



AUSOM database: Common Data Model, Vocabulary

Dukyong Yoon



ECG-ViEW II (<http://ecgview.org/>)

Electrocardiogram Vigilance with Electronic data Warehouse
longitudinal observational QT database for surveillance on proarrhythmic potential of marketed drugs

Home | About ECG-ViEW | Data Request | Downloads | Publications | Contacts

Welcome to the Electrocardiogram Vigilance with Electronic data Warehouse II(ECG-ViEW II), the database comprising numeric parameters (RR interval, PR interval, QRS duration, QT interval, QTc interval, P axis, QRS axis, and T axis) of Electrocardiogram for surveillance of proarrhythmic potential of marketed drugs.

The database contains more than ~0.7 million of QT/QTc measurements and associated clinical data, including diagnoses(~7.7 million), drug prescriptions (~49 million) and selected laboratory-test results (~3.8 million), from ~461 thousands patients over the 19-years study period. The data is intended for researchers interested in surveillance on proarrhythmic potential of marketed drugs.

Scanned ECG printout → cropping → Cropped text area → Background cleansing → Recognized ECG parameters → Text-parsing

Vent. rate: 77 bpm
PR interval: 168 ms
QRS duration: 88 ms
QT/QTc: 424/479 ms
P-R-T axes: 67 32 153

Health examination Outpatient Emergency

Inpatient All

<Overall process to extract ECG parameters from scanned ECG> <The QTc distribution and size of study populations in ECG-ViEW>

좋아요 공유하기 39명이 좋아합니다. 친구들이 무엇을 좋아하는지 알아보려면 가입하기

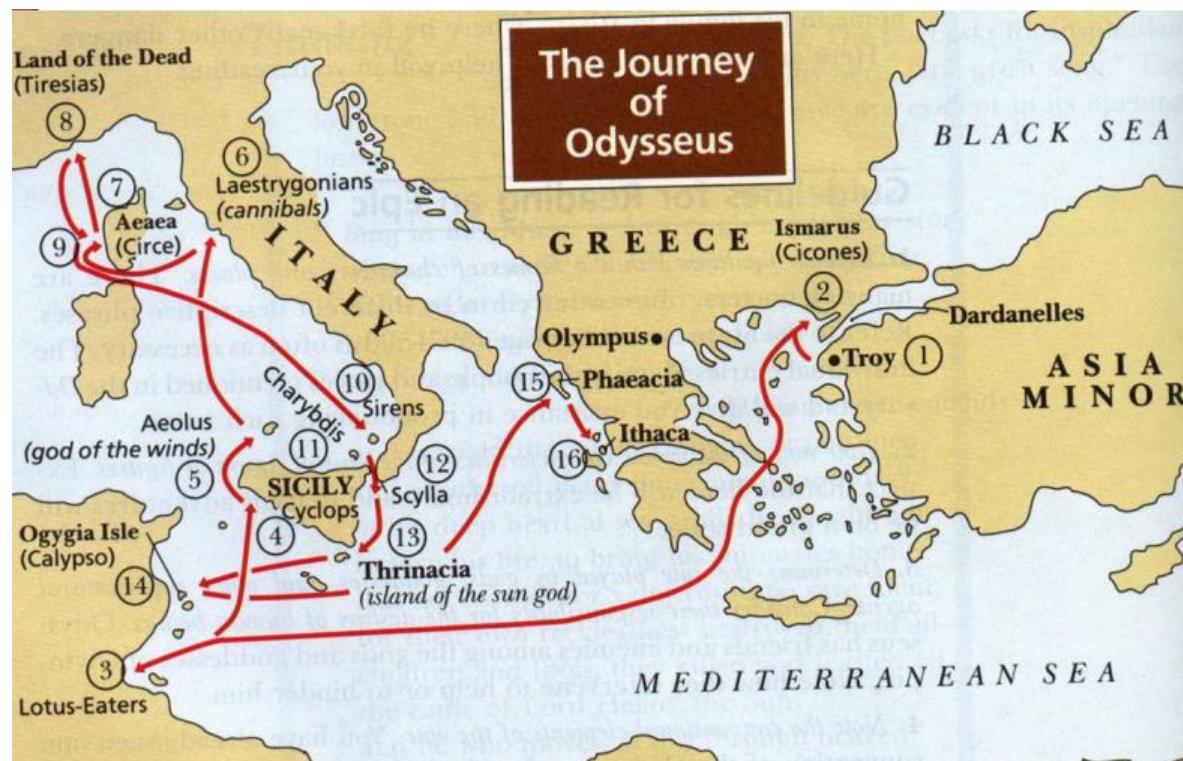
This site is maintained by ABMI

Systems Biomedical Informatics
National Core Research Center



OHDSI (Observational Health Data Sciences and Informatics)

The odyssey to evidence generation





OHDSI (Observational Health Data Sciences and Informatics)

The odyssey to evidence generation





OHDSI (Observational Health Data Sciences and Informatics)

The odyssey to evidence generation

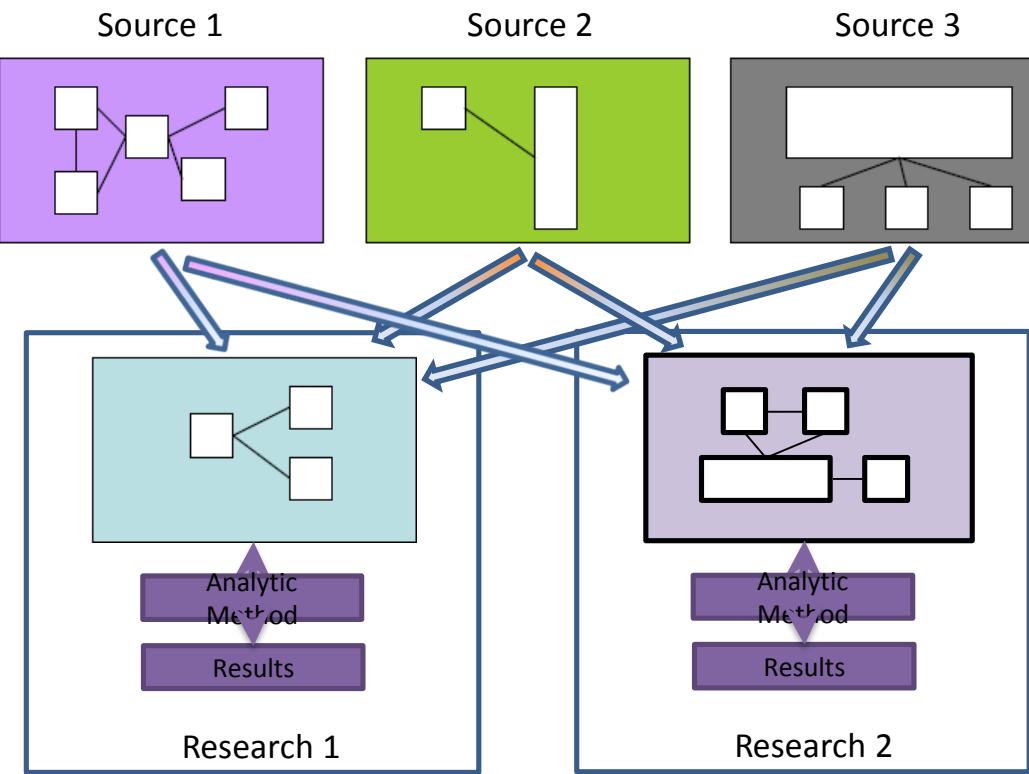




Why Common Data Model?

