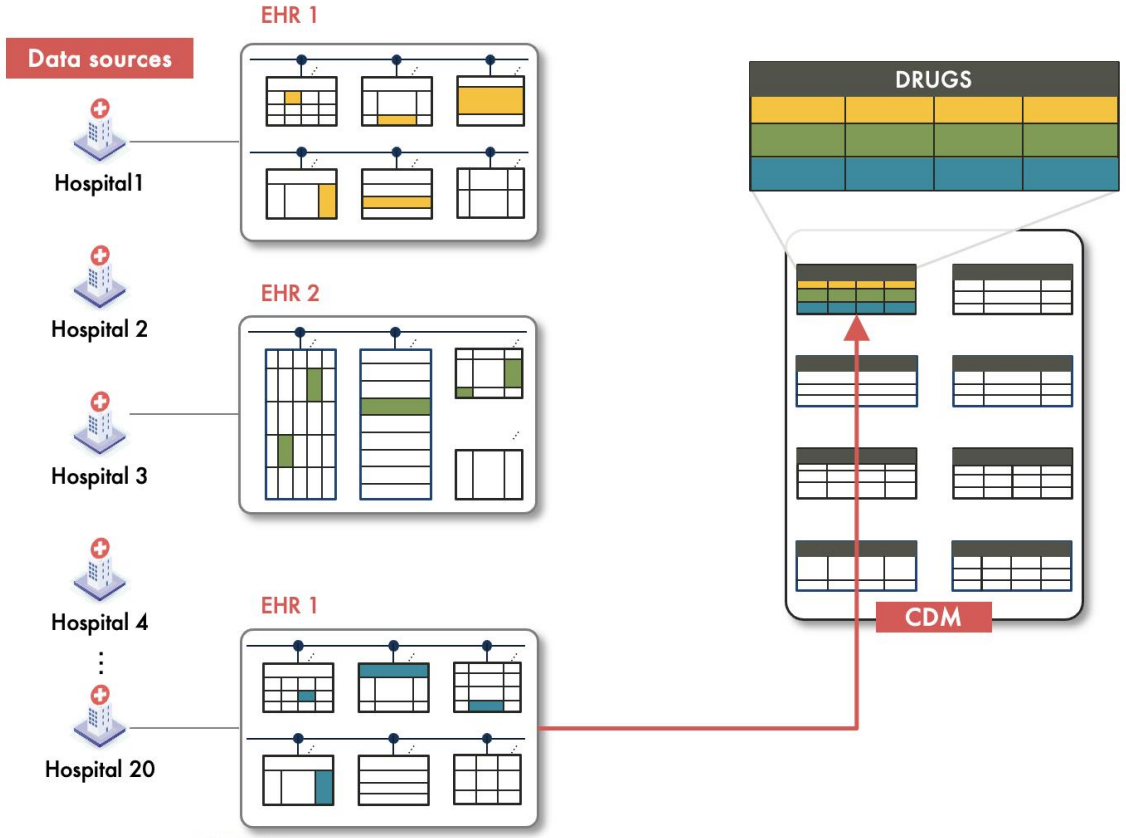
PATIENTS: 1,000

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

DATA SOURCES: EHR - 20 hospitals

PATIENTS: 50,000

Common data models (CDM)



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: 10 sites

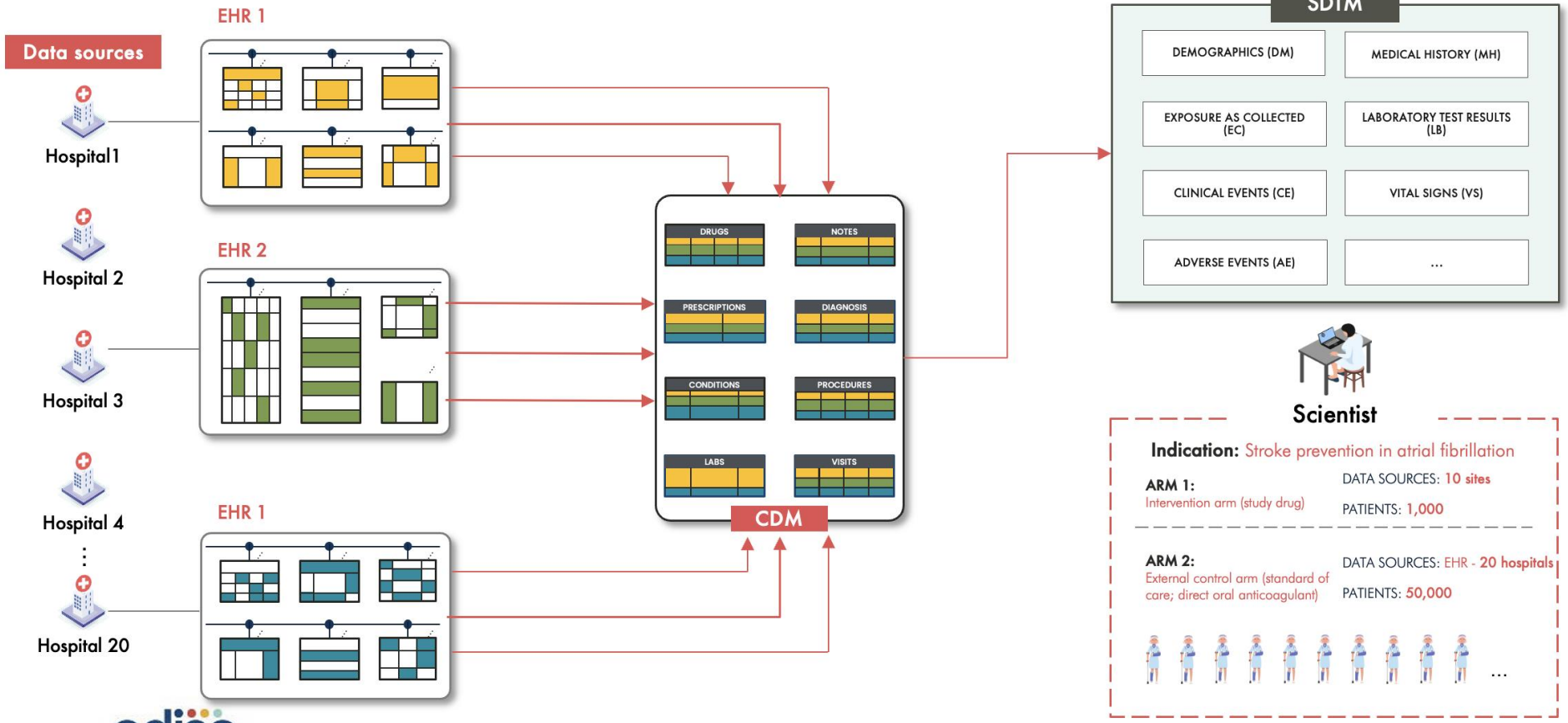
PATIENTS: 1,000

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

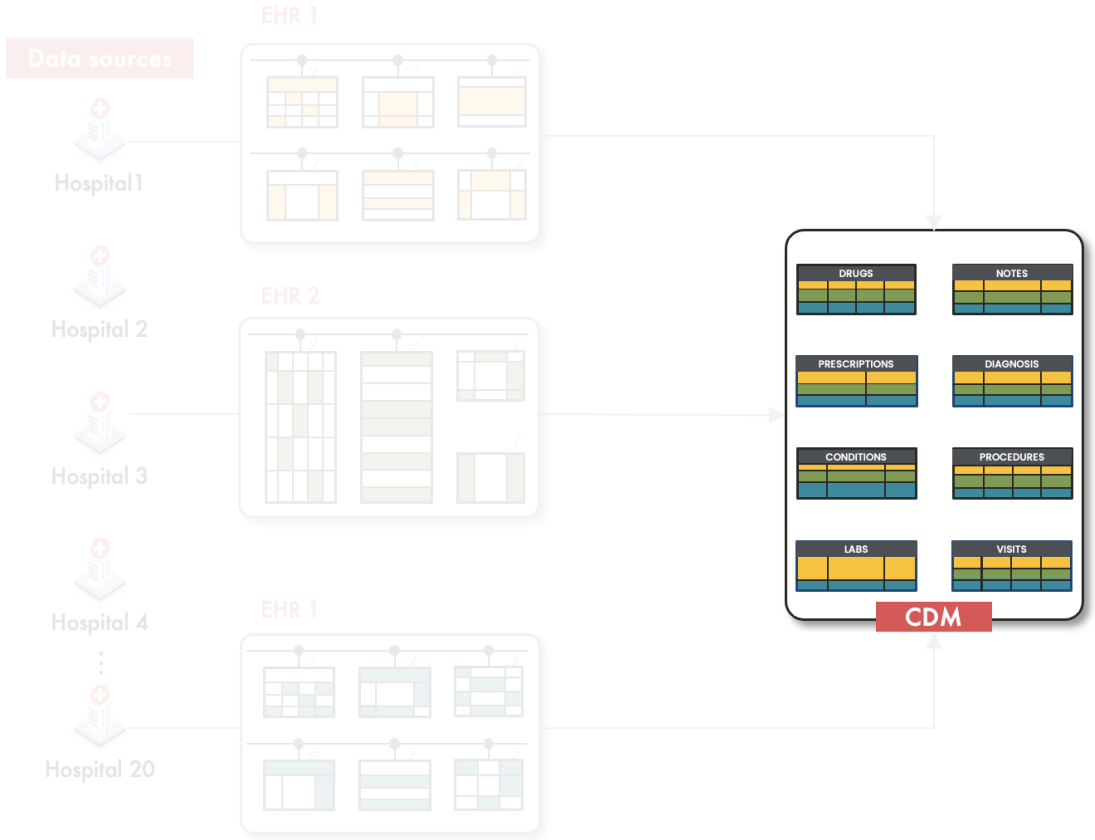
DATA SOURCES: EHR - 20 hospitals

PATIENTS: 50,000

Common data models (CDM)



Common data models (CDM)



FDA

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

> *...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:
Intervention arm (study drug)

DATA SOURCES: **10 sites**

PATIENTS: **1,000**

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

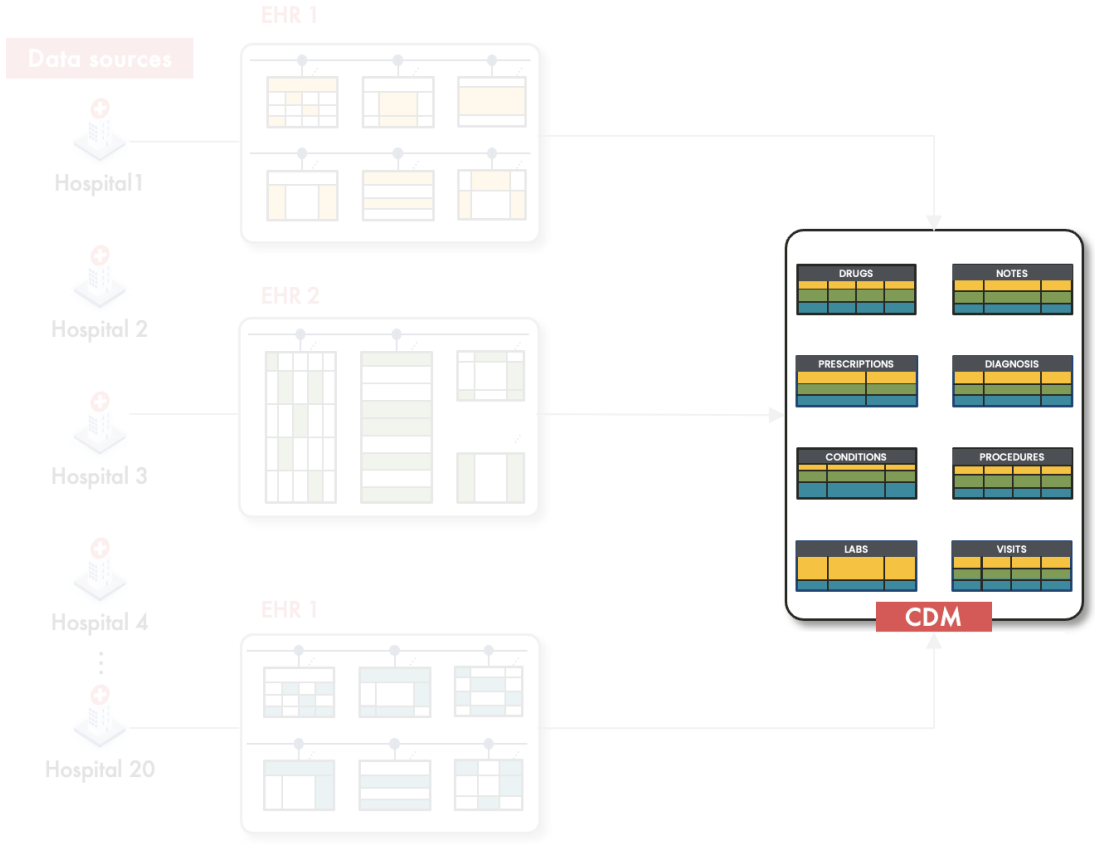
DATA SOURCES: EHR - **20 hospitals**

PATIENTS: **50,000**

...

