



# Data quality of OHDSI APAC: CDM Inspection study

2022-05-24

Community call

OHDSI APAC Study Team 4



# Background

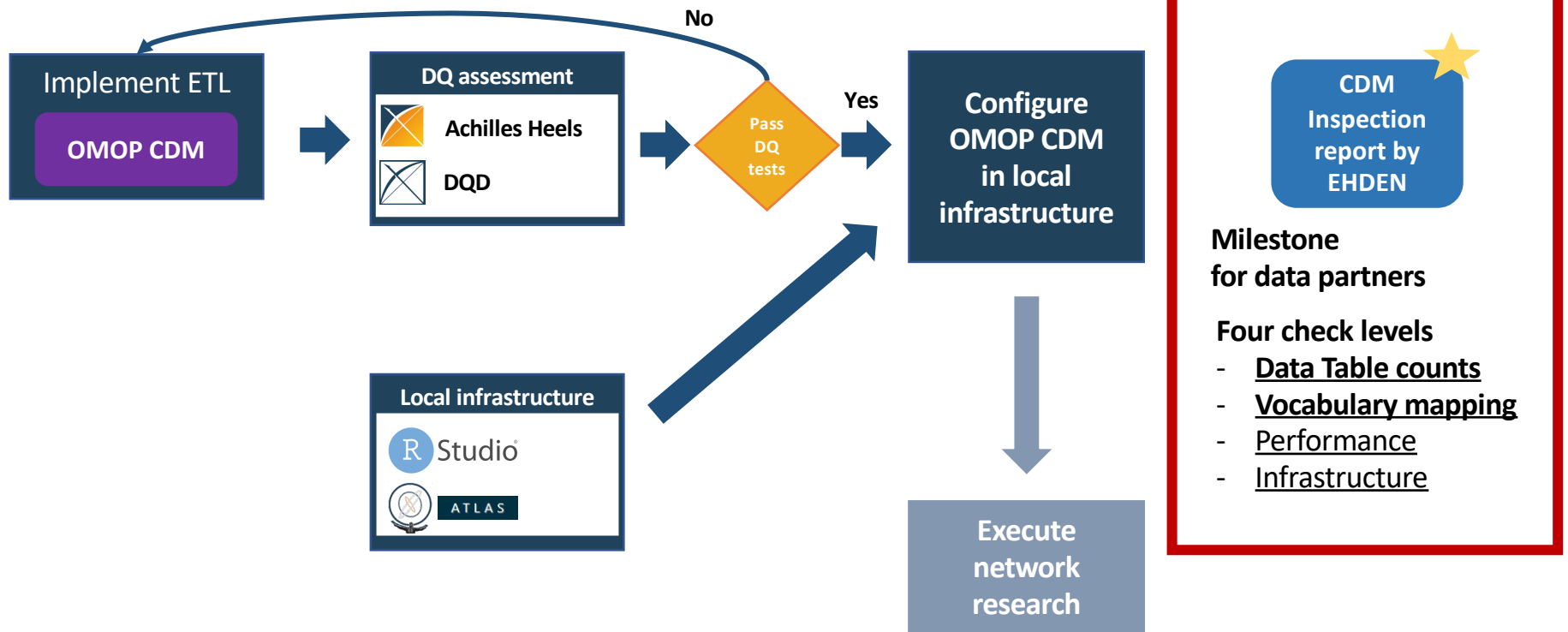


Figure is recreated from Maxim Moinat slides (21/11/10, at OHDSI community call)



# Objectives

## **What is this study for?**

- Collecting CDM Inspection reports from APAC community

## **Why this study is needed?**

- To check the current status of CDMs, get insights from the CDMs, and improve their data quality

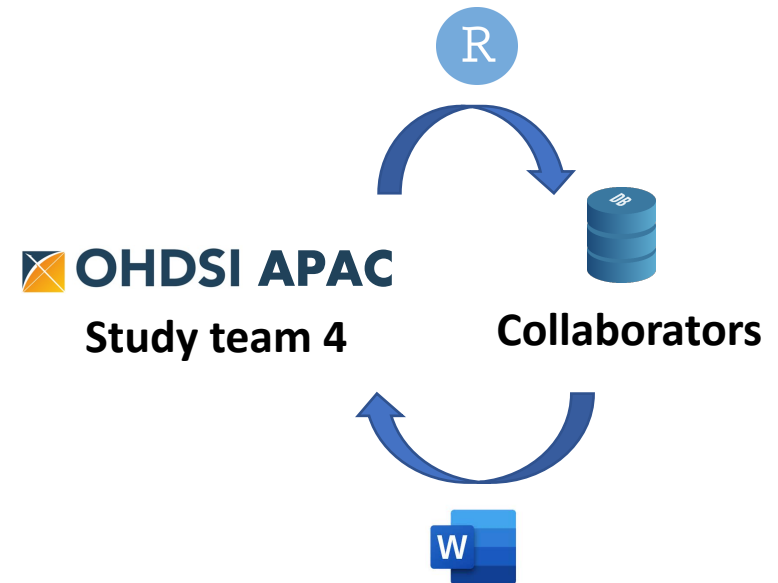
## **What is the final goal?**

- Disclosure of current status of conversion, contents, and data distribution of CDMs of the OHDSI APAC community.
- To provide the basic statistics which can be used as references for future CDM conversion