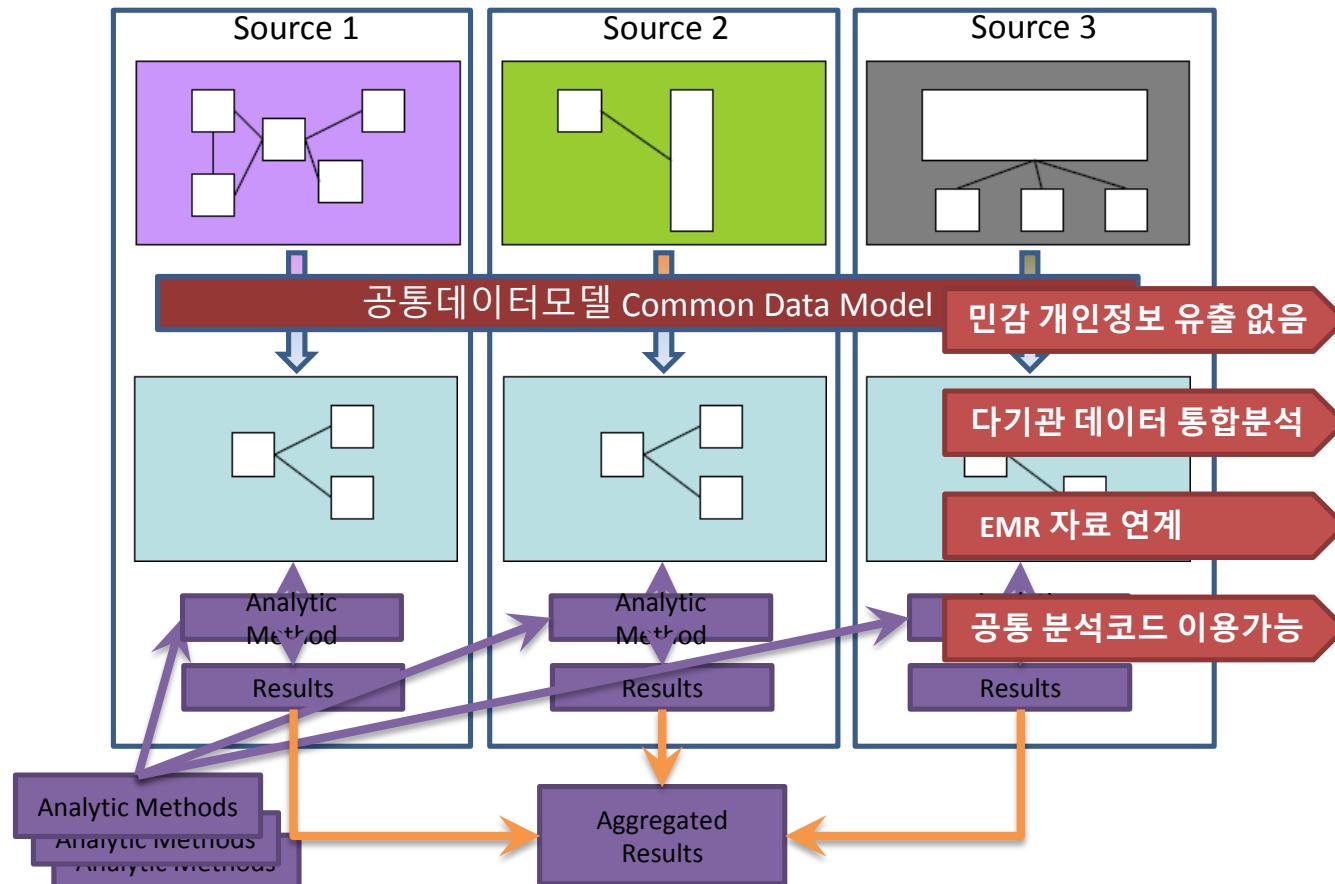
기존의 다기관 연구방법

연구 수행 때마다 데이터 모델을 맞추는 변환 작업을 수행해야 함





CDM in Distributed Research Network





OMOP 공통데이터모델

- OMOP (Observational Medical Outcomes Partnerships) CDM (common data model)
 - 2008년 관찰 데이터를 이용한 연구 기반 구축을 위해 시작된 공공-민간 협력 파트너십
 - OMOP에서 개발한 공통데이터모델은 계속 진화·발전하고 있음



간결한 구조

- 24개의 변환 테이블로 구성

국제 표준 모델

- 전 세계적으로 동일한 구조와 의미
- 높은 재사용성
- 공동 연구 및 비교 연구 용이

표준 용어 이용

- SNOMED-CT, LOINC 등

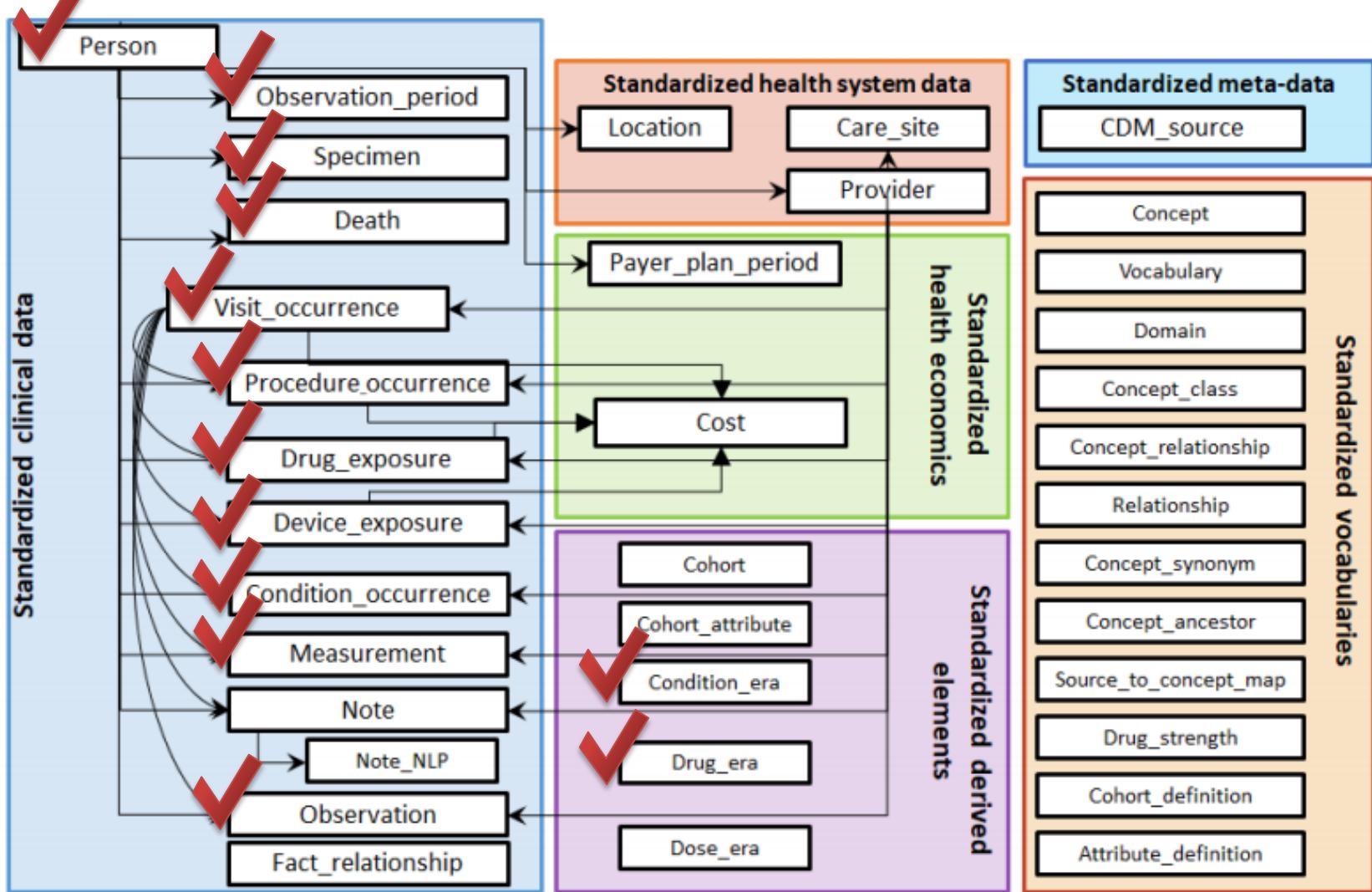
오픈 정책 채택

- 다양한 오픈소스 툴 보유
- 많은 사용자 및 개발자
- 높은 창의성과 혁신성



OMOP-CDM Version 5.3

<https://github.com/OHDSI/CommonDataModel/wiki>





OMOP CDM Principles

- Patient centric
- Vocabulary and Data Model are blended
- Domain-oriented concepts
- Accommodates data from various sources
- Preserves data provenance
- Extendable & Evolving
- Database Platform Independent

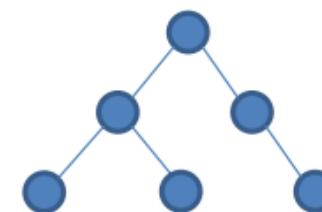


OMOP CDM Standard Domain Features

Feature	Description & Purpose	Field Name Convention	Example
Patient centric	Every domain table has patient identifier . Patient data can be retrieved independently from other domains.	<code>person_id</code>	<code>person_id</code> 123
Unique domain identifiers	Every domain table has a unique primary key to identify domain entities.	<code><entity>_id</code>	<code>condition_occurrence_id</code> 470985
Standard concept from a respective vocabulary domain	Integration with the Vocabulary. Foreign key into the Standard Vocabulary for Standard Concept .	<code><entity>_concept_id</code>	<code>condition_concept_id</code> 313217 (SNOMED "Atrial Fibrillation")
Source value	Provenance. Verbatim information from the source data, not to be used by any standard analytics.	<code><entity>_source_value</code>	<code>condition_source_value</code> 427.31 (ICD9CM "Atrial Fibrillation")
Source concept from a respective vocabulary domain	Provenance. Foreign key into Standard Vocabulary for Source Concept .	<code><entity>_source_concept_id</code>	<code>condition_source_concept_id</code> 44821957 (ICD9CM "Atrial Fibrillation")
Source type	Provenance. Foreign key into Vocabulary for the origin of the data .	<code><entity>_type_concept_id</code>	<code>condition_type_concept_id</code> 38000199 ("Inpatient header – primary")