cdisc

Common data models (CDM)



FDA

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

> *...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

...

cdisc

Common data models (CDM)

FDA

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

> *...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

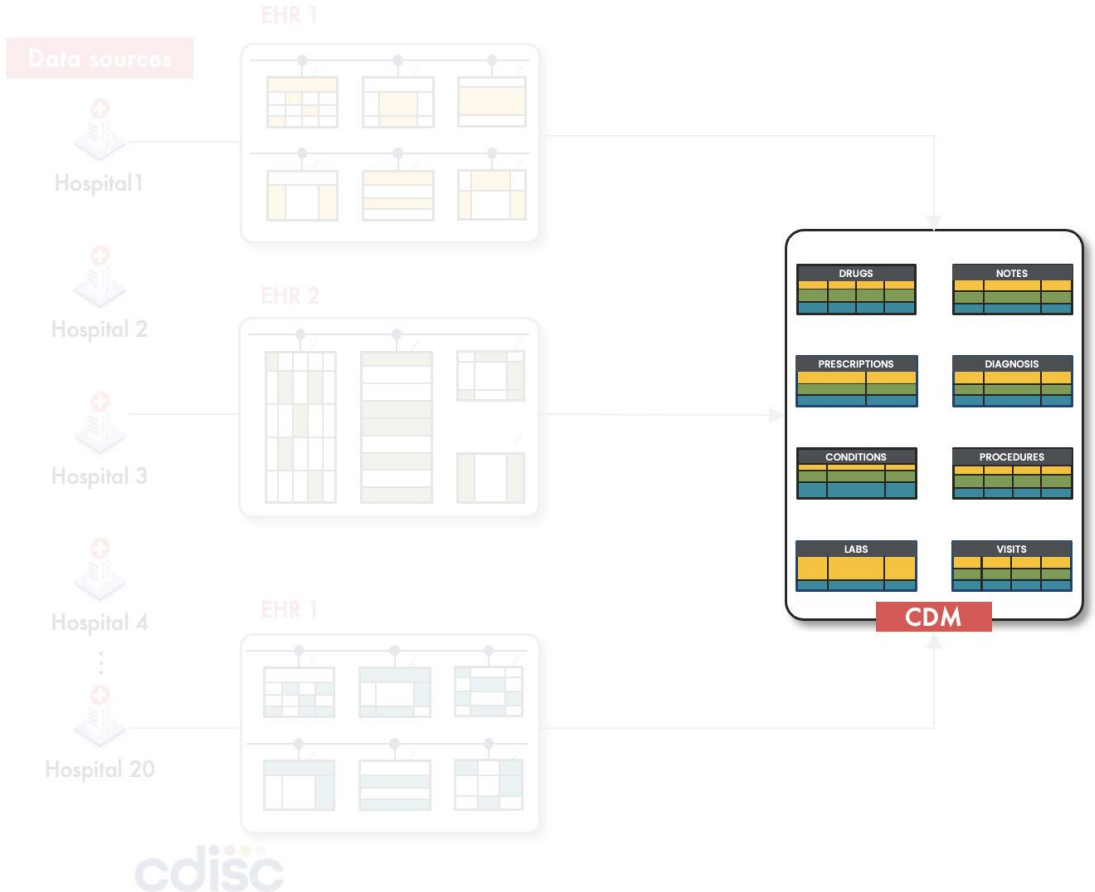
ARM 2:
External control arm (standard of care; direct oral anticoagulant)

DATA SOURCES: EHR - 20 hospitals

PATIENTS: 50,000

...

Data loss in common data models



Scientist

Indication: Stroke prevention in atrial fibrillation

ARM 1:
Intervention arm (study drug)