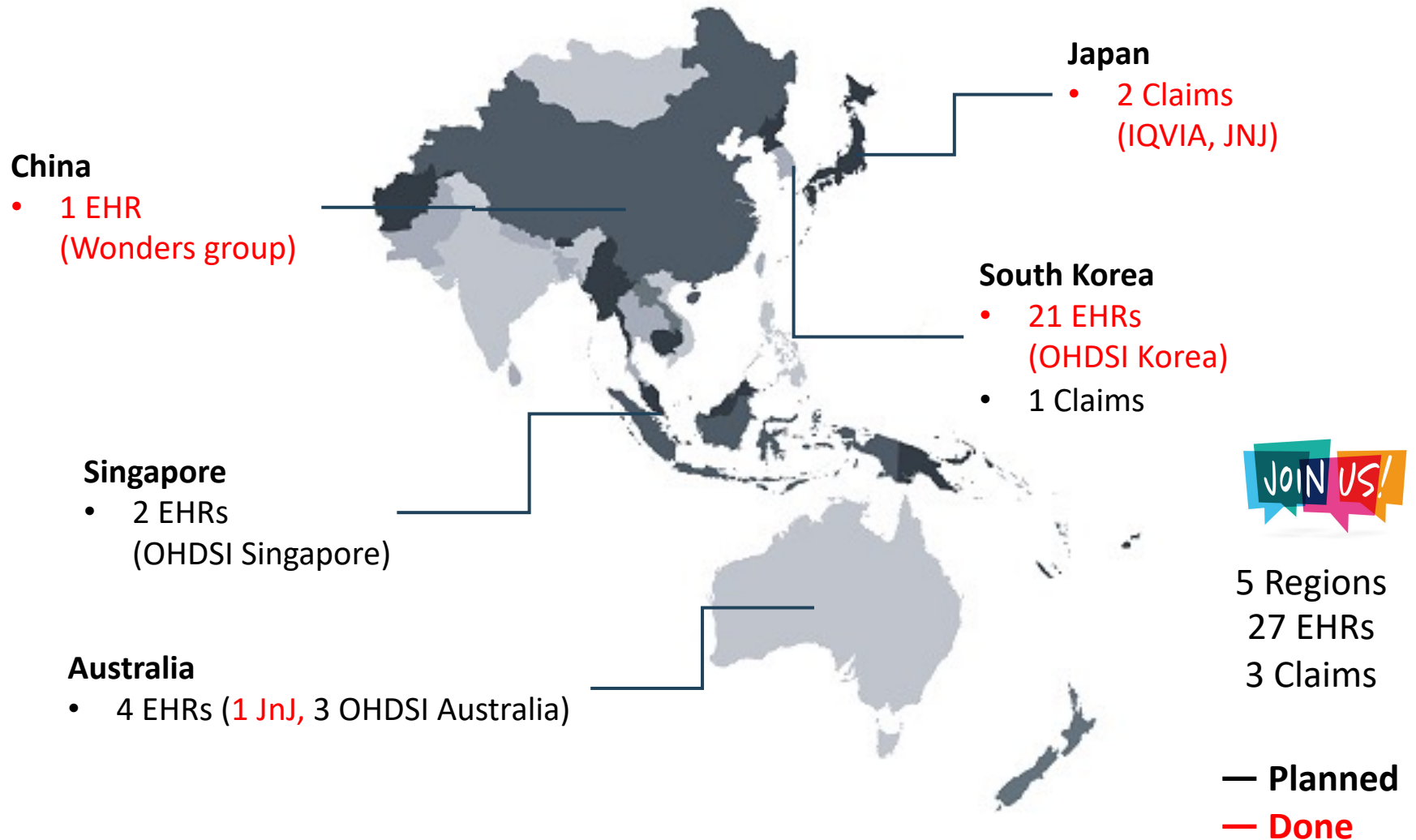
# Study package

- Data sources: CDM databases from OHDSI APAC community
- Collecting inspection reports from each site.
- R package for automatically creating inspection reports.
- Collectibles
  - Number of record, person
  - Number of unique concepts per person
  - Source-CDM mapping ratio
  - Proportion of standard concepts in mapped codes
  - Drug mapping level (granularity)
  - Frequent concept list in each domain
  - Achilles heel result (error / notification / warnings)
  - Sample cohort generation