Structure of OMOP Vocabulary



All content: concepts in
concept

Direct relationships between
concepts in
concept_relationship

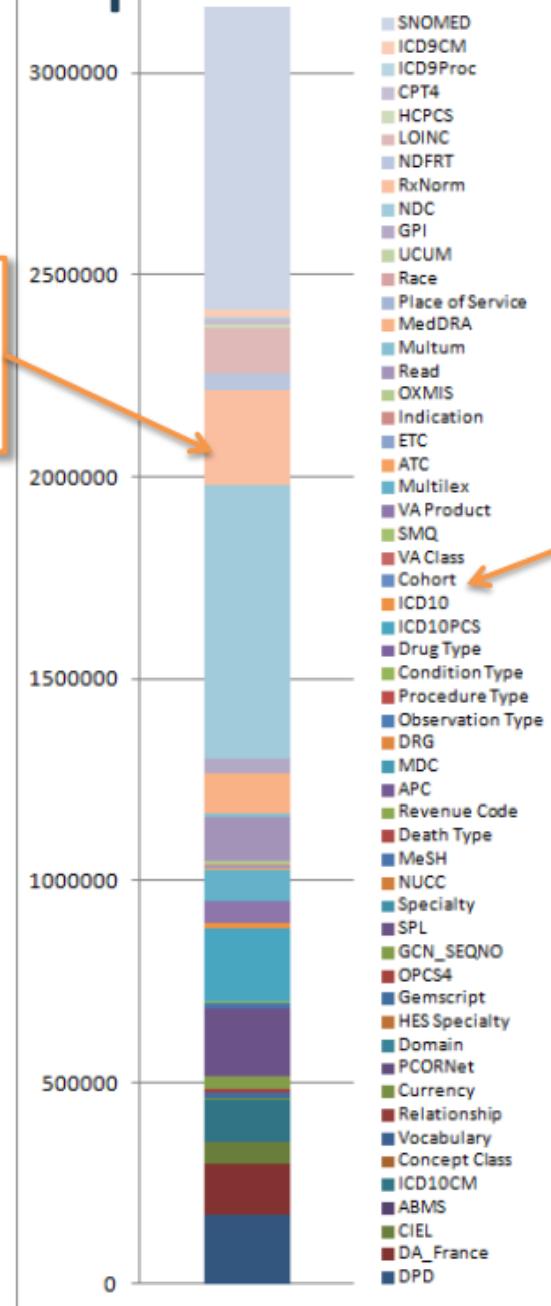
Multi-step hierarchical
relationships pre-processed
into
concept_ancestor



Single Concept Reference Table

All vocabularies
stacked up in one
table

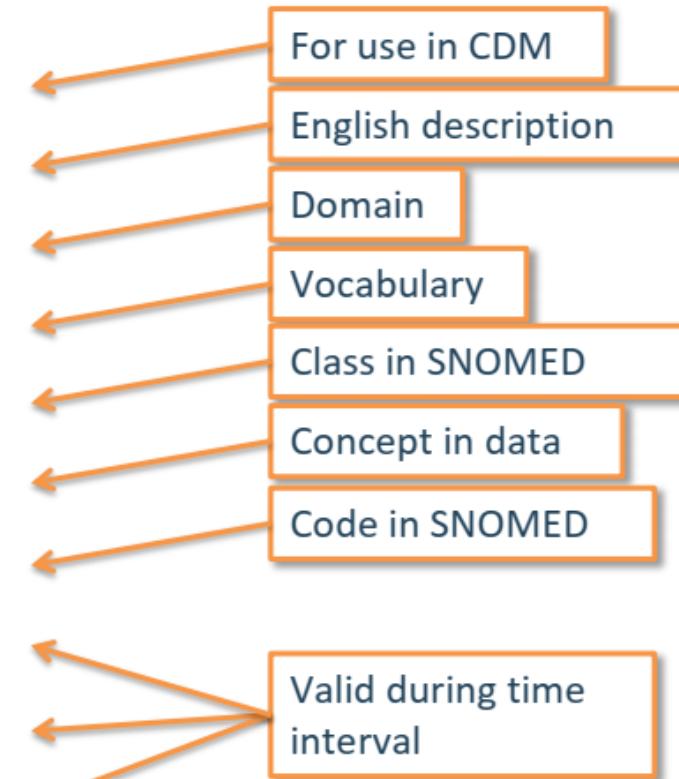
Vocabulary ID





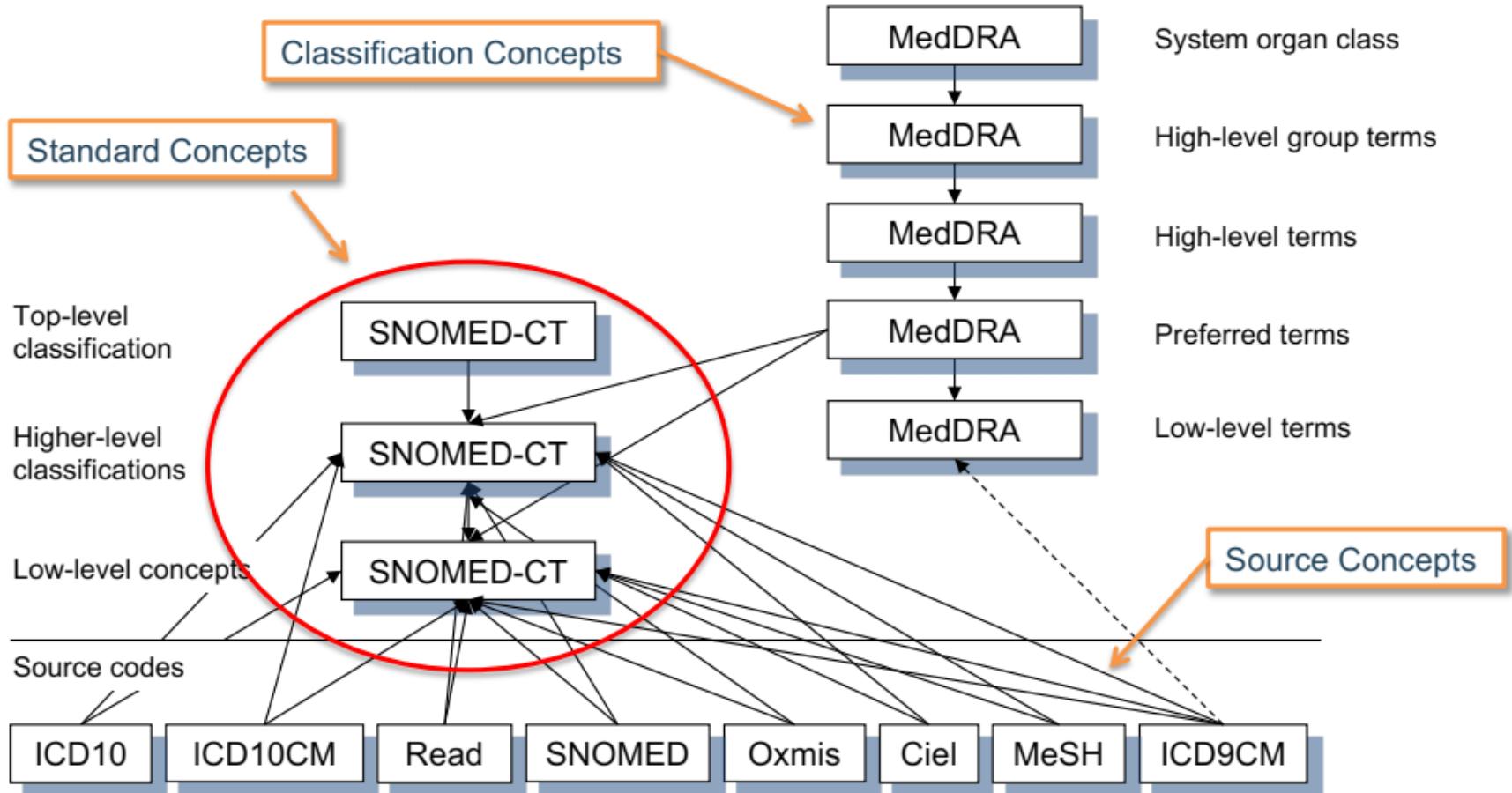
What's in a Concept

CONCEPT_ID	313217
CONCEPT_NAME	Atrial fibrillation
DOMAIN_ID	Condition
VOCABULARY_ID	SNOMED
CONCEPT_CLASS_ID	Clinical Finding
STANDARD_CONCEPT	S
CONCEPT_CODE	49436004
VALID_START_DATE	01-Jan-1970
VALID_END_DATE	31-Dec-2099
INVALID_REASON	