DATA SOURCES: **10 sites**

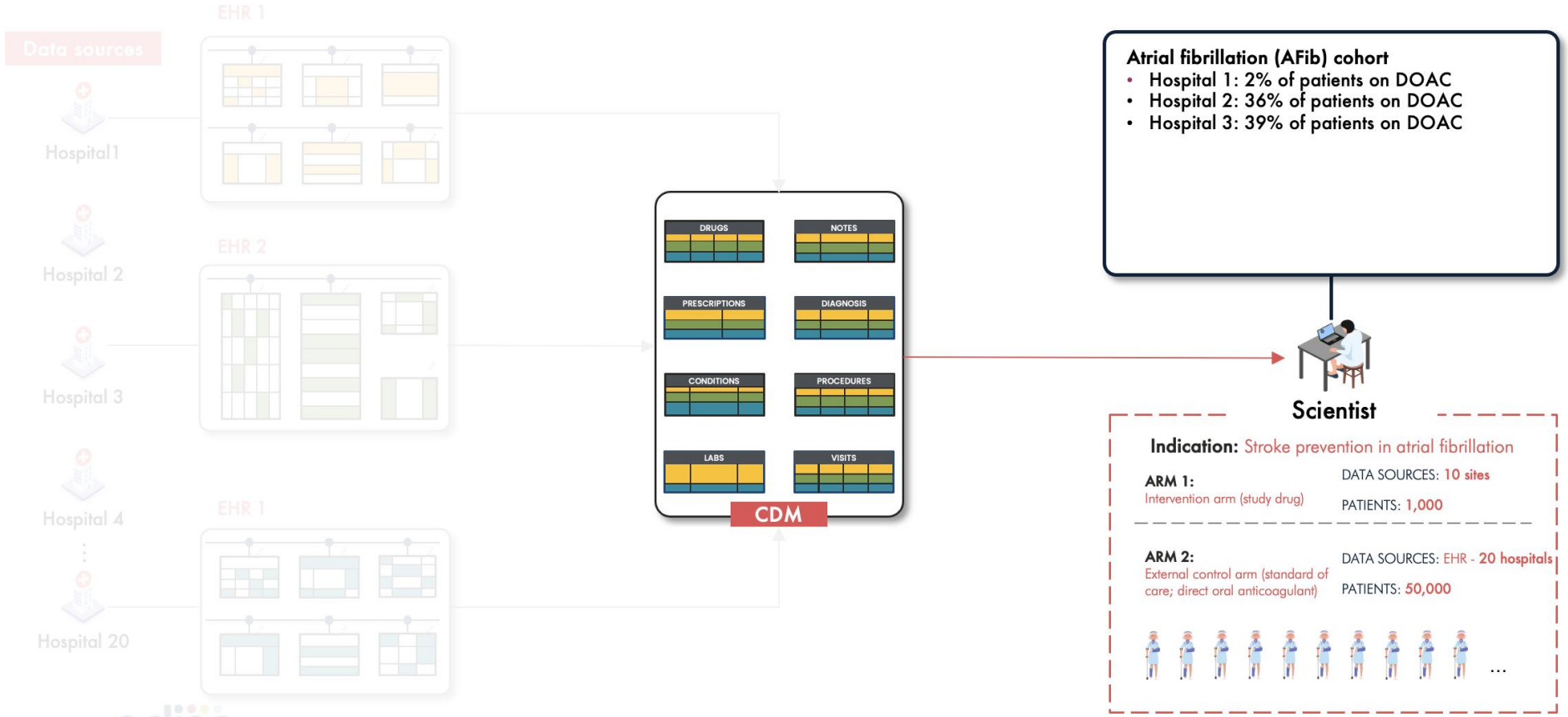
PATIENTS: **1,000**

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

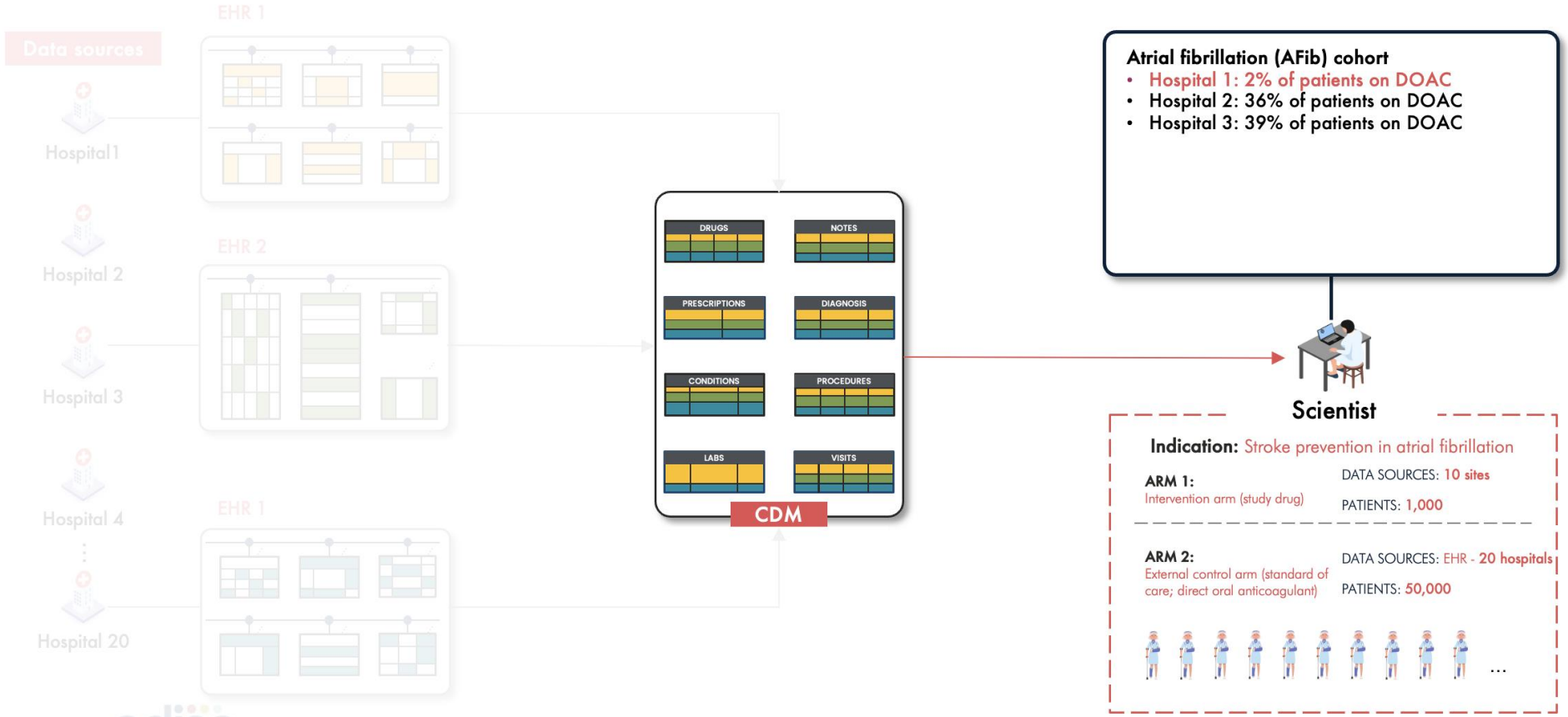
DATA SOURCES: EHR - **20 hospitals**

PATIENTS: **50,000**

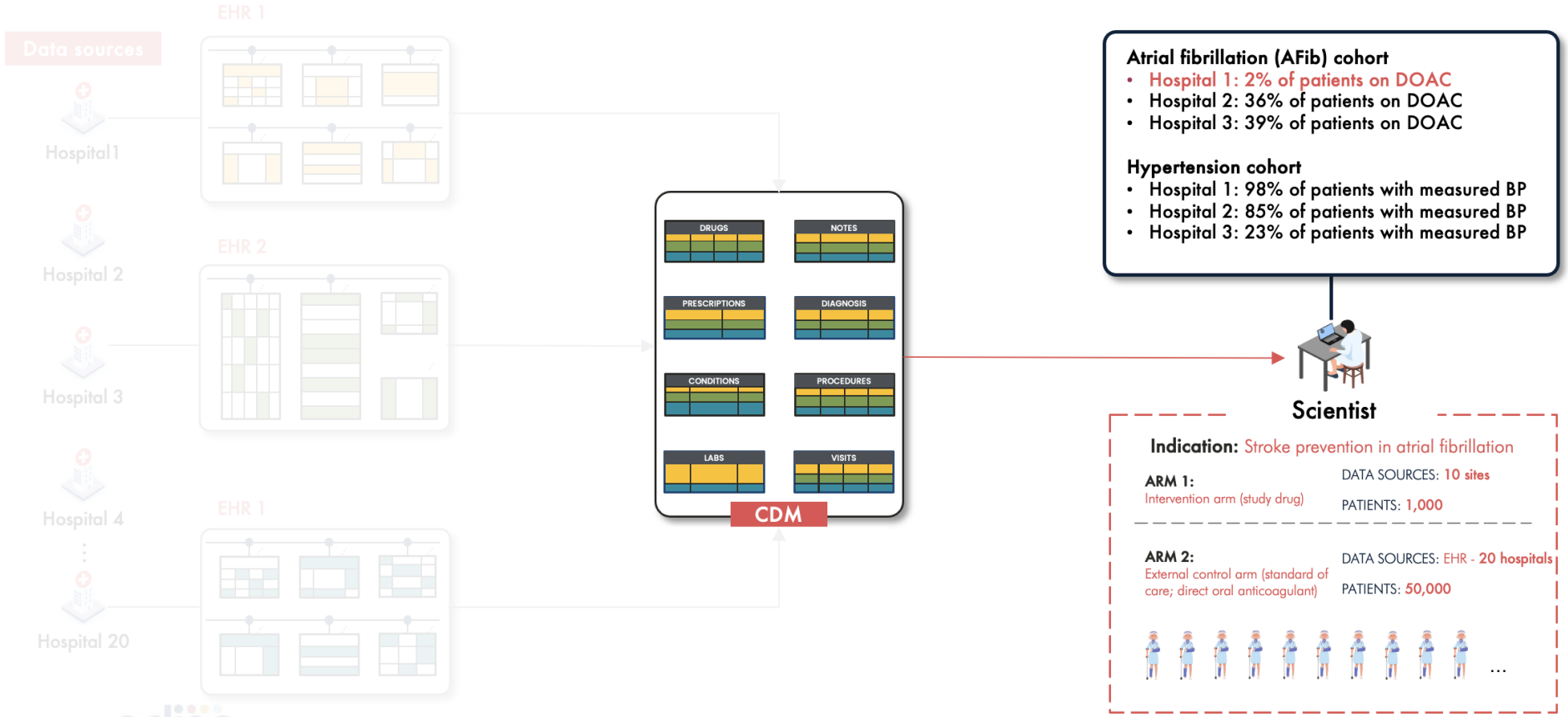
Data loss in common data models



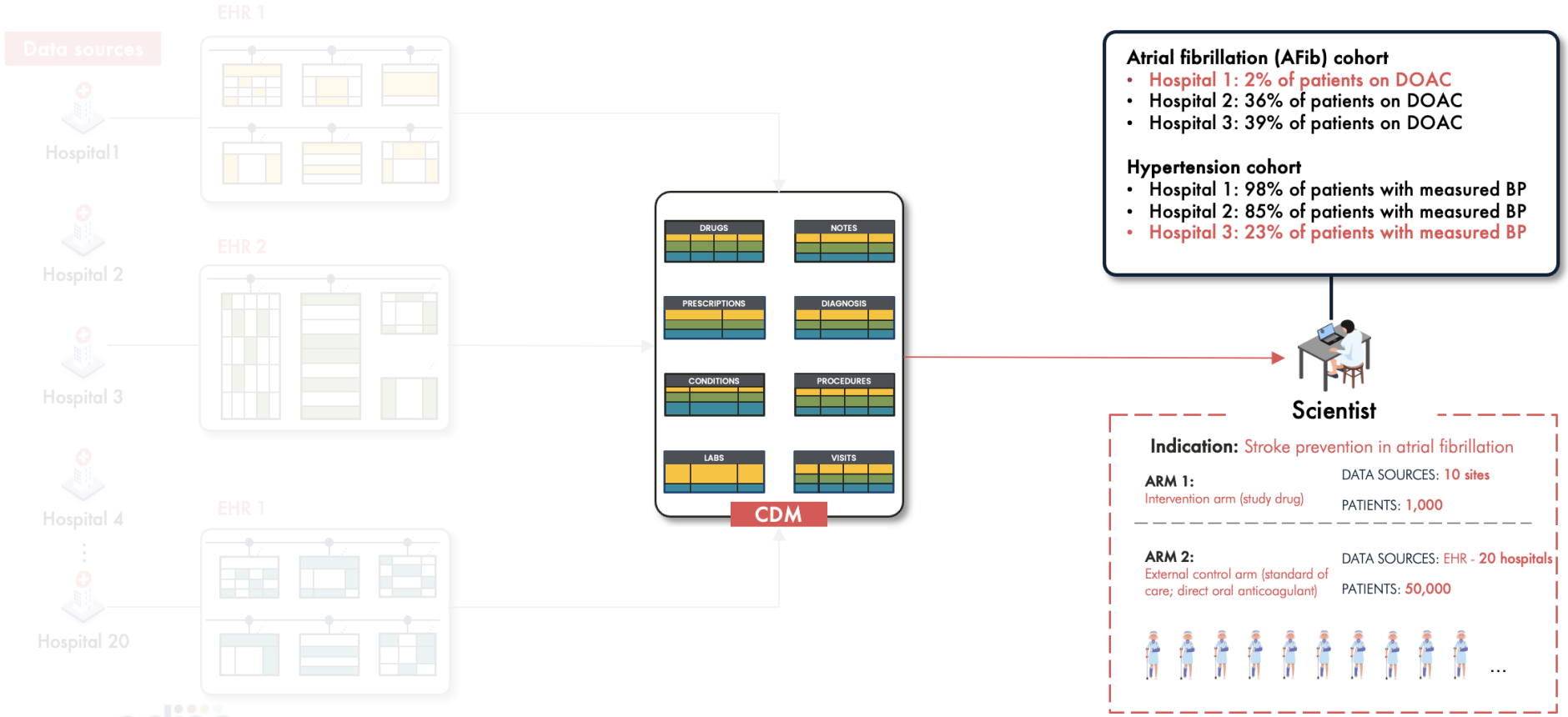
Data loss in common data models



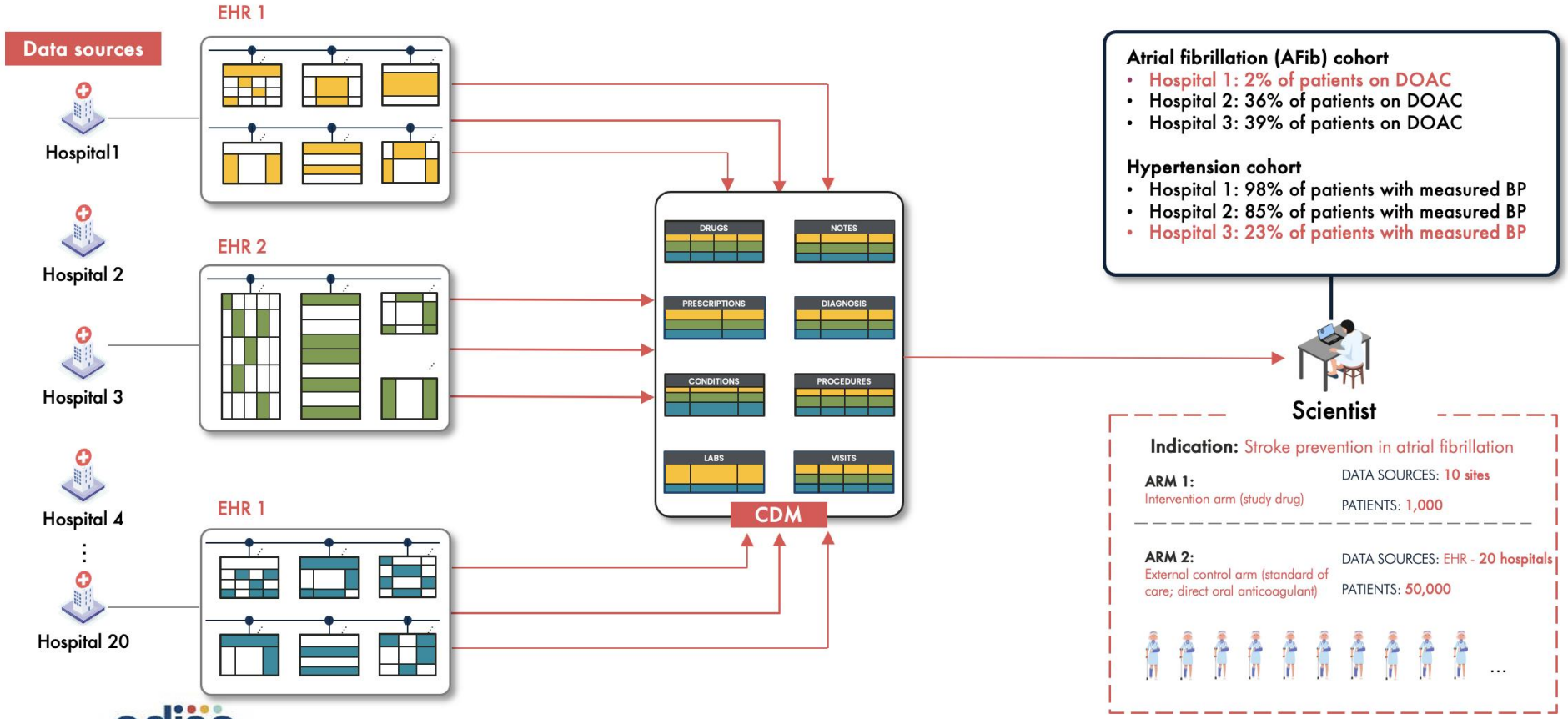
Data loss in common data models



