Use of electronic health records to evaluate treatment pathways –
a Common Data Model approach

Tan Hui Xing
Data Analyst
Vigilance and Compliance Branch
Health Sciences Authority, Singapore
Objectives

• Demonstrate the use of a common data model (CDM) to harmonize electronic medical records (EMR)
  – To transform data from an acute care hospital in Singapore to the CDM developed by the Observational Medical Outcomes Partnership (OMOP)
  – To apply analytical tools on the data to uncover treatment patterns of patients newly diagnosed with diabetes mellitus, hypertension or depression.
Methodology

1. Conversion of source data files to the OMOP CDM
2. Modified query on the transformed data
Conversion of source data files to OMOP CDM

• Data source:
  – Approximately 250,000 patients
  – Tertiary care hospital in Singapore
  – Data from January 2013 to December 2016 comprising the following:
    • 1.1 million rows of diagnoses
    • 5.2 million rows of ordered medications
    • 15.5 million lab records
Conversion of source data files to OMOP CDM

• Data vocabularies employed:
  – Systematic Nomenclature of Medicine Clinical Terms (SNOMED CT) for diagnosis codes,
  – **RxNorm** Extension for drugs, and
  – Logical Observation Identifiers Names and Codes (LOINC) for laboratory tests and vitals measurements

• **Extract, Transform, Load (ETL)** process

• Mapped to OMOP table based on “Concept Name”
Conversion of source data files to OMOP CDM

- Percentage of records converted:

| OMOP CDM Tables       | Source Tables                   | Number of rows of records | Number of rows of records | Proportion migrated |
|-----------------------|---------------------------------|---------------------------|---------------------------|--------------------|
| Table name            |                                 |                           |                           |                    |
| person                | t_demographics                  | 245,561                   | 258,038                   | 95.2%              |
| condition_occurrence  |                                 |                           |                           |                    |
| Primary: 210,830      | t_primary_diagnosis             | 222,554                   |                           | 94.7%              |
| Secondary: 799,169    | t_secondary_diagnosis           | 839,265                   |                           | 95.2%              |
| measurement           | t_lab_result                    | 14,116,544                | 15,523,576                | 90.9%              |
| visit_occurrence      | t_encounter                     | 1,041,587                 | 1,057,263                 | 98.5%              |
| drug_exposure         |                                 |                           |                           |                    |
| 4,378,657             | t_eprescription Dispensing*     | 2,147,505                 |                           | 84.8%              |
|                       | t_inpatient_med_order†          | 3,015,159                 |                           |                    |

*Refers to outpatient pharmacy orders and inpatient discharge prescriptions
†Refers to medications used during inpatient ward stay
Modified query on the transformed data

- Selection criteria of diabetes, hypertension and depression cohorts

- Start of observation
  - At least 6 months of prior observation
  - No exposure to drugs of interest for 6 months prior to index exposure
  - At least one condition occurrence of disease of interest and no condition occurrence of any excluded disease between all time prior to index and all time after index

- INDEX: First exposure
  - At least 3 months of observation post exposure
  - At least one drug exposure after index
## Modified query on the transformed data

- **Diseases of interest, excluded diseases and drugs used in each cohort**

| Cohort      | Disease of interest | Excluded disease            | Drug classes included                                                                 |
|-------------|---------------------|----------------------------|---------------------------------------------------------------------------------------|
| Diabetes    | Diabetes mellitus   | Findings related to pregnancy | Insulins, biguanides, sulfonylureas, DPP4 inhibitors*, alpha-glucosidase inhibitors, SGLT2 inhibitors^ |
| Hypertension| Hypertensive disorder | Findings related to pregnancy | Antihypertensives, diuretics, peripheral vasodilators, beta blockers, calcium channel blockers, agents acting on the renin-angiotensin-aldosterone system |
| Depression  | Depressive disorder | Findings related to pregnancy, bipolar I disorder, schizophrenia | Antidepressants                                                                      |

*dipeptidyl peptidase 4 inhibitors

^sodium-glucose transport protein 2 inhibitors
Results and Discussion

• Number of patients identified per cohort
  – Diabetes mellitus (n = 1,006)
  – Hypertension (n = 3,175)
  – Depression (n = 251)

• Sequence of drug exposures in these patients were tracked, and plotted on sunburst diagrams
Results and Discussion

Diabetes Mellitus (n=1,006)

- Metformin most often prescribed as the first medication (53.5%).
- Sulfonylureas (SU) were the most common second line agent used in diabetes.
- Among SU, glipizide was the most common (43.3%), followed by tolbutamide (7.1%).
- Newer generation alternatives such as gliclazide and glimepiride were used less frequently, at 5.5 and 0.6%, respectively.

Diabetes Mellitus
Period of Data: Jan 2013 to Dec 2016
Observations

• Fair proportion of patients using insulin as first-line medication
• Patients could have obtained medication from other healthcare providers.
• No observable difference in Hba1c levels between patients given insulin vs patients given other agents as first line medication over 4 readings
• Converting EHR to CDM provides opportunities for such analysis to be carried out efficiently

Error bars showing mean ± 1 s.d. of first 4 HbA1C readings by treatment group
Average duration between readings: 3.2 months
Results and Discussion

Hypertension (n=3,175)

- Amlodipine (22.1%) was the most commonly used first line medication for hypertension.
- Considerable heterogeneity for first-line treatment for hypertension.
- However, for second-line options, most common drugs used were enalapril, losartan, nifedipine and amlodipine.
Observations

• In PNAS treatment pathway study (published in 2015), the most prevalent drugs were hydrochlorothiazide, lisinopril, amlodipine

• In our KTPH study, the most prevalent were amlodipine, losartan, enalapril (very few diuretics)
  – Data were from Jan 2013 to Dec 2016
  – Clinicians likely adopted a newer version of the treatment guidelines i.e. JNC-8
Results and Discussion

**Depression (n=251)**

- Mirtazapine (37.1%) and fluvoxamine (19.1%) were the most commonly used first line medications for depression.

- Overall, a large variety of drug choices across all levels of treatment. Little consensus on prescribing patterns.

**Depression**

Period of Data: Jan 2013 to Dec 2016
Limitations with Sunburst Diagrams

• Dosages and dosing frequency of different treatments not compared

• Unable to identify whether treatment was stopped or switched
Conclusion

• Considerable heterogeneity in treatment patterns for hypertension and depression, whereas for diabetes, metformin was the most common first-line agent (53.5%)

• Use of CDM and federated query\(^1\) were feasible for the data source

• These models provide drug regulators valuable insights on real world drug utilization patterns and adherence to recommended treatment guidelines

1. Hripcsak, G., et al. Observational Health Data Sciences and Informatics (OHDSI): Characterising treatment pathways at scale using the OHDSI network. Proc Natl Acad Sci U.S.A, 2016. 113(27):7329-36.
Any Questions?