Condition Concepts





Finding the Right Concept #1

1. ..if I know the ID

```
SELECT * FROM concept WHERE concept_id = 313217
```

CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE	VALID_START_DATE	VALID_END_DATE	INVALID_REASON
313217	Atrial fibrillation	Condition	SNOMED	Clinical Finding	S	49436004	01-Jan-1970	31-Dec-2099	

2. ..if I know the code

```
SELECT * FROM concept WHERE concept_code = '49436004'
```

SNOMED code

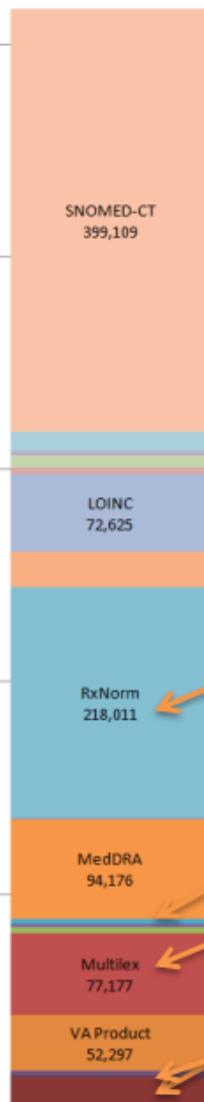
CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE	VALID_START_DATE	VALID_END_DATE	INVALID_REASON
313217	Atrial fibrillation	Condition	SNOMED	Clinical Finding	S	49436004	01-Jan-1970	31-Dec-2099	



Concept ID versus Concept Code

```
SELECT *
FROM concept
WHERE concept_code = '1001';
```

Same code



Concept_Name	Concept Class	Vocabulary_ID	Concept_Code
Antipyrine	Ingredient	RxNorm	1001
Aceprometazine maleate	Ingredient	BDPM	1001
Serum	Specimen	CIEL	1001
methixene hydrochloride	Ingredient	Multilex	1001
Brompheniramine Maleate, 10 mg/mL injectable solution	Multum	Multum	1001
ABBOTT COLD SORE BALM 4%/0.06% W/	Drug Product	LPD_Australia	1001
Residential Treatment - Psychiatric	Revenue Code	Revenue Code	1001



Finding the Right Concept #2

3. ..if I know the name

```
SELECT * FROM concept WHERE concept_name = 'Atrial fibrillation';
```

CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE
313217	Atrial fibrillation	Condition	SNOMED	Clinical Finding	S	49436004
44821957	Atrial fibrillation	Condition	ICD9CM	5-dig billing code		427.31
35204953	Atrial fibrillation	Condition	MedDRA	PT	C	10003658
45500085	Atrial fibrillation	Condition	Read	Read		G573000
45883018	Atrial fibrillation	Meas Value	LOINC	Answer	S	LA17084-7



Finding the Right Concept #3

- if don't know any of this, but I know the code in another vocabulary

ICD-9 is not a Standard Concept

```
SELECT * FROM concept WHERE concept_code = '427.31';
```

CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE
44821957	Atrial fibrillation	Condition	ICD9CM	5-dig billing code		427.31

```
SELECT * FROM concept_relationship WHERE concept_id_1 = 44821957;
```

Mapping to different vocabularies

Kind of relationship

ID_1	CONCEPT_ID_2	RELATIONSHIP_ID	VALID_START_DATE	VALID_END_DATE	INVALID_REASON
44821957	21001551	ICD9CM - FDB Ind	01-Oct-13	31-Dec-2099	
44821957	35204953	ICD9CM - MedDRA	01-Jan-70	31-Dec-2099	
44821957	44824248	Is a	01-Oct-14	31-Dec-2099	
44821957	44834731	Is a	01-Oct-14	31-Dec-2099	
44821957	313217	Maps to	01-Jan-70	31-Dec-2099	



Mapping = Translating

Step 1. Lookup the Source Concept

```
SELECT * FROM concept WHERE concept_code = '427.31';
```

CONCEPT_ID	CONCEPT_NAME	DOMAIN_ID	VOCABULARY_ID	CONCEPT_CLASS_ID	STANDARD_CONCEPT	CONCEPT_CODE
44821957	Atrial fibrillation	Condition	ICD9CM	5-dig billing code		427.31

I

Langenscheidt
Universal-Wörterbuch

Englisch

Englisch – Deutsch
Deutsch – Englisch

Step 2. Translate to Standard

```
SELECT * FROM concept_relationship WHERE concept_id_1 = 44821957  
AND relationship_id = 'Maps to';
```

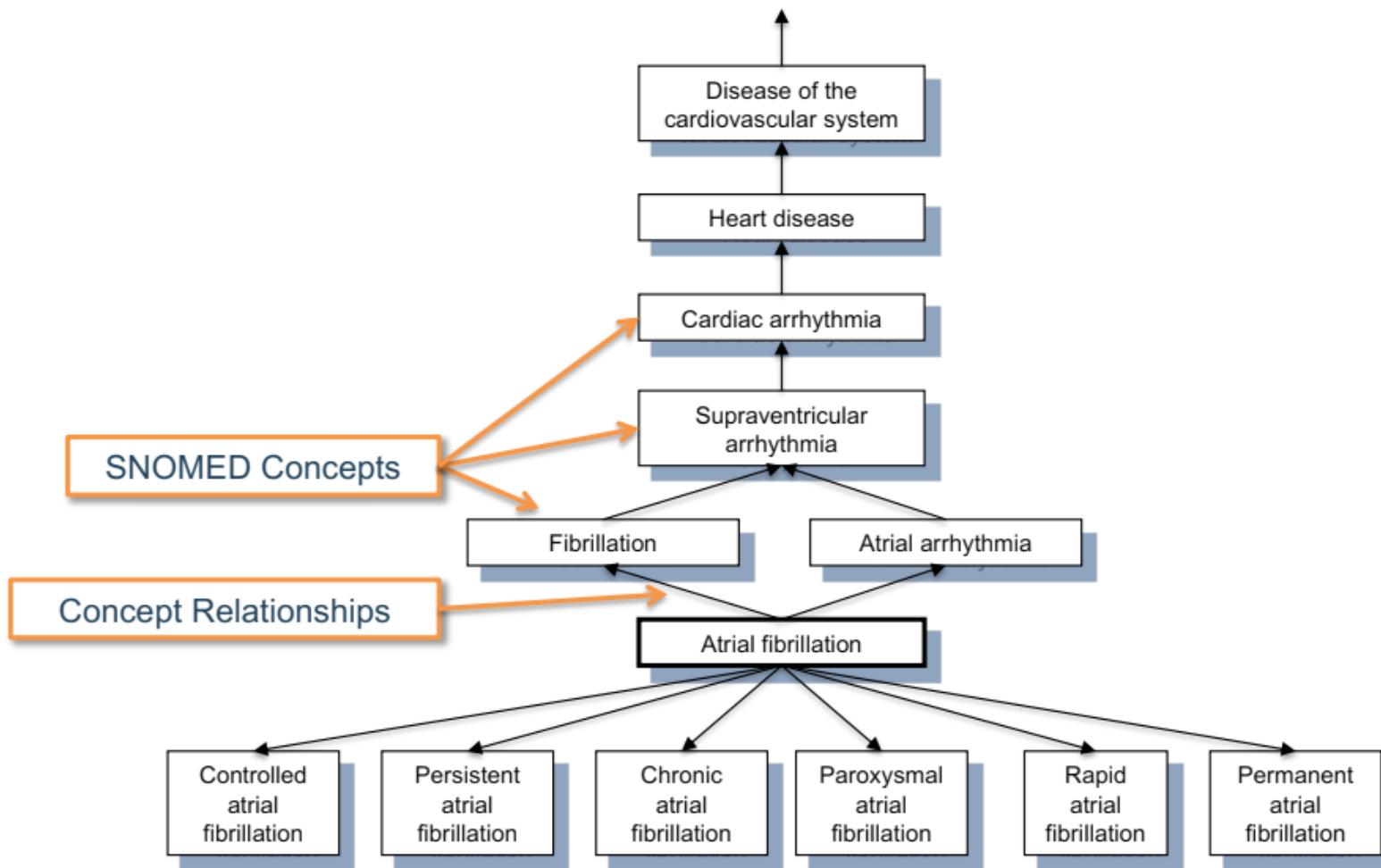
CONCEPT_ID_1	CONCEPT_ID_2	RELATIONSHIP_ID	VALID_START_DATE	VALID_END_DATE	INVALID_REASON
44821957	313217	Maps to	01-Jan-1970	31-Dec-2099	

Step 3. Check out the translated Concept

```
SELECT * FROM concept WHERE concept_id = 313217;
```



Reason #2: Disease Hierarchy





Exploring Relationships

```
SELECT *
FROM concept_relationship
WHERE concept_id_1 = 313217
```