Data loss in common data models

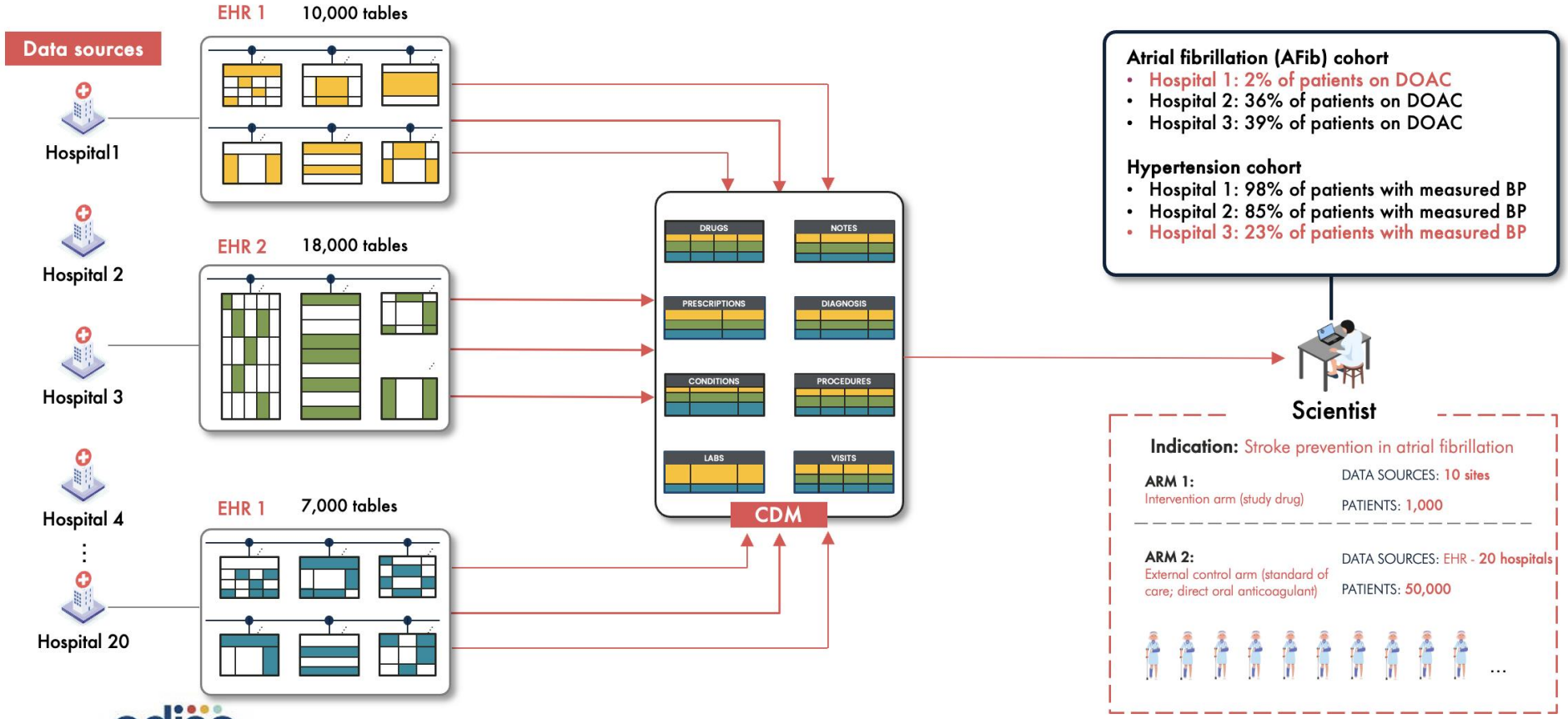


Data loss in common data models



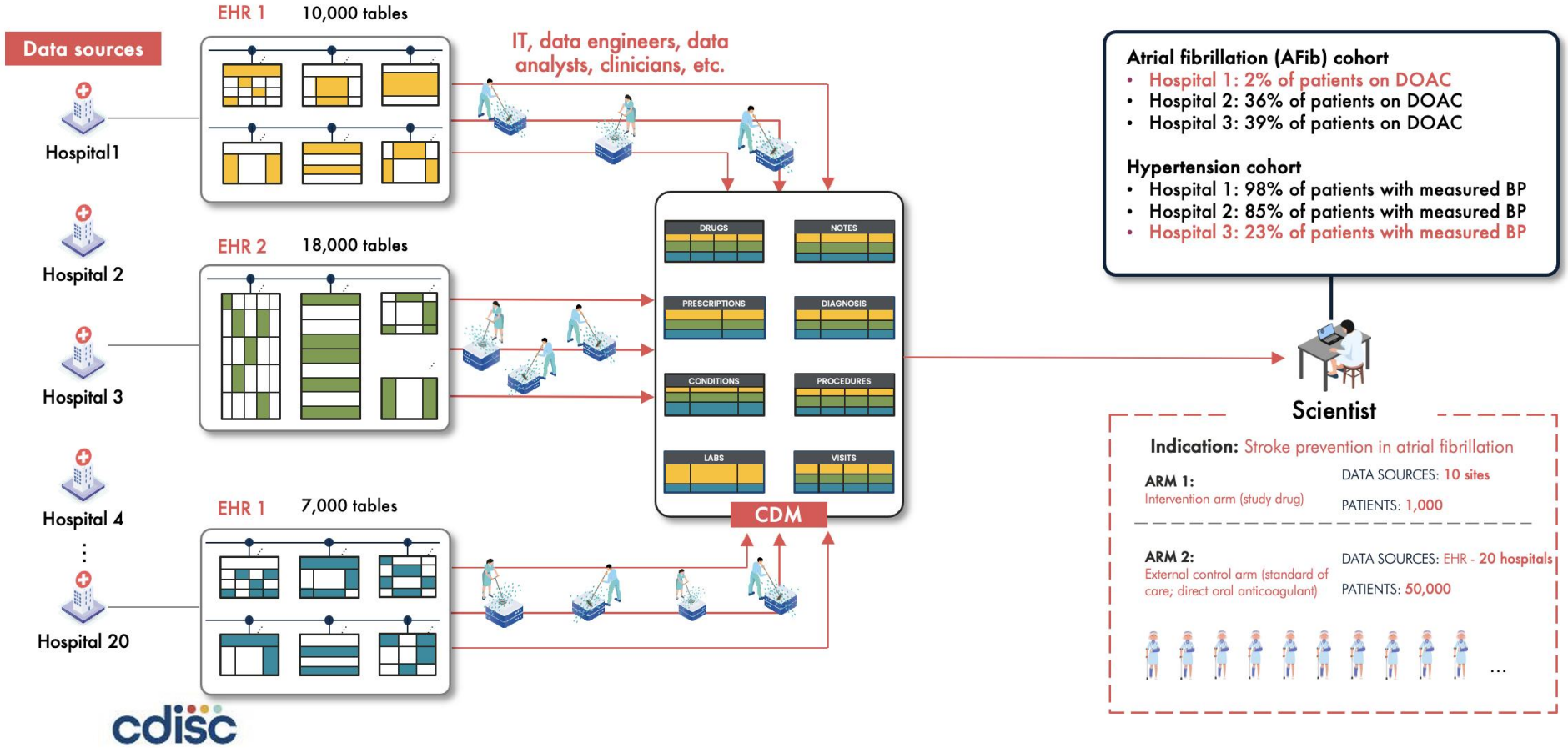
cdisc

Data loss in common data models



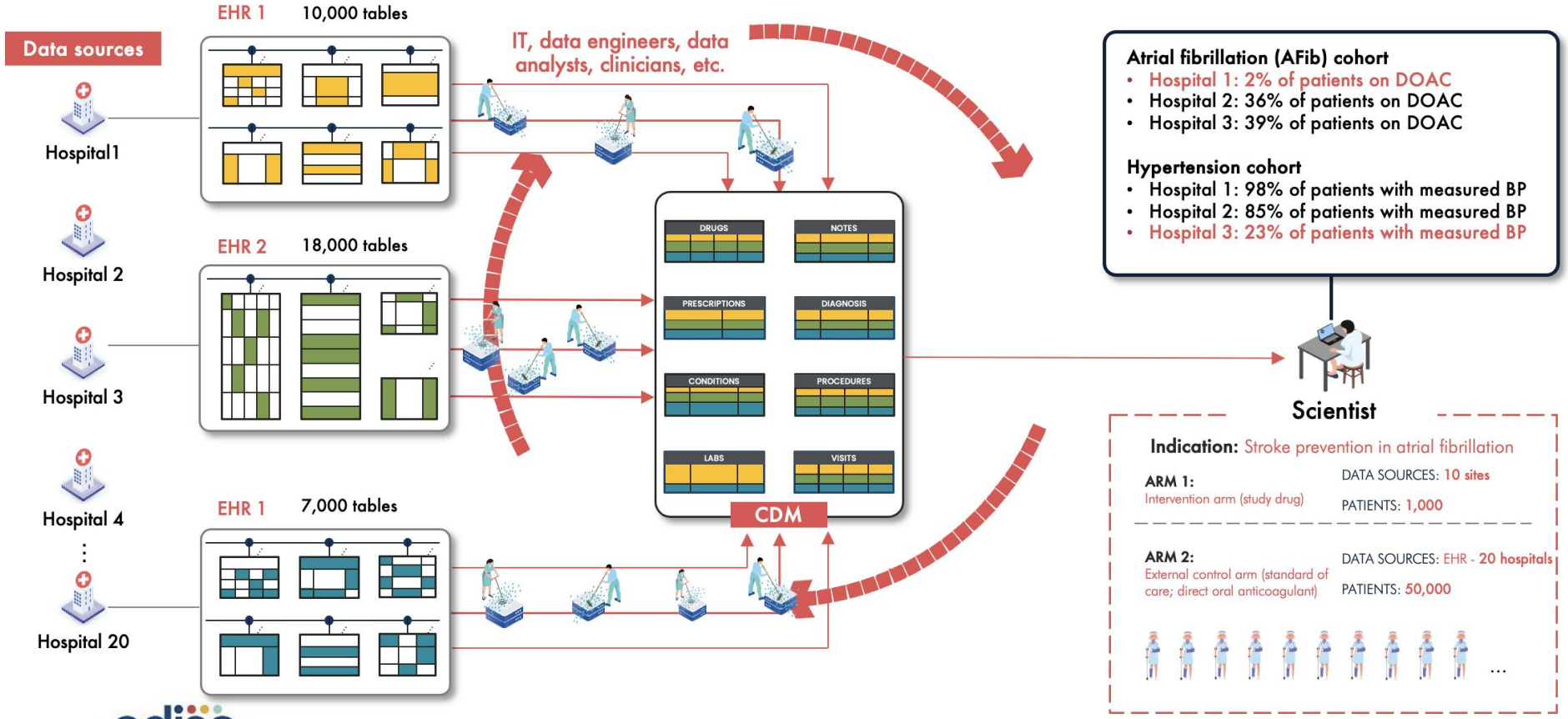
cdisc

Data loss in common data models



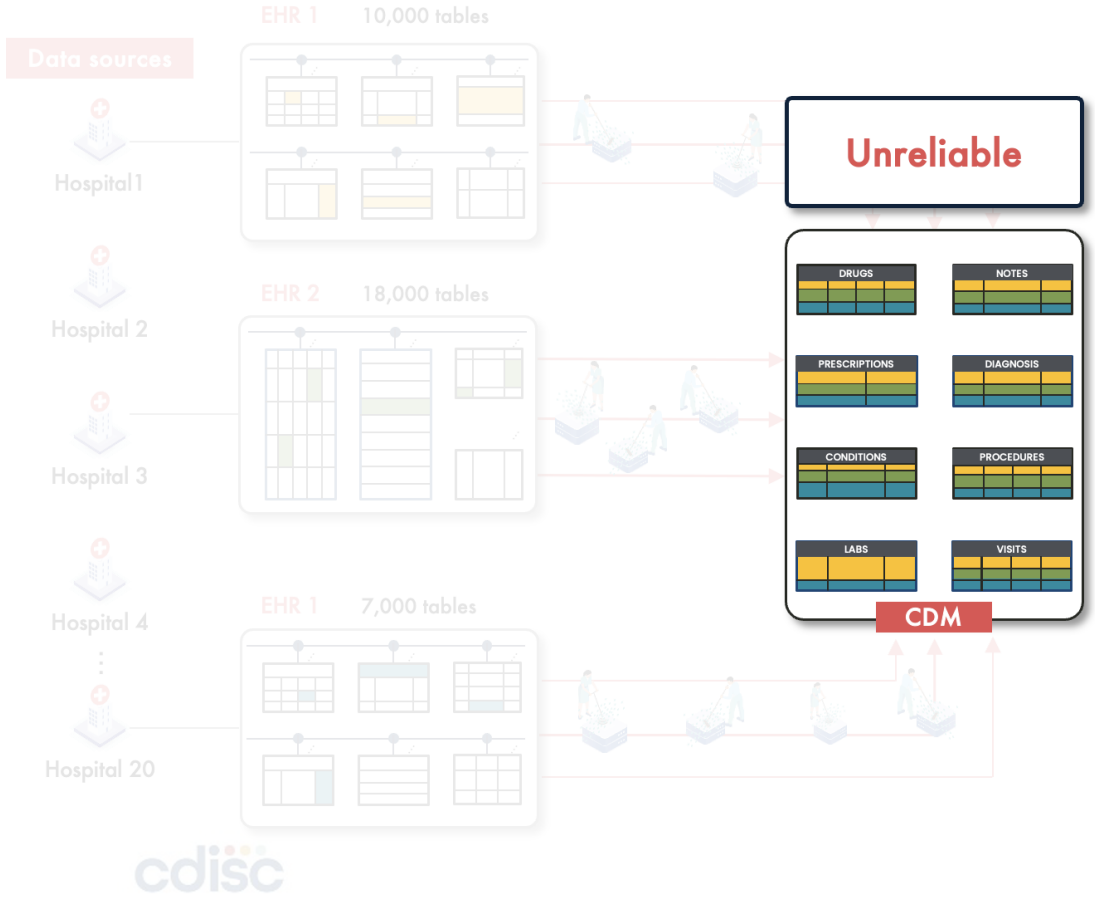
cdisc

Data loss in common data models



cdisc

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:
Intervention arm (study drug)