# Study participants



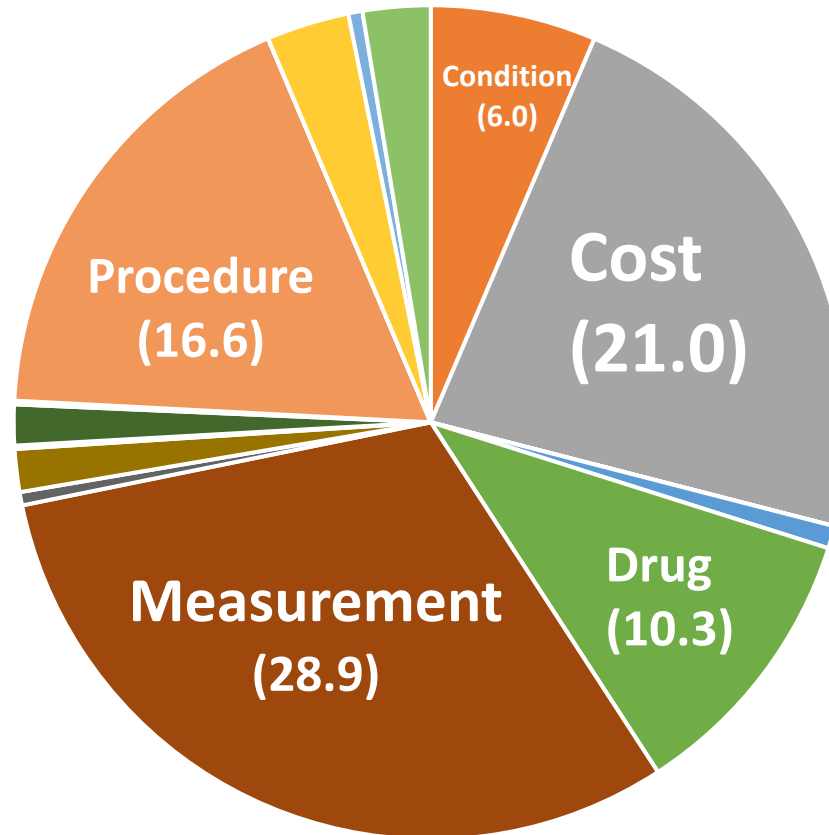


# Preliminary results

Records Proportion (%) in 23 databases

**Total n of records  
= 36,096,359,491**

**Total n of persons  
= 48,127,420**



- Care site
- Condition occurrence
- Cost
- Death
- Device exposure
- Drug exposure
- Location
- Measurement
- Note
- Observation
- Observation period
- Payers plan period
- Person
- Procedure occurrence
- Provider
- Specimen
- Visit details
- Visit occurrence



# Preliminary results

**Table 2. Summary result of record mapping to the OMOP concept from common data model databases**

| Domain               | Mapping codes / source codes | Mapped records / total records | Mapped as standard / Mapped records |
|----------------------|------------------------------|--------------------------------|-------------------------------------|
|                      | Median [Q1, Q3]              | Median [Q1, Q3]                | Median [Q1, Q3]                     |
| Condition occurrence | 97.5 [87.6, 99.6]            | 99.5 [94.8, 100.0]             | 100.0 [99.0, 100.0]                 |
| Device exposure      | 56.6 [48.6, 78.2]            | 75.4 [64.7, 92.2]              | 79.3 [54.2, 100.0]                  |
| Drug exposure        | 85.0 [74.8, 90.6]            | 97.0 [96.0, 98.2]              | 98.3 [97.7, 99.1]                   |
| Measurement          | <b>50.1 [25.4, 87.1]</b>     | <b>97.0 [69.1, 99.7]</b>       | 100.0 [99.7, 100.0]                 |
| Measurement-unit     | 96.7 [29.7, 100.0]           | 100.0 [40.8, 100.0]            | 100.0 [100.0, 100.0]                |
| Measurement-value    | <b>13.3 [4.4, 46.0]</b>      | <b>7.5 [4.1, 49.4]</b>         | 100.0 [100.0, 100.0]                |
| Observation          | 100.0 [98.3, 100.0]          | 100.0 [93.0, 100.0]            | 100.0 [100.0, 100.0]                |
| Observation-unit     | 100.0 [61.1, 100.0]          | 100.0 [50.4, 100.0]            | 97.8 [44.4, 100.0]                  |
| Observation-value    | 50.0 [50.0, 100.0]           | 92.1 [77.8, 100.0]             | 100.0 [100.0, 100.0]                |
| Procedure occurrence | <b>66.3 [57.5, 97.2]</b>     | <b>38.7 [22.4, 95.1]</b>       | 100.0 [85.1, 100.0]                 |
| Visit occurrence     | 100.0 [100.0, 100.0]         | 100.0 [100.0, 100.0]           | 100.0 [100.0, 100.0]                |

**Table 3. Vocabulary granularity in drug exposure table**

| Vocabulary         | Classification      | N of records       | Mapped records / total records, Median [Q1, Q3] |
|--------------------|---------------------|--------------------|---|
| RxNorm (Extension) | <b>Branded Drug</b> | <b>607,710,428</b> | <b>40.1 [1.1, 57.1]</b>                         |
|                    | Clinical Drug       | 722,805,309        | <b>9.4 [2.3, 46.6]</b>                          |
|                    | Quant Branded Drug  | 266,331,050        | <b>17.3 [0.3, 30.6]</b>                         |



Thank you



Chungsoo Kim  
ted9219@ajou.ac.kr