Related Concepts

Relationship ID

CONCEPT_ID_1	CONCEPT_ID_2	RELATIONSHIP_ID
313217	4232697	Subsumes
313217	4181800	Focus of
313217	35204953	SNOMED - MedDRA eq
313217	4203375	Asso finding of
313217	4141360	Subsumes
313217	4119601	Subsumes
313217	4117112	Subsumes
313217	4232691	Subsumes
313217	4139517	Due to of
313217	4194288	Asso finding of
313217	44782442	Subsumes
313217	44783731	Focus of
313217	21003018	SNOMED - ind/CI
313217	40248987	SNOMED - ind/CI
313217	21001551	SNOMED - ind/CI
313217	21001540	SNOMED - ind/CI
313217	45576876	Mapped from
313217	44807374	Asso finding of
313217	21013834	SNOMED - ind/CI
313217	21001572	SNOMED - ind/CI
313217	21001606	SNOMED - ind/CI
313217	21003176	SNOMED - ind/CI
313217	42263991	is a
313217	500001801	SNOMED - HOI
313217	500002401	SNOMED - HOI
313217	4119602	Subsumes
313217	40631039	Subsumes
313217	4108832	Subsumes
313217	21013671	SNOMED - ind/CI
313217	21013390	SNOMED - ind/CI
313217	313217	Maps to
313217	44821957	Mapped from
313217	2617597	Mapped from
313217	45500085	Mapped from
313217	313217	Mapped from
313217	45951191	Mapped from
313217	21013856	SNOMED - ind/CI
313217	21001575	SNOMED - ind/CI
313217	21001594	SNOMED - ind/CI



Exploring Relationships

```
SELECT cr.relationship_id, c.*  
FROM concept_relationship cr  
JOIN concept c ON cr.concept_id_2 = c.concept_id  
WHERE cr.concept_id_1 = 313217
```

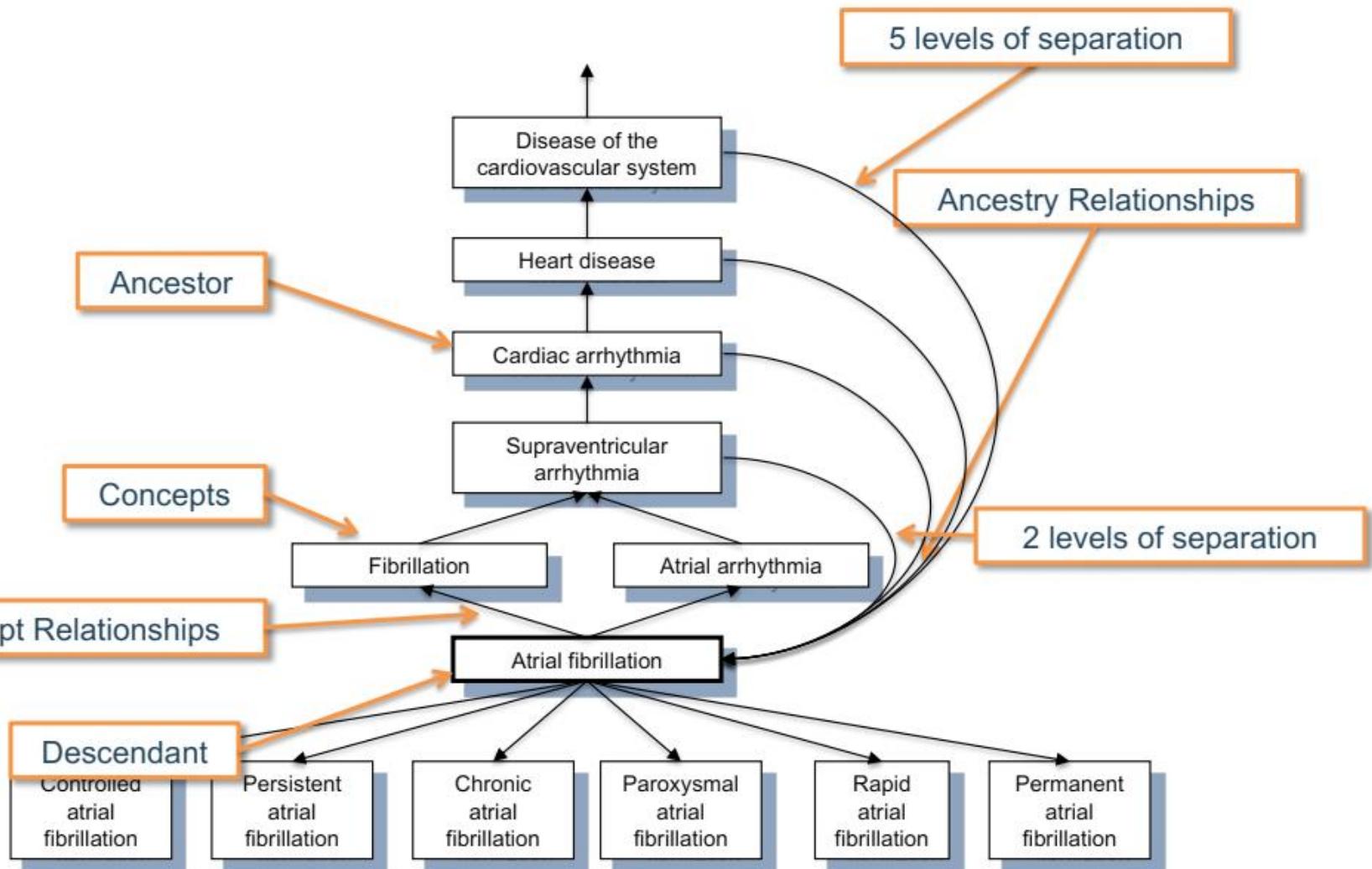
Find out related concept

relationship_id	concept_id	concept_name	domain_id	vocabulary_id	concept_class_id	standard_concept	concept_code	valid_start_date	valid_end_date	invalid_reason
Anso finding of	4194288		Observation	SNOMED	Context-dependent S		312442005	1/1/1970 0:00	12/31/2099 0:00	NULL
Anso finding of	4203375		Observation	SNOMED	Context-dependent S		433276002	1/31/2009 0:00	12/31/2099 0:00	NULL
Anso finding of	42689685	Atrial fibrillation not otherwise specified	Observation	SNOMED	Context-dependent S	1.06706E+15	4/1/2017 0:00	12/31/2099 0:00	NULL	
Anso finding of	44807374	Atrial fibrillation - excluded	Observation	SNOMED	Context-dependent S	8.16401E+14	4/1/2014 0:00	12/31/2099 0:00	NULL	
Concept poss_eq from	40323929	Fibrillation - atrial	Condition	SNOMED	Clinical Finding	NULL	155364009	1/1/1970 0:00	3/11/2016 0:00 U	
Concept poss_eq from	40345197	Fibrillation - atrial	Condition	SNOMED	Clinical Finding	NULL	266306001	1/1/1970 0:00	3/11/2016 0:00 U	
Due to of	4139517	Transient cerebral ischemia due to atrial fibrillation	Condition	SNOMED	Clinical Finding	S	426814001	1/1/1970 0:00	12/31/2099 0:00	NULL
Focus of	4209991	Insertion of pacemaker for control of atrial fibrillation	Procedure	SNOMED	Procedure	S	449863006	1/31/2012 0:00	12/31/2099 0:00	NULL
Has finding site	4242112	Atrial structure	Spec Anatomic Site	SNOMED	Body Structure	S	59652004	1/1/1970 0:00	12/31/2099 0:00	NULL
Is a	4226399	Fibrillation	Condition	SNOMED	Clinical Finding	S	40593004	1/1/1970 0:00	12/31/2099 0:00	NULL
Is a	4068155	Atrial arrhythmia	Condition	SNOMED	Clinical Finding	S	17366009	1/1/1970 0:00	12/31/2099 0:00	NULL
Mapped from	40323929	Fibrillation - atrial	Condition	SNOMED	Clinical Finding	NULL	155364009	1/1/1970 0:00	3/11/2016 0:00 U	
Mapped from	2617597	Patient with heart failure and atrial fibrillation documented to be on warfarin therapy	Observation	HCPCS	HCPCS	NULL	G8183	1/1/1970 0:00	11/11/2014 0:00 D	
Mapped from	45576876	Unspecified atrial fibrillation	Condition	ICD10CM	5-char billing code	NULL	I48.91	12/30/2006 0:00	12/31/2099 0:00	NULL
Mapped from	45500085	Atrial fibrillation	Condition	Read	Read	NULL	G573000	1/1/1970 0:00	12/31/2099 0:00	NULL
Mapped from	45611600	Atrial Fibrillation	Condition	MeSH	Main Heading	NULL	D001281	1/1/1970 0:00	12/31/2099 0:00	NULL
Mapped from	40345197	Fibrillation - atrial	Condition	SNOMED	Clinical Finding	NULL	266306001	1/1/1970 0:00	3/11/2016 0:00 U	
Mapped from	45951191	Atrial Fibrillation	Condition	CIEL	Diagnosis	NULL	148203	11/3/2007 0:00	12/31/2099 0:00	NULL
Mapped from	313217	Atrial fibrillation	Condition	SNOMED	Clinical Finding	S	49436004	1/1/1970 0:00	12/31/2099 0:00	NULL
Mapped from	44821957	Atrial fibrillation	Condition	ICD9CM	5-dig billing code	NULL	427.31	1/1/1970 0:00	12/31/2099 0:00	NULL
Maps to	313217	Atrial fibrillation	Condition	SNOMED	Clinical Finding	S	49436004	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - HOI	500002401	OMOP Atrial Fibrillation 1	Condition	Cohort	Cohort	C	500002401	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - HOI	500001801	OMOP Qt Prolongation/Torsade De Pointes 1	Condition	Cohort	Cohort	C	500001801	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - Ind/CI	21005673	Prevention of Thromboembolism in Chronic Atrial Fibrillation	Drug	Indication	Indication	C	5673	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - Ind/CI	21003176	Tachyarrhythmia	Drug	Indication	Indication	C	3176	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - Ind/CI	21001542	Supraventricular Tachycardia	Drug	Indication	Indication	C	1542	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - Ind/CI	21001594	Disease of Cardiovascular System	Drug	Indication	Indication	C	1594	1/1/1970 0:00	12/31/2099 0:00	NULL
SNOMED - MedDRA eq	35204953	Atrial fibrillation	Condition	MedDRA	PT	C	10003658	1/1/1970 0:00	12/31/2099 0:00	NULL
Subsumes	4117112	Controlled atrial fibrillation	Condition	SNOMED	Clinical Finding	S	300996004	1/1/1970 0:00	12/31/2099 0:00	NULL
Subsumes	4119601	Lone atrial fibrillation	Condition	SNOMED	Clinical Finding	S	233910005	1/1/1970 0:00	12/31/2099 0:00	NULL
Subsumes	4232697	Persistent atrial fibrillation	Condition	SNOMED	Clinical Finding	S	440059007	1/31/2009 0:00	12/31/2099 0:00	NULL
Subsumes	4141360	Chronic atrial fibrillation	Condition	SNOMED	Clinical Finding	S	426749004	1/1/1970 0:00	12/31/2099 0:00	NULL
Subsumes	44782442	Atrial fibrillation with rapid ventricular response	Condition	SNOMED	Clinical Finding	S	1.20041E+14	1/31/2014 0:00	12/31/2099 0:00	NULL
Subsumes	4199501	Rapid atrial fibrillation	Condition	SNOMED	Clinical Finding	S	314208002	1/1/1970 0:00	12/31/2099 0:00	NULL
Subsumes	4119602	Non-rheumatic atrial fibrillation	Condition	SNOMED	Clinical Finding	S	233911009	1/1/1970 0:00	12/31/2099 0:00	NULL