DATA SOURCES: 10 sites

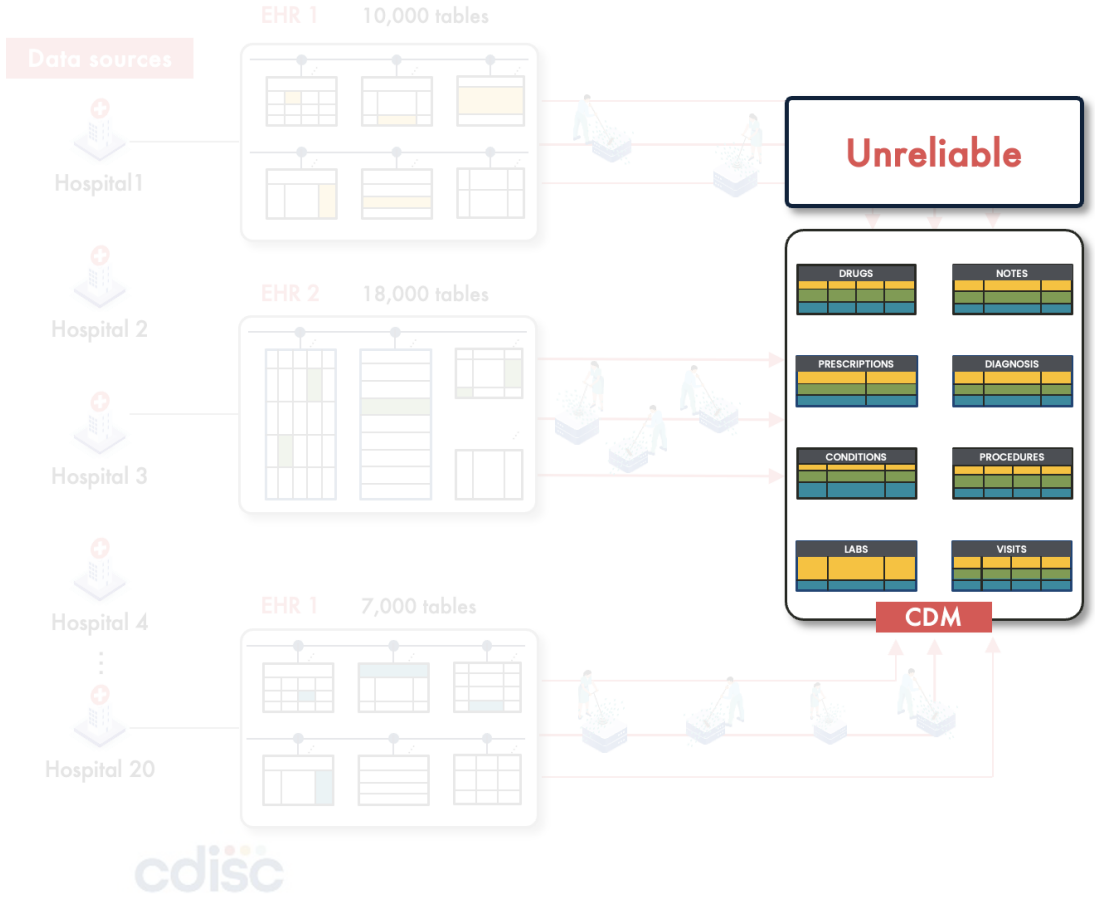
PATIENTS: 1,000

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



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:
Intervention arm (study drug)