Descendant concepts



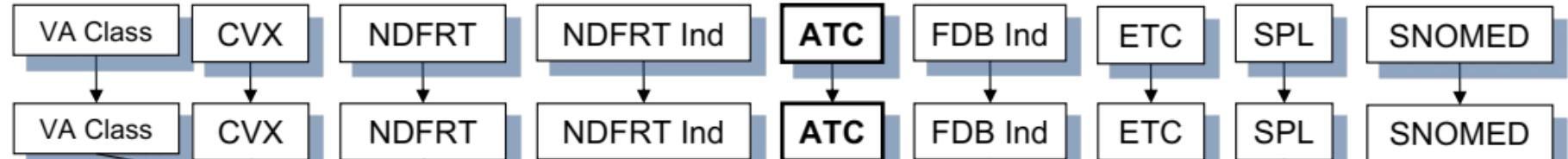
Ancestry Relationships: Higher-Level Relationships





Drug Hierarchy

Classifications



Drugs

Ingredients

Standard Concepts
Drug forms and components

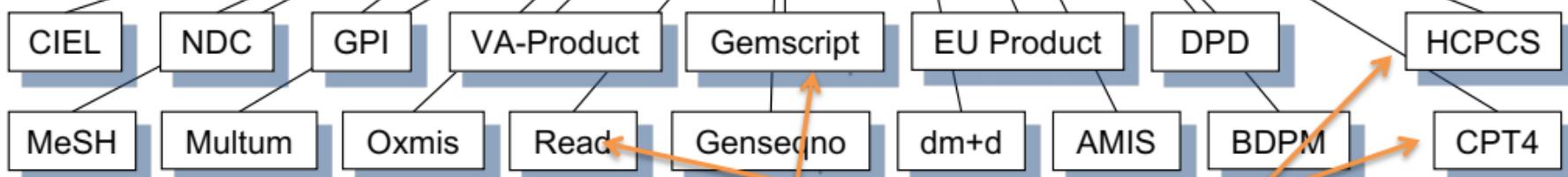
RxNorm RxNorm Extension

RxNorm RxNorm Extension

Drug products

RxNorm RxNorm Extension

Source codes



Source Codes

Procedure Drugs



PERSON

필드명	필수사항	데이터타입	설명 (한글)
person_id	Yes	integer	개별 환자별로 부여되는 개인 식별자
gender_concept_id	Yes	integer	(외래키) 성별에 대한 OMOP concept ID
year_of_birth	Yes	integer	출생년도. 데이터 소스에 따라 생년월일 데이터에서 년도만 추출되어질 수 있음
month_of_birth	No	integer	출생월
day_of_birth	No	integer	출생일
time_of_birth	No	Time	출생 시간
race_concept_id	Yes	integer	(외래키) 인종에 대한 concept ID
ethnicity_concept_id	Yes	integer	(외래키) 민족성에 대한 concept ID
location_id	No	integer	(외래키) 주소에 대한 자세한 정보가 담겨 있는 location table에 있는 환자의 거주지역을 참조하기 위한 키



PERSON

필드명	필수사항	데이터타입	설명 (한글)
provider_id	No	integer	(외래키) Provider 테이블에 있는 일차적인 치료를 제공하고 있는 의료진을 참조하기 위한 키
care_site_id	No	integer	(외래키) 치료 장소(care site)에 대한 자세한 정보가 담겨 있는 care_site table에 있는 환자의 일차적인 치료 장소를 참조하기 위한 키
person_source_value	No	varchar(50)	원천 데이터에서 사용되는 (암호화된) 환자 식별 키. 원천 데이터와 연계할 필요가 있을 경우 이 필드의 데이터가 사용될 수 있음
gender_source_value	No	varchar(50)	원천 데이터 상에서 보여지는 환자의 성별에 대한 코드 값. 원천 데이터의 환자의 성별은 표준 사전 내에 정의된 성별에 대한 컨셉과 매핑되며, 원천 데이터의 값이 참고용으로 이 필드에 저장됨
gender_source_concept_id	No	integer	(외래 키) 원천 데이터에서 사용되는 성별 구분 코드 체계에 대한 컨셉을 참조함
race_source_value	No	varchar(50)	원천 데이터 상에서 보여지는 환자의 인종(유전적 인종, race)에 대한 코드 값. 원천 데이터의 환자의 인종은 표준 사전 내에 정의된 인종에 대한 컨셉과 매핑되며, 원천 데이터의 값이 참고용으로 이 필드에 저장됨
race_source_concept_id	No	integer	(외래 키) 원천 데이터에서 사용되는 인종 구분 코드 체계에 대한 컨셉을 참조함
ethnicity_source_value	No	varchar(50)	원천 데이터 상에서 보여지는 환자의 민족성(ethnicity)에 대한 코드 값
ethnicity_source_concept_id	No	integer	(외래 키) 원천 데이터에서 사용되는 민족성 구분 코드 체계에 대한 컨셉을 참조함



DEATH

필드명	필수사항	데이터타입	설명 (한글)
person_id	Yes	integer (8)	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
death_date	Yes	date	환자가 사망한 일자
death_type_concept_id	Yes	integer (8)	(외래키) 사망 선고 타입에 대한 컨셉. 원천 데이터에서 사용되는 진단 종류 구분 코드 체계에 대한 컨셉을 참조함
cause_concept_id	No	Integer (8)	(외래키) 사망 원인. 원천 데이터에서 사용되는 진단 구분 코드 체계에 대한 컨셉을 참조함
cause_source_value	No	varchar(50)	사망 원인 원천 데이터
cause_source_concept_id	No	Integer (8)	(외래키) 사망 원인 원천 데이터에서 사용하고 있는 코드 체계 컨셉



(NOTE)