DATA SOURCES: 10 sites

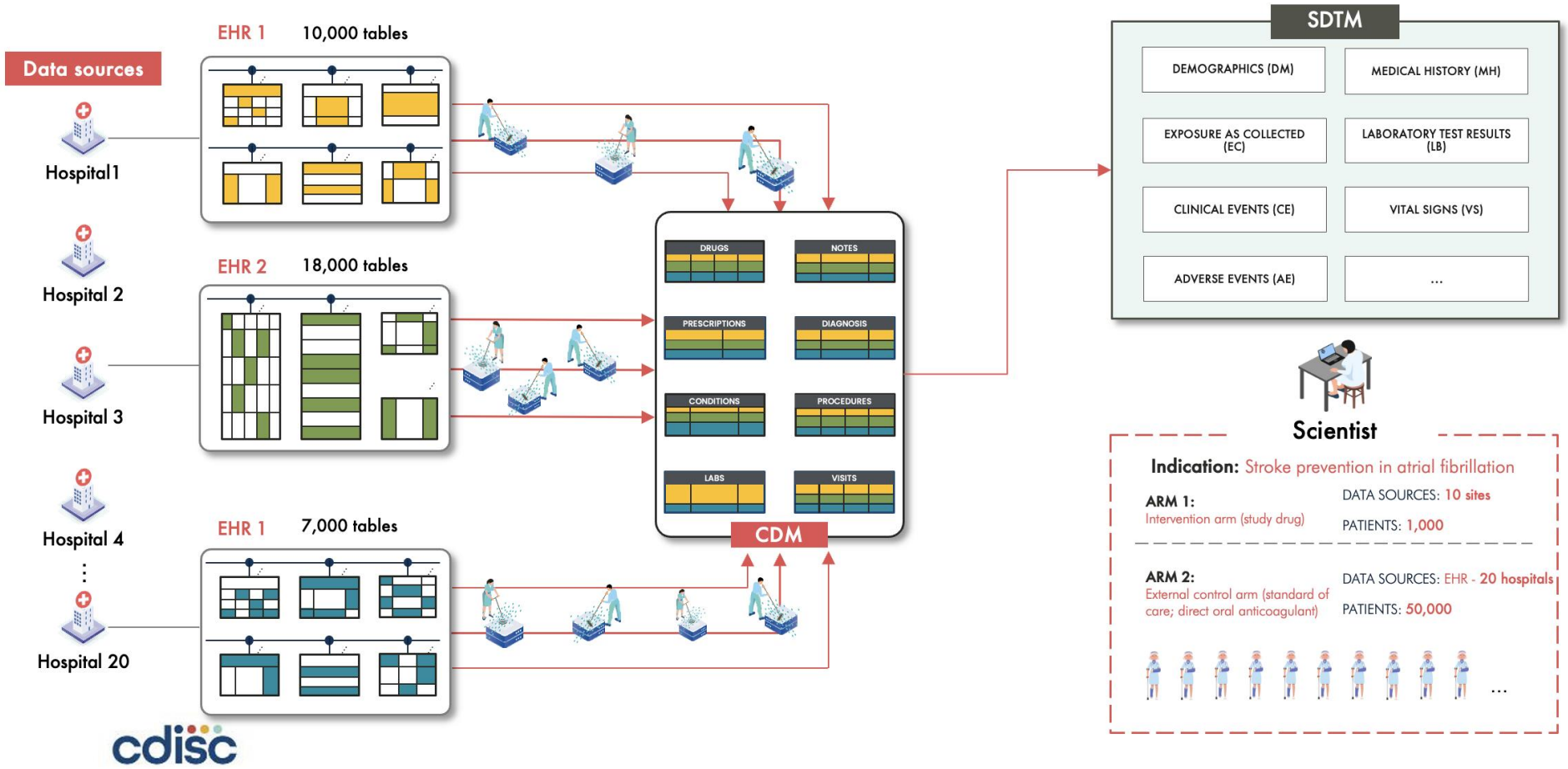
PATIENTS: 1,000

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

DATA SOURCES: EHR - 20 hospitals

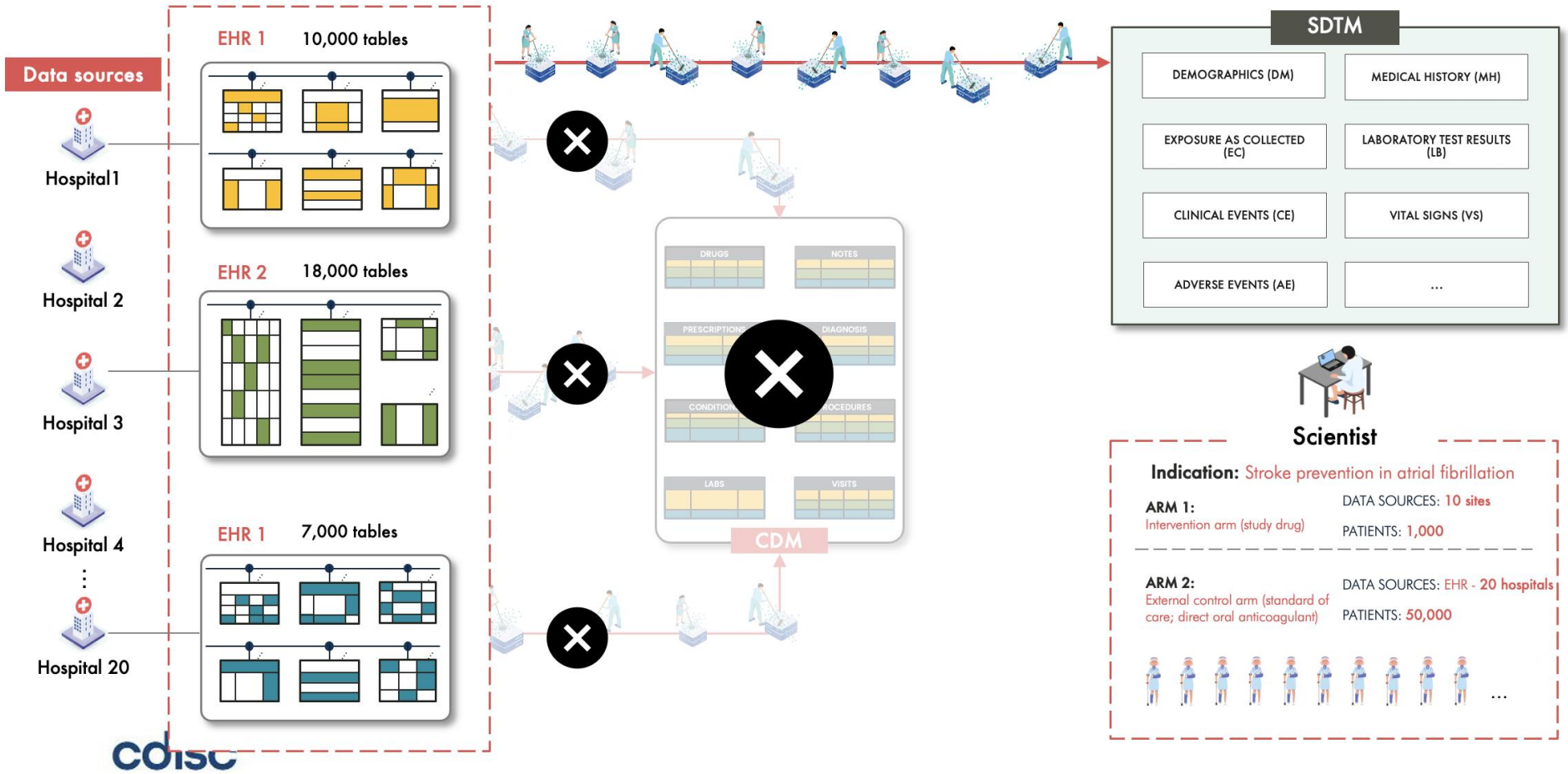
PATIENTS: 50,000

Data loss in common data models

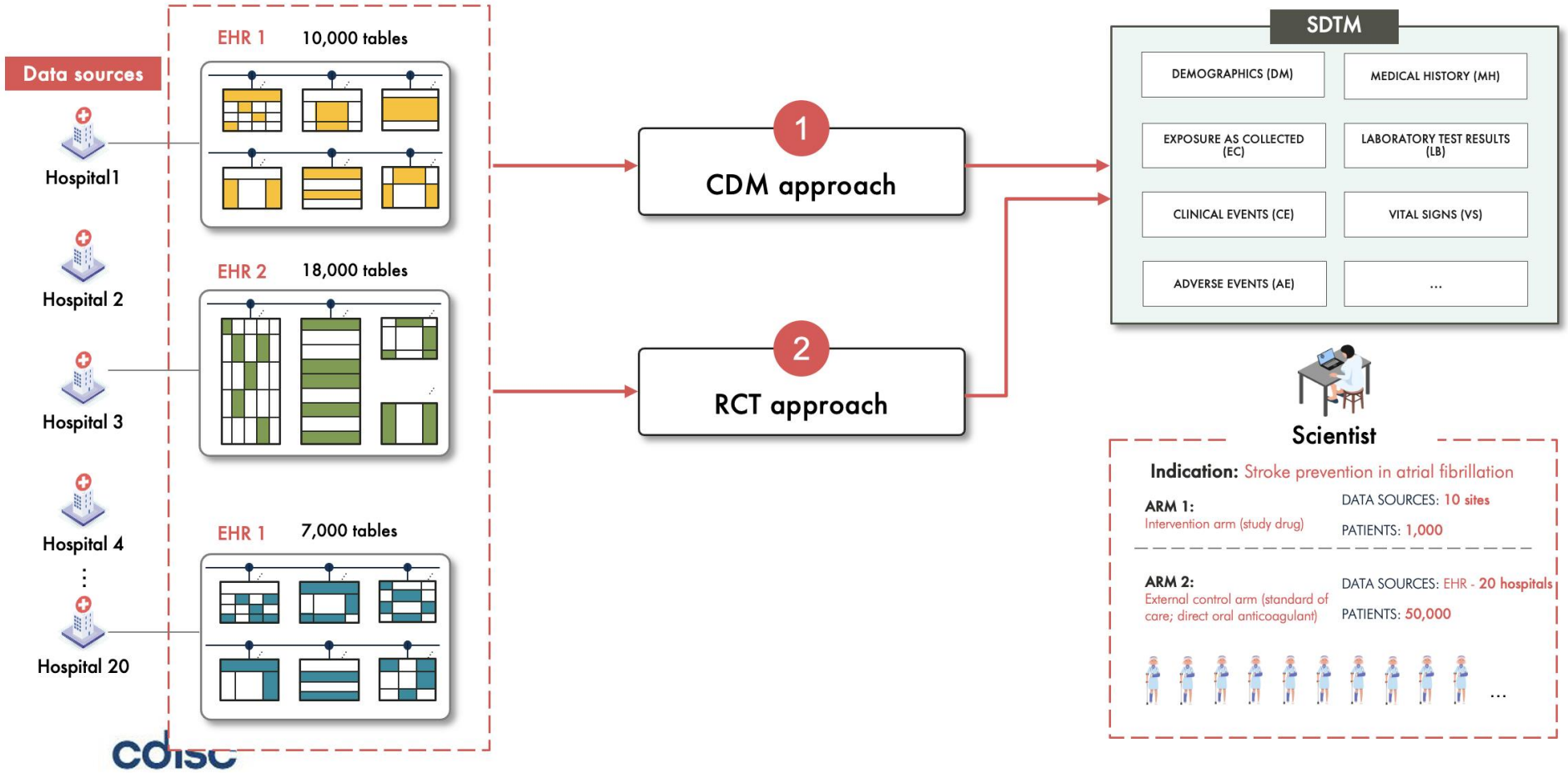


cdisc

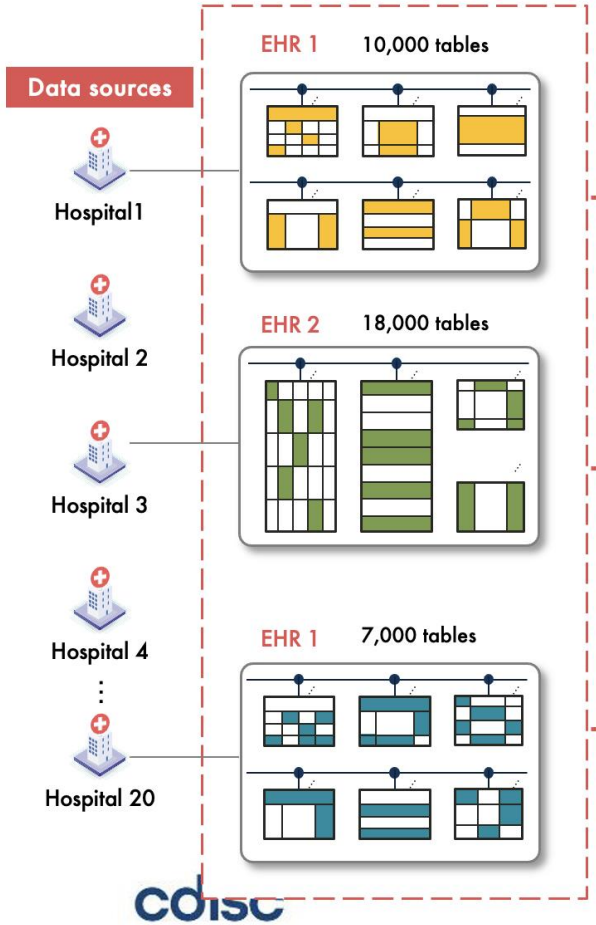
Traditional data preparation approach

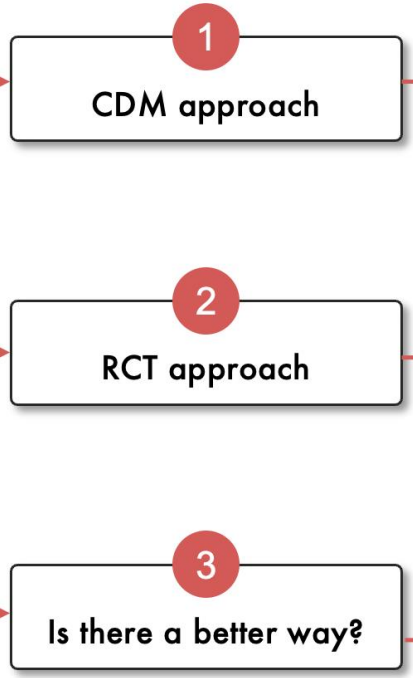


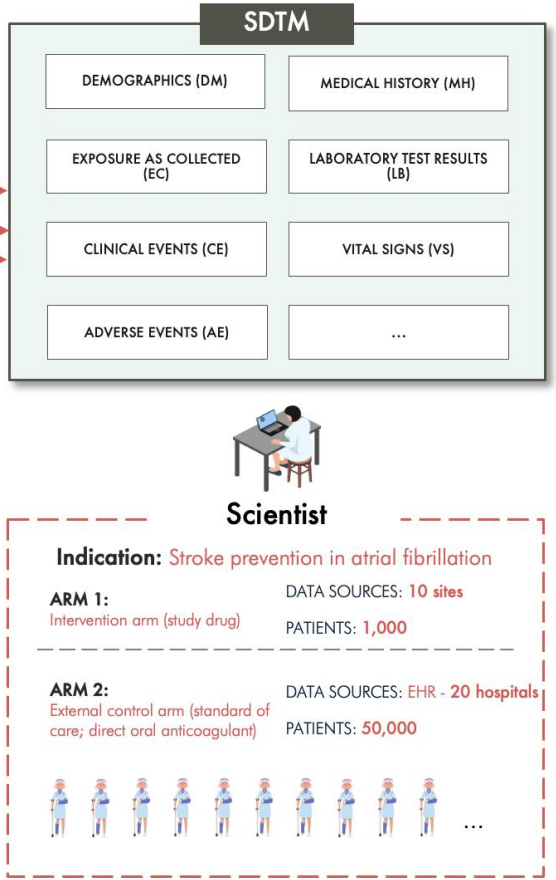
Options


Data sources
Hospital1
Hospital 2
Hospital 3
Hospital 4
Hospital 20
EHR 1
10,000 tables
EHR 2
18,000 tables
EHR 1
7,000 tables
1
CDM approach
2
RCT approach
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
...

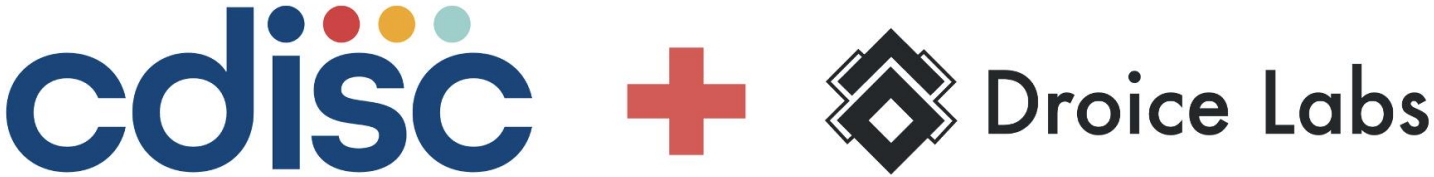
Options


Data sources
EHR 1
10,000 tables
Hospital1
Hospital 2
EHR 2
18,000 tables
Hospital 3
Hospital 4
EHR 1
7,000 tables
Hospital 20
cdisc


1
CDM approach
2
RCT approach
3
Is there a better way?


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
...