필드명	필수사항	데이터타입	설명 (한글)
note_id	Yes	integer	각 note에 대한 유일한 식별 키
person_id	Yes	Integer	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
note_date	Yes	date	note가 기록된 날짜
note_datetime	No	datetime	note가 기록된 날짜와 시간
note_type_concept_id	Yes	Integer	(외래키) note의 종류, 기원 등에 대해 사전에 정의한 표준용어로 입력한 외래키
note_class_cocnept_id	Yes	Integer	(외래키) HL7 LOINC 문서 타입에 대해 사전에 정의된 표준용어로 입력한 외래키
note_title	No	varchar (250)	note의 제목
note_text	Yes	varchar (MAX)	note의 내용
encoding_concept_id	Yes	varchar(50)	(외래키) note의 character의 인코딩을 사전에 정의한 표준용어로 입력한 외래키



(NOTE)

필드명	필수사항	데이터타입	설명 (한글)
language_concept_id	Yes	Integer	note의 언어에 대해 사전에 정의한 표준용어로 입력한 외래키
provider_id	No	integer	(외래키) 해당 상태를 결정 한 (진단을 내린) 의료 제공자 (provider 테이블 내)를 참조하기 위한 외래키
visit_occurrence_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문(visit 테이블)을 참조하기 위한 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
note_source_value	No	varchar (50)	원천 note의 출처



OBSERVATION PERIOD

필드명	필수사항	데이터타입	설명 (한글)
observation_period_id	Yes	integer	개별 관찰기간을 식별하기 위한 고유 식별키
person_id	Yes	integer	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
observation_period_start_date	Yes	date	관찰 기간의 시작 일자
observation_period_end_date	Yes	date	관찰 기간의 종료 일자
period_type_concept_id	Yes	Integer (8)	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 관찰 기간 정보 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키



VISIT OCCURRENCE

필드명	필수사항	데이터타입	설명 (한글)
visit_occurrence_id	Yes	integer	각 환자의 의료제공 기관 방문에 대한 고유한 식별자
person_id	Yes	integer	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
visit_concept_id	Yes	integer	(외래키) 환자 방문과 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
visit_start_date	Yes	date	내원 시작 날짜
visit_start_time	No	time	내원 시작 시간
visit_end_date	Yes	date	내원 종료 일자. 만약 하루만 방문한 경우 내원 시작 날짜와 동일함
visit_end_time	No	time	내원 종료 시간



VISIT OCCURRENCE

필드명	필수사항	데이터타입	설명(한글)
visit_type_concept_id	Yes	integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 방문 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
provider_id	No	integer	(외래키) 의료서비스 제공자로, provider 테이블에 대한 외래키
care_site_id	No	integer	(외래키) 요양기관 id. Care_site 테이블에 대한 외래키
visit_source_value	No	varchar(50)	visit_occurrence_id에 대한 원천 데이터
visit_source_concept_id	No	Integer	(외래키) visit_occurrence_id에 대한 원천 데이터의 코드 체계 concept
admitted_from_source_value	No	varchar(50)	입원 경로에 대한 원천 데이터
admitted_from_concept_id	Yes	Integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 입원 경로에 대한 서비스 장소가 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
discharge_to_source_value	No	varchar(50)	퇴원에 대한 원천 데이터
discharge_to_concept_id	Yes	Integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 퇴원에 대한 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
preceding_visit_detail_id	No	Integer	(외래키) 해당 방문 직전 방문에 대한 외래 키



(VISIT DETAIL)

필드명	필수사항	데이터타입	설명 (한글)
visit_detail_id	Yes	integer	각 환자의 의료제공 기관의 부서 방문에 대한 교유한 식별자 (외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키.
person_id	Yes	integer	관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
visit_detail_concept_id	Yes	integer	(외래키) 환자가 방문한 기관의 부서와 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
visit_detail_start_date	No	date	내원 시작 날짜
visit_detail_start_datetime	Yes	datetime	내원 시작 시간
visit_detail_end_date	No	date	내원 종료 일자. 만약 하루만 방문한 경우 내원 시작 날짜와 동일함
visit_detail_end_datetime	Yes	datetime	내원 종료 시간
visit_detail_type_concept_id	Yes	Integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 방문 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
provider_id	No	integer	(외래키) 의료서비스 제공자로, provider 테이블에 대한 외래키
care_site_id	No	integer	(외래키) 요양기관 id. Care_site 테이블에 대한 외래키



(VISIT DETAIL)

필드명	필수사항	데이터타입	설명 (한글)
visit_detail_source_value	No	string(50)	visit_detail_id에 대한 원천 데이터
visit_detail_source_concept_id	Yes	Integer	(외래키) visit_detail_id에 대한 원천 데이터의 코드 체계 concept
admitted_from_source_value	No	Varchar(50)	입원 경로에 대한 원천 데이터
admitted_from_concept_id	Yes	Integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 입원 경로에 대한 서비스 장소가 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
discharge_to_source_value	No	Varchar(50)	퇴원에 대한 원천 데이터
discharge_to_concept_id	Yes	Integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 퇴원에 대한 서비스 장소가 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
preceding_visit_detail_id	No	Integer	(외래키) 해당 방문 직전 방문에 대한 외래 키
visit_detail_parent_id	No	Integer	visit_detail_id 에 대한
visit_occurrence_id	Yes	Integer	(외래키) visit_detail_id가 발생한 visit_occurrence_id



CONDITION OCCURRENCE

필드명	필수사항	데이터타입	설명 (한글)
condition_occurrence_id	Yes	integer	각 상태(진단명 등) 발생 건에 대한 고유한 식별자 (시스템에 의해 자동 생성)
person_id	Yes	integer	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음
condition_concept_id	Yes	integer	(외래키) 상태와 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
condition_start_date	Yes	date	상태가 기록된 날짜
condition_end_date	No	date	해당 상태가 종료되었다고 판단되는 날짜. 만약 해당 정보가 없을 경우, NULL로 설정.
condition_type_concept_id	Yes	integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 상태 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키. 예를 들어, 주진단 혹은 부진단, 문제 목록 등



CONDITION OCCURRENCE

필드명	필수사항	데이터타입	설명 (한글)
stop_reason	No	varchar(20)	환자 상태에 대한 기록이 더 이상 기록되지 않은 사유를 원천 데이터에 있는 대로 기록(기록 가능한 경우에 한함). 퇴원, 혹은 호전 등이 가능. 주의: 종료 사유는 상태가 더 이상 발생하지 않는다는 것을 의미하지는 않는다
provider_id	No	integer	(외래키) 해당 상태를 결정 한 (진단을 내린) 의료 제공자 (provider 테이블 내)를 참조하기 위한 외래키
visit_occurrence_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문(visit 테이블)을 참조하기 위한 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
condition_source_value	No	varchar(50)	원천 데이터내에 표현되어 있는 진단에 대한 코드값. 원천 데이터의 코드는 표준 사전의 concept와 매핑이 되며, 원천 데이터의 코드가 이 필드에 저장됨. 상태는 일반적으로 다음의 코드 체계에 의해 기록됨: ICD-9-CM diagnosis codes(청구 데이터) 혹은 discharge status/visit diagnosis codes(전자의무기록)
condition_source_concept_id	No	integer	(외래키) 원천 데이터에서 사용되고 있는 코드 체계를 지칭하는 컨셉에 대한 외래 키



DRUG EXPOSURE

필드명	필수사항	데이터타입	설명 (한글)
drug_exposure_id	Yes	Integer (10)	각 약물 사용건에 대한 고유한 식별자
person_id	Yes	Integer (8)	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
drug_concept_id	Yes	Integer (8)	(외래키) 약물과 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
drug_exposure_start_date	Yes	date	현재 약물 사용 사례의 시작 날짜. 다음의 데이터들이 사용될 수 있음: 처방의 시작일, 처방전이 작성된 날짜, 혹은 약물 투여가 기록된 날짜
drug_exposure_end_date	No	date	현재 약물 사용 사례의 종료 날짜
drug_type_concept_id	Yes	Integer (8)	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 약물 노출 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키. 이 필드는 약물 노출이 어떻게 원천 데이터에 기록되었는지를 알려준다. 예: 약물 투여 기록 등.
stop_reason	No	varchar(20)	약물을 중단하게 된 사유 (가능한 경우만 사용). 예: 치료 요법 완료, 약물 변경, 제거 등
refills	No	integer (3)	첫 처방 이후 추가 수령이 가능한 횟수. 약제의 첫 불출은 포함되지 않으며, 0으로 표시. (주의: 우리나라 의료시스템에서는 이 필드에 대한 개념이 없음)
quantity	No	float (4)	원천 처방전 혹은 지급 기록에 기록되어 있는 약의 양