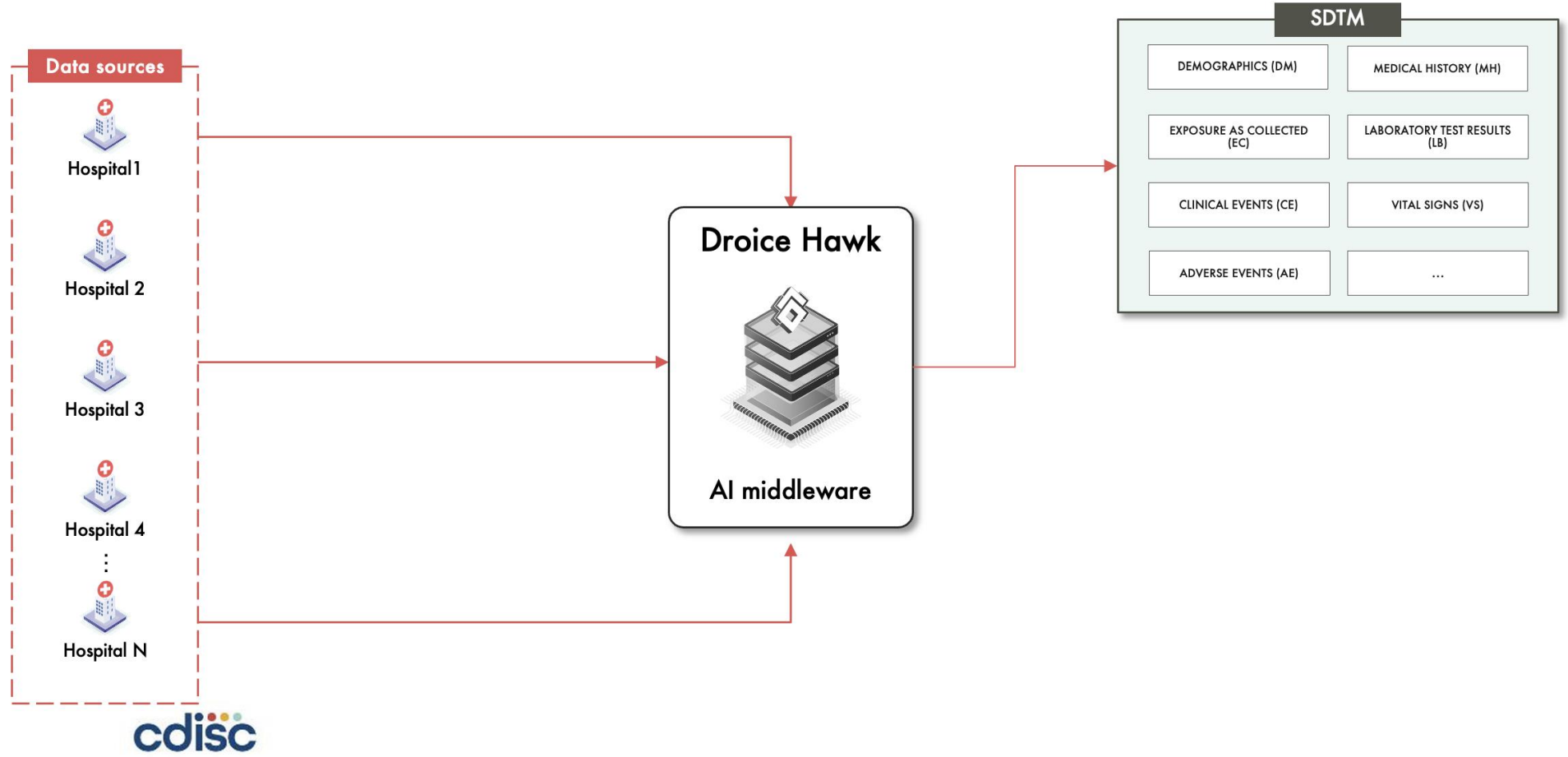
RWD: Straight to SDTM project



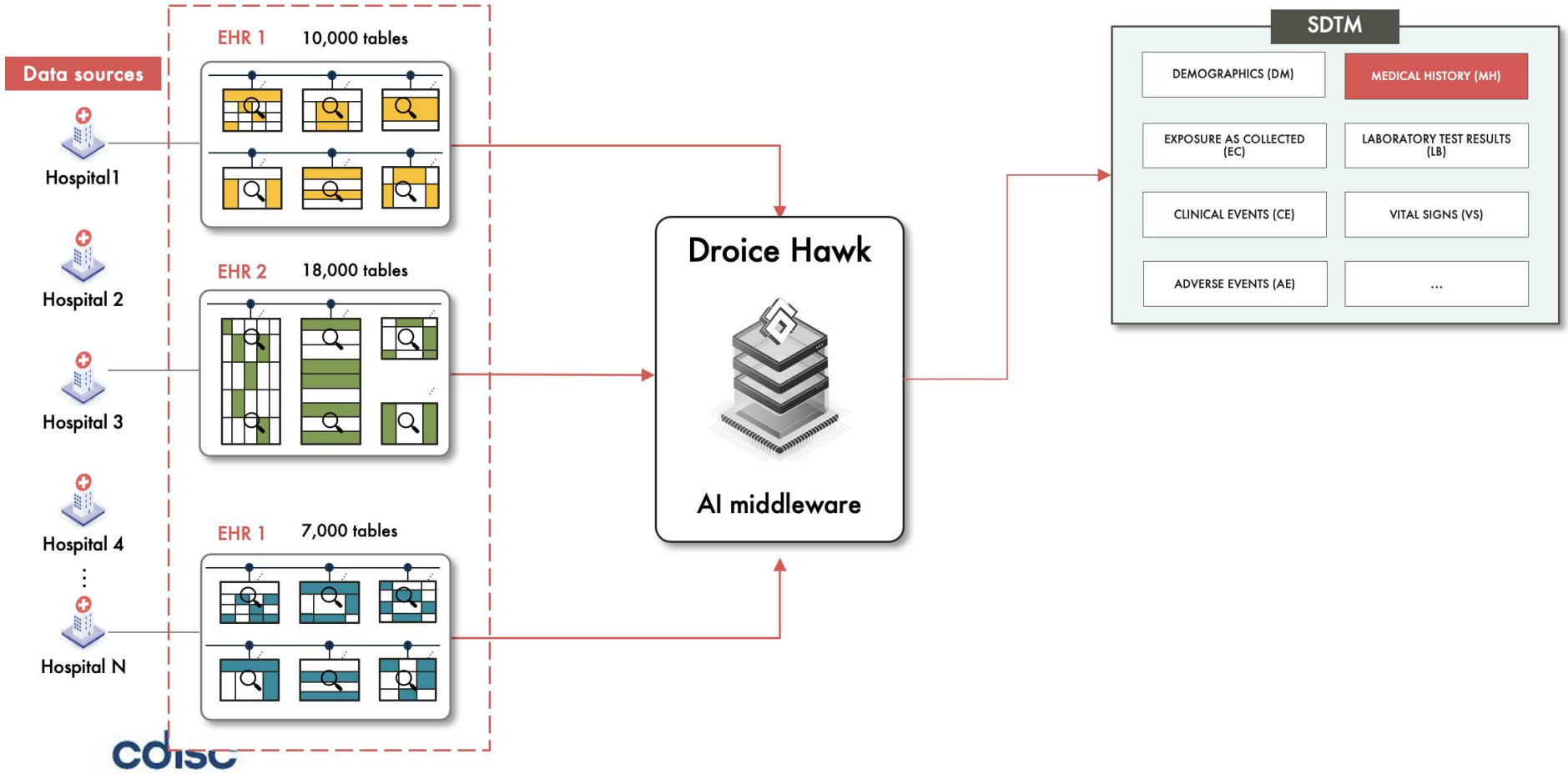
(standards expertise)

(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

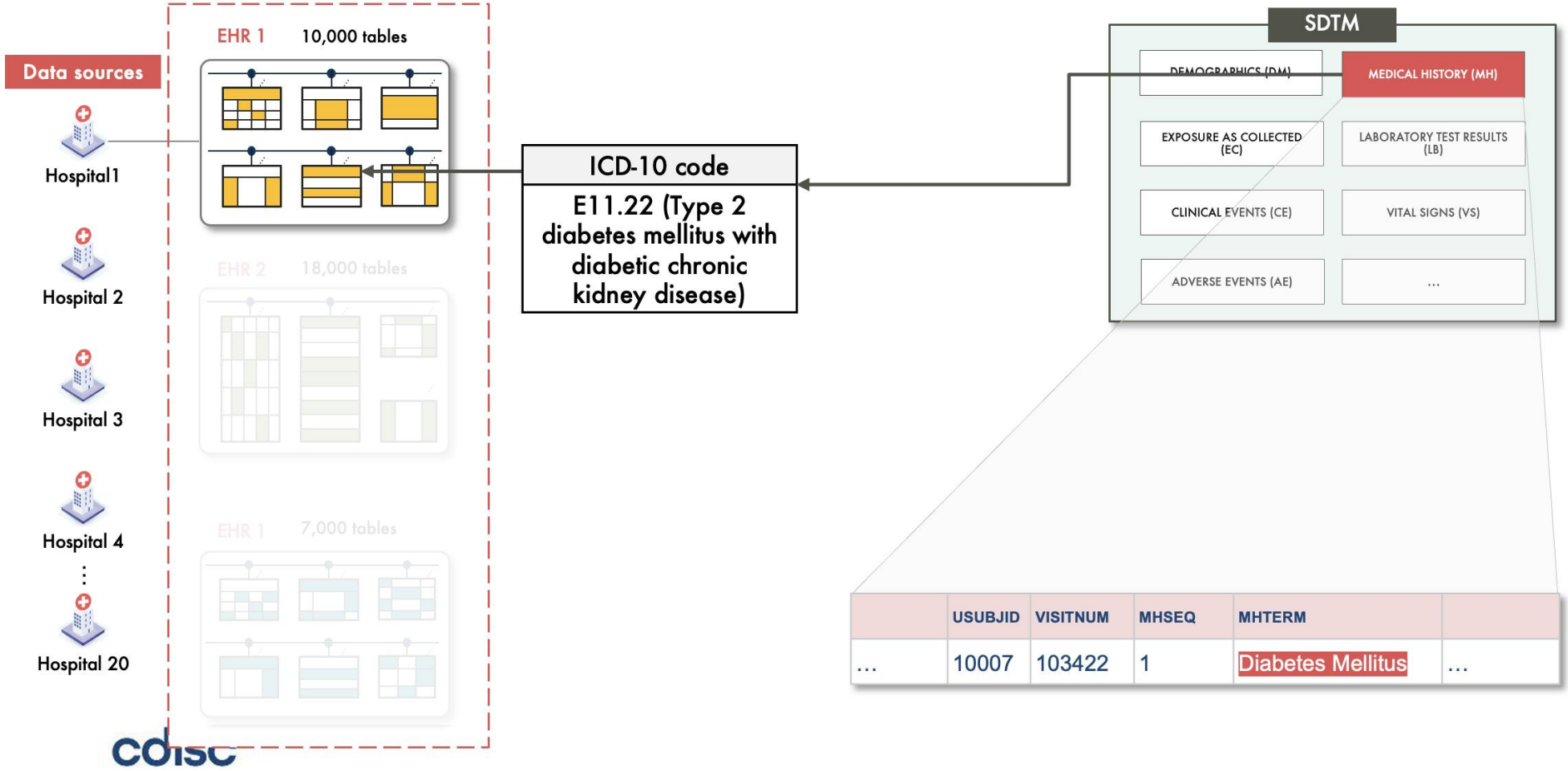
RWD: Straight to SDTM project

Medical history (MH)

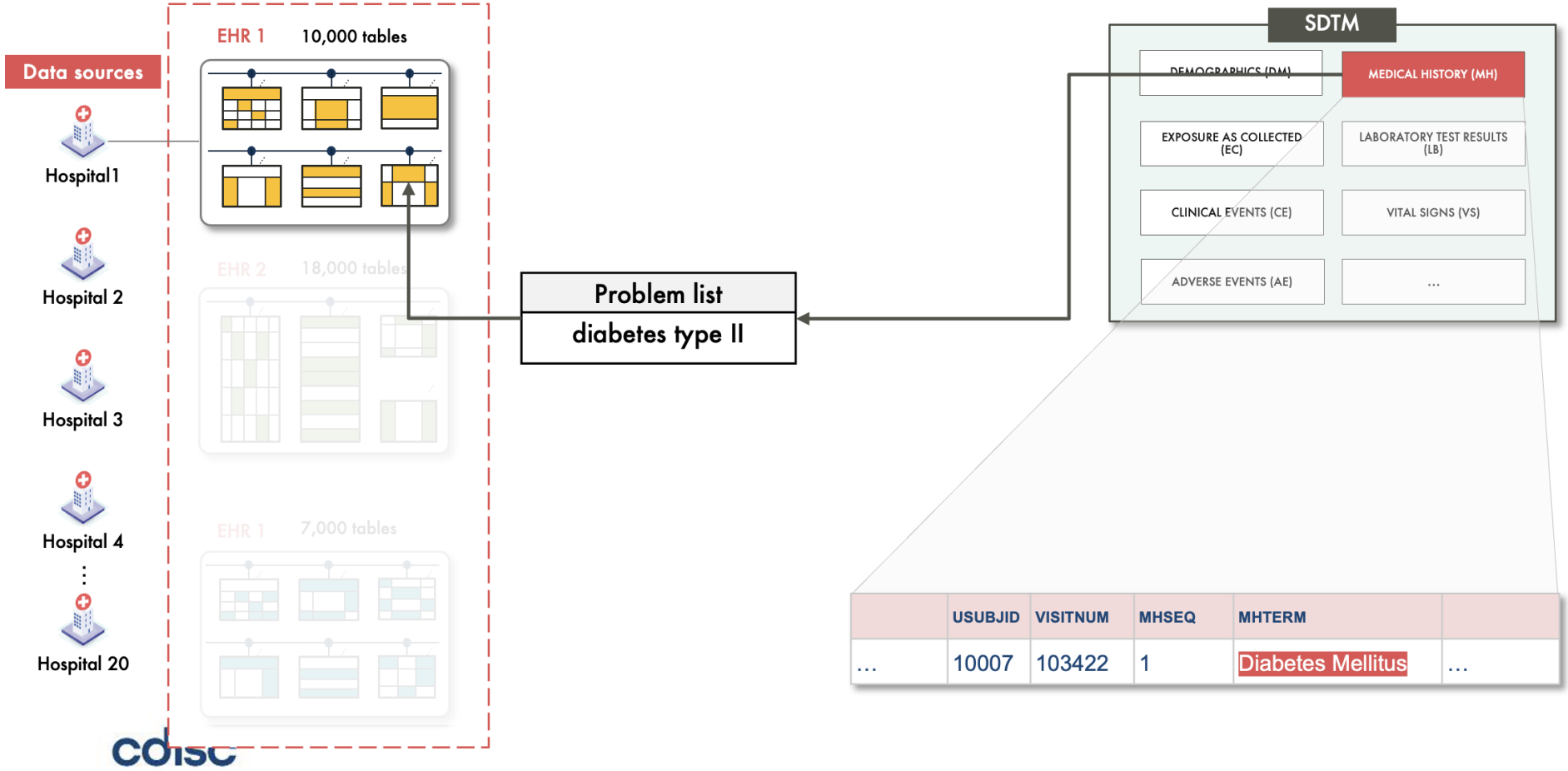
STUDYID	DOMAIN	USUBJID	VISITNUM	MHSEQ	MHTERM	MHCAT	MHDTC	MHSTDTC
EX-030	MH	10005	115407	8	Hypothyroidism	PAST MEDICAL HISTORY	2024-10-07	2024-10-07
EX-030	MH	10006	101357	1	Chronic Heart Failure	PAST MEDICAL HISTORY	2021-08-11	2021-08-11
EX-030	MH	10006	101357	2	Hypothyroidism	PAST MEDICAL HISTORY	2021-08-11	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

cdisc

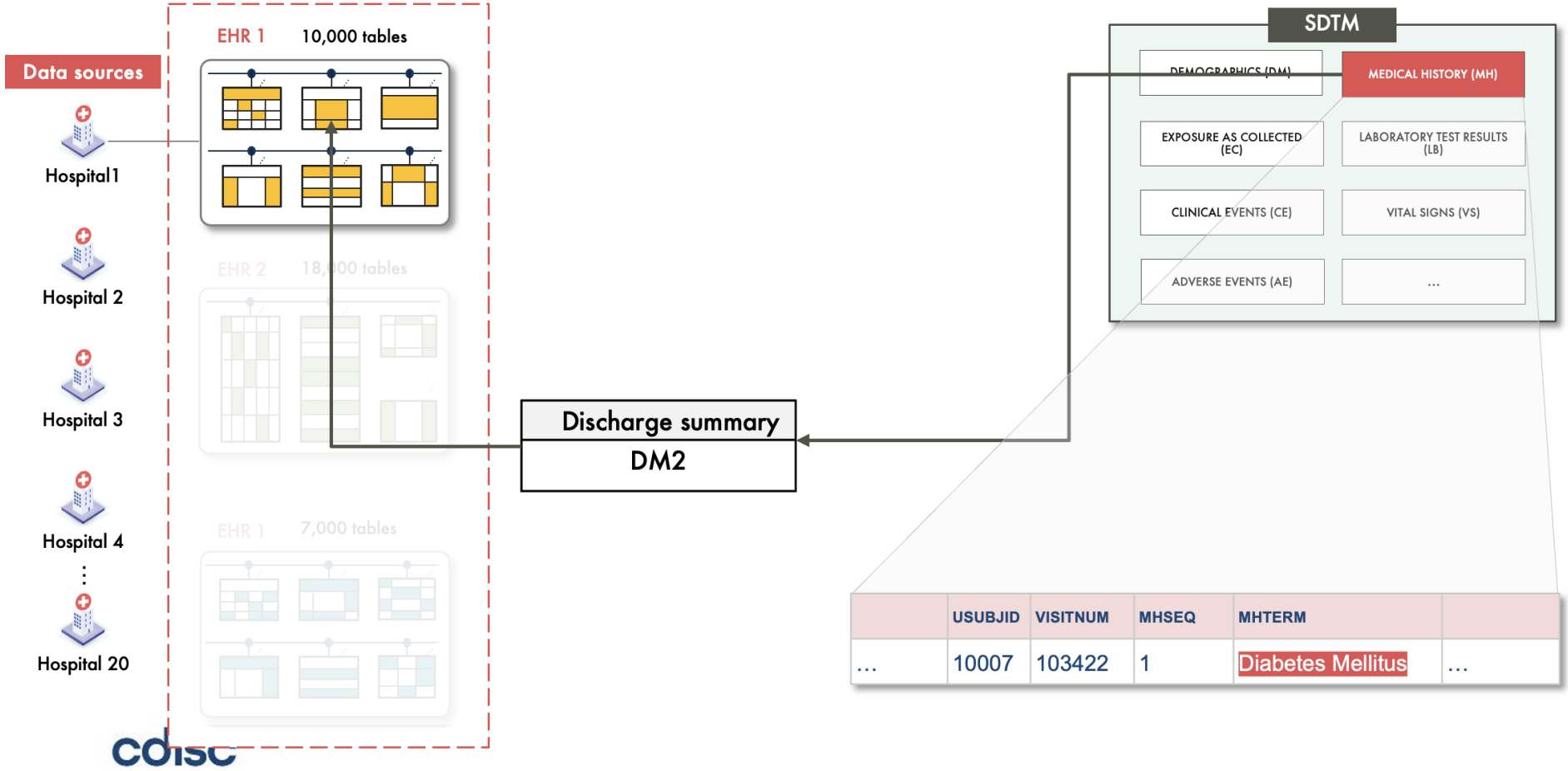
RWD: Straight to SDTM project



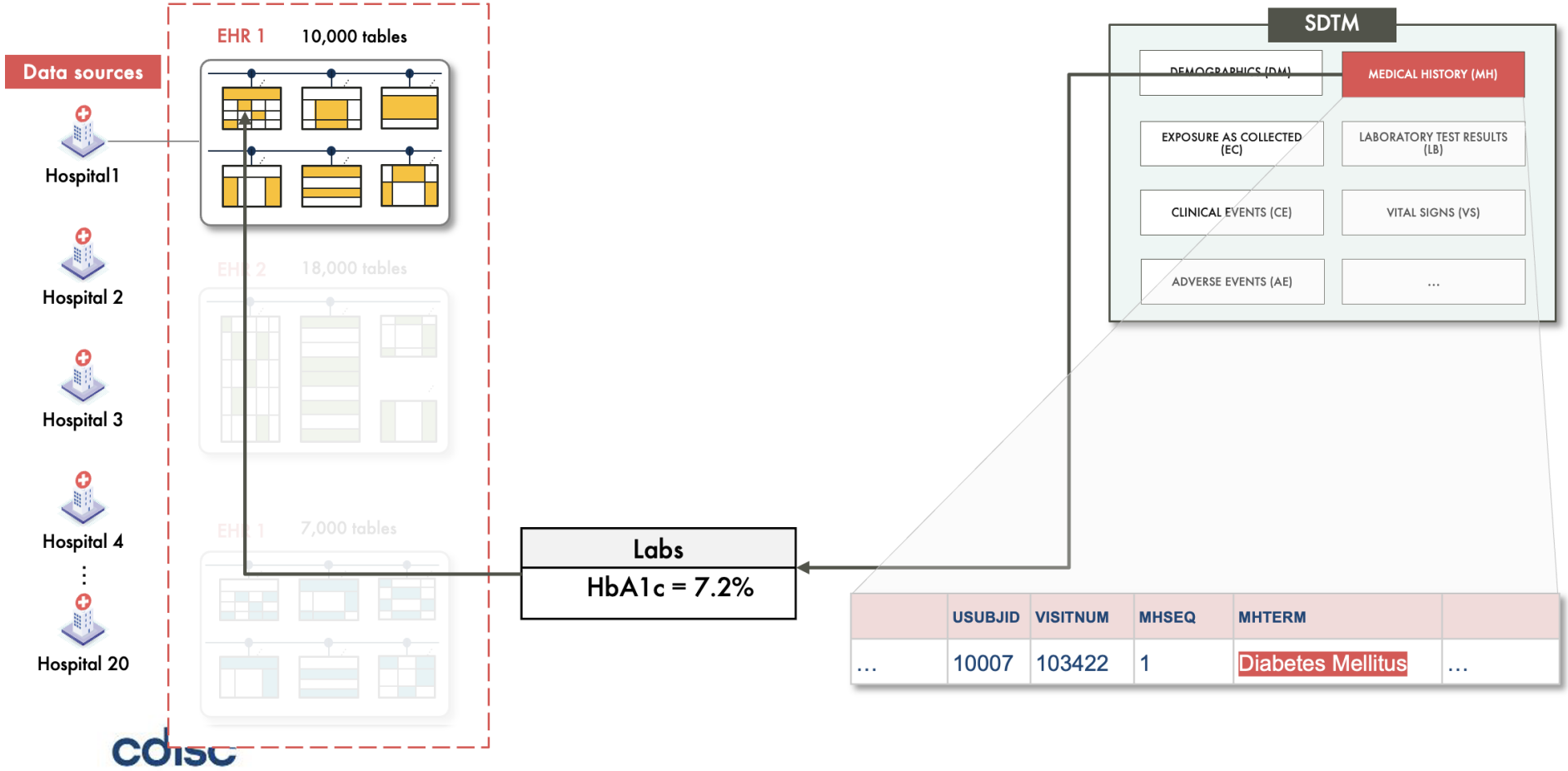
RWD: Straight to SDTM project



RWD: Straight to SDTM project



RWD: Straight to SDTM project

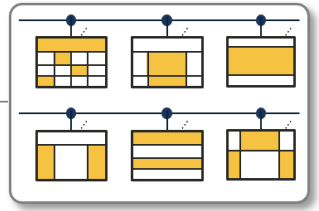


RWD: Straight to SDTM project

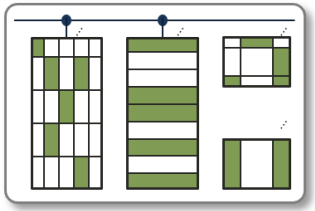
Data sources

- Hospital1
- Hospital 2
- Hospital 3
- Hospital 4
- ⋮
- Hospital 20

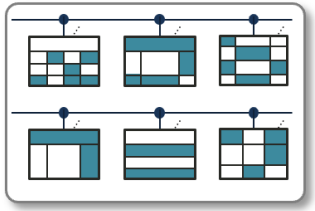
EHR 1 10,000 tables



EHR 2 18,000 tables



EHR 1 7,000 tables



SDTM

- DEMOGRAPHICS (DM)
- MEDICAL HISTORY (MH)
- EXPOSURE AS COLLECTED (EC)
- LABORATORY TEST RESULTS (LB)
- CLINICAL EVENTS (CE)
- VITAL SIGNS (VS)
- ADVERSE EVENTS (AE)
- ...

FDA

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

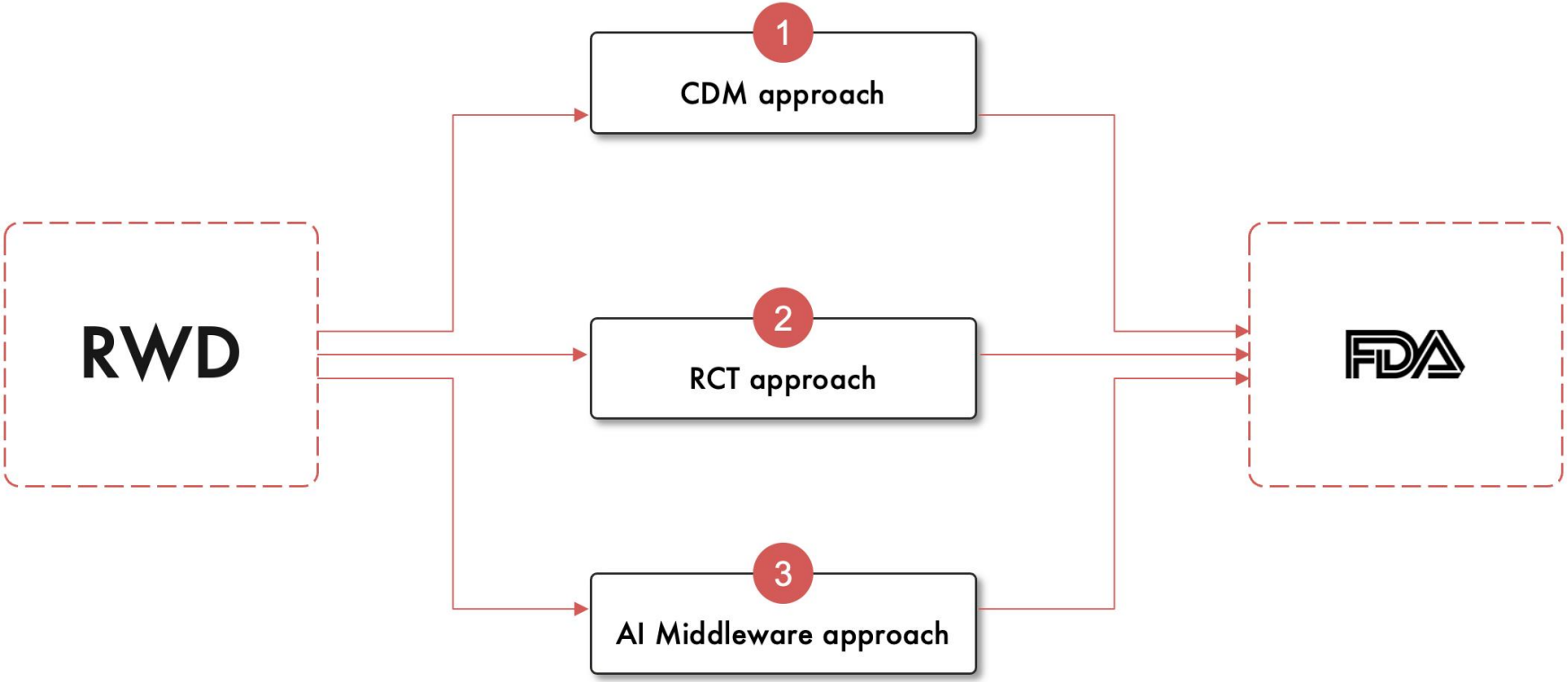
Guidance for Industry

DRAFT GUIDANCE

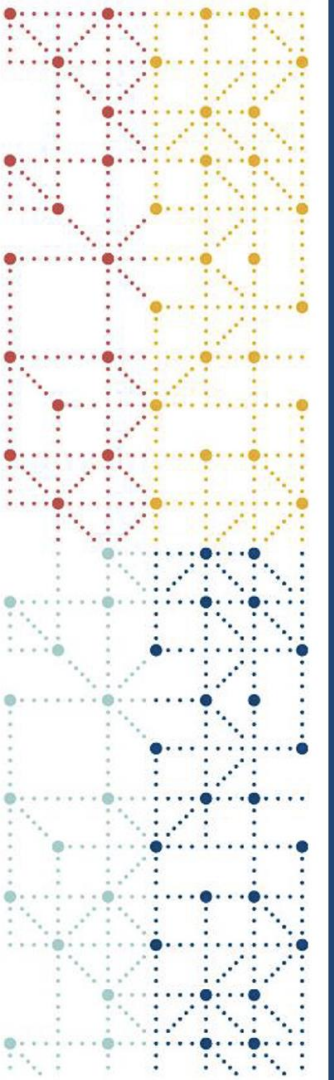
> *Provenance of core data elements to allow tracking of these elements back to their respective points of origin, with clear documentation of modifications that may have occurred. (R 1097-1099)*

CDISC

Takeaways



cdisc



Thank You!