DRUG EXPOSURE

필드명	필수사항	데이터타입	설명(한글)
days_supply	No	integer (4)	원천 처방전 혹은 지급 기록에 기록되어 있는 약의 1일 투여 수
sig	No	varchar (MAX)	원천 처방전 (그리고 약 통에 기록되어 있는) 혹은 지급 기록에 기록되어 있는 약물 복용 방법
route_concept_id	No	Integer (8)	(외래키) 약물 투여 경로를 의미하는 식별자. 표준 사전 내 컨셉을 참조하는 외래키
effective_drug_dose	No	float (8,2)	약물 투여 용량에 대한 수치값
dose_unit_concept_id	No	integer (8)	(외래키) 약물 투여 용량 단위를 나타내는 식별자. 표준 사전 내 컨셉을 참조하는 외래키
lot_number	No	varchar(50)	제품이 어디서 유래되었는지에 대한 식별자
provider_id	No	integer (8)	(외래키) 해당 약물 노출을 시작(처방) 한 의료 제공자 (provider 테이블 내)를 참조하기 위한 외래키
visit_occurrence_id	No	integer (8)	(외래키) 해당 약물 노출이 시작된 방문(visit 테이블)을 참조하기 위한 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
drug_source_value	No	varchar(50)	원천 데이터내에 표현되어 있는 약물에 대한 코드값



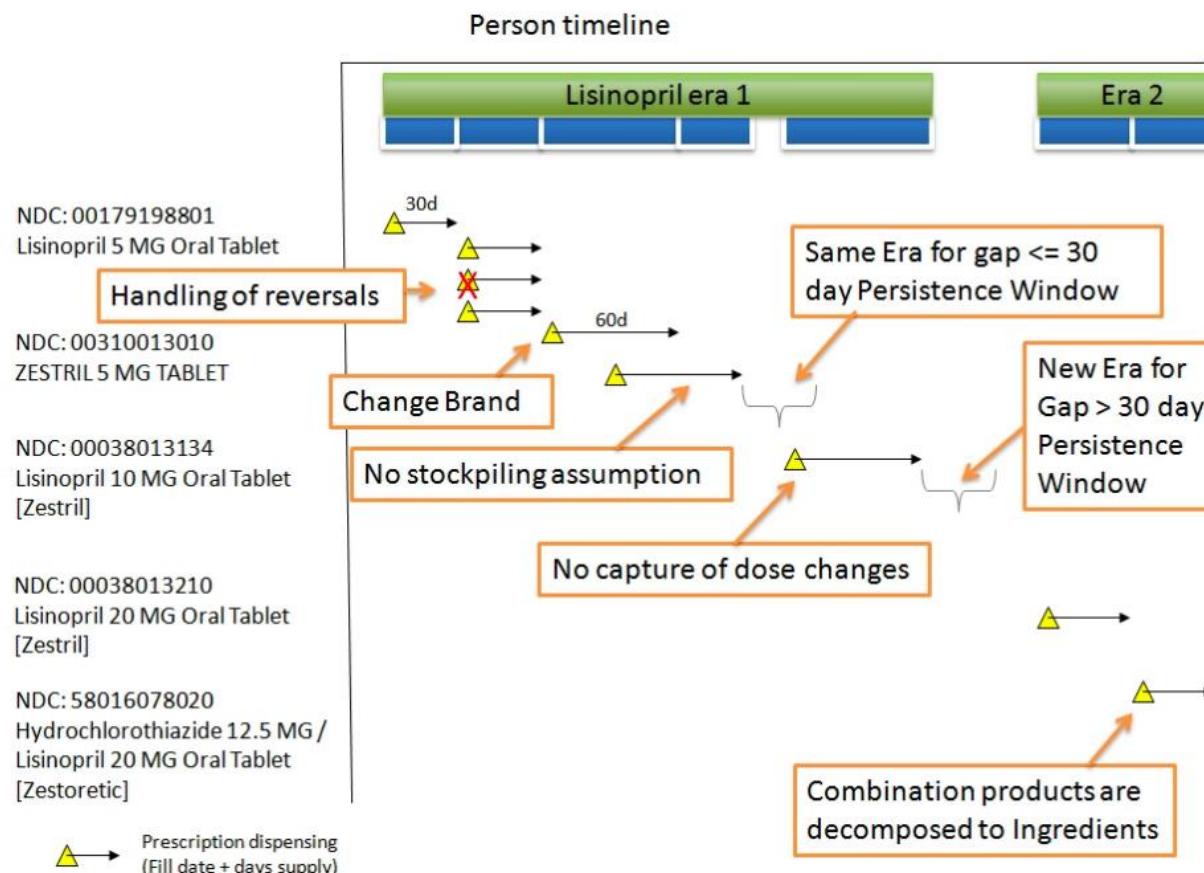
DRUG EXPOSURE

필드명	필수사항	데이터타입	설명 (한글)
drug_source_concept_id	No	Integer (8) (외래키)	원천 데이터에서 사용되고 있는 코드 체계를 지칭하는 컨셉에 대한 외래 키
route_source_value	No	varchar(50)	원천 데이터 내 기록되어 있는 약물 투여 경로 상세 내역에 대한 정보
dose_unit_source_value	No	varchar(50)	원천 데이터 내 기록되어 있는 용량 단위에 대한 정보



DRUG ERA

DRUG_ERA





DRUG ERA

필드명	필수사항	데이터타입	설명 (한글)
drug_era_id	Yes	integer	연속적 약물 사용 에피소드에 대한 고유한 식별자
person_id	Yes	integer	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음.
drug_concept_id	Yes	integer	(외래키) 약물과 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
drug_era_start_datetime	Yes	datetime	현재 약물 사용 사례의 시작 날짜. 다음의 데이터들이 사용될 수 있음: 처방의 시작일, 처방전이 작성된 날짜, 혹은 약물 투여가 기록된 날짜
drug_era_end_datetime	Yes	datetime	현재 약물 사용 사례의 종료 날짜
drug_exposure_count	No	integer	하나의 drug_era_id로 간주된 약물 노출 횟수
gap_days	No	integer	하나의 연속적인 약물 사용 에피소드로 간주된 기간과 (drug_era_start_date and drug_era_end_date), 참고된 여러 약물 처방 기록의 날짜와의 차이



PROCEDURE OCCURRENCE

필드명	필수사항	데이터타입	설명 (한글)
procedure_occurrence_id	Yes	integer (8)	각 시술 발생에 대한 고유한 식별자
person_id	Yes	integer (8)	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음
procedure_concept_id	Yes	integer (8)	(외래키) 시술과 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
procedure_date	Yes	date	시술이 시행된 날짜
procedure_type_concept_id	Yes	Integer (8)	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 시술 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
modifier_concept_id	No	Integer (8)	(외래키) 시술에 대한 표현을 한정하기 위한 표준 컨셉 identifiedr를 참조하는 외래키 (예. 양측)
quantity	No	integer (4)	시술이 처방되거나 수행된 양
provider_id	No	integer (8)	(외래키) Provider 테이블에 있는 의료 제공자 중 시술을 수행하는데 책임이 있는 의료 제공자를 참조하는 외래키



PROCEDURE OCCURRENCE

필드명	필수사항	데이터타입	설명 (한글)
visit_occurrence_id	No	integer (8)	(외래키) 시술이 수행되었던 시점과 연관있는 visit 테이블 내 병원 방문 기록을 참조하는 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
procedure_source_value	No	varchar(50)	원천 데이터내에 표현되어 있는 시술에 대한 코드값. 원천 데이터의 코드는 표준 사전의 concept와 매핑이 되며, 원천 데이터의 코드가 이 필드에 저장됨. 시술코드는 일반적으로 다음의 코드 체계로 되어 있음: ICD-9-Proc, CPT-4, HCPCS or OPSCS-4 codes.
procedure_source_concept_id	No	Integer (8)	(외래키) 원천 데이터에서 사용되고 있는 코드 체계를 지칭하는 컨셉에 대한 외래 키
modifier_source_value	No	varchar(50)	원천 데이터 내에서 사용되고 있는 시술에 대한 수식어구. 원천 데이터에 기록되어 있는 그대로 표현



MEASUREMENT

필드명	필수사항	데이터타입	설명 (한글)
measurement_id	Yes	integer (10)	각 측정(혈액 검사, 혈압 등)에 대한 고유한 식별자
person_id	Yes	integer (8)	(외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음
measurement_concept_id	Yes	integer (8)	(외래키) 측정 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
measurement_date	Yes	date	측정 날짜
measurement_time	No	time	측정 시간
measurement_type_concept_id	Yes	integer (8)	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 측정 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
operator_concept_id	No	integer (8)	(외래키) 측정 결과에 적용되는 수학적 기호를 참조하는 외래키 수치값 및 <, ≤, =, ≥, >
value_as_number	No	float (8,3)	숫자로 저장되는 측정값. 측정 결과가 숫자값으로 표현되는 측정에 적용 가능함
value_as_concept_id	No	integer (8)	(외래키) 컨셉 identifier로 저장되는 측정 값. 측정 결과가 표준 사전에 있는 표준 컨셉으로 표현되는 경우 적용 가능함 (예: 양성/음성, 존재/없음, 낮음/높음 등)



MEASUREMENT

필드명	필수사항	데이터타입	설명 (한글)
unit_concept_id	No	integer (8)	(외래키) 표준 사전 내에 정의된 측정 단위에 대한 표준 컨셉 identifier를 참조하는 외래키
range_low	No	float (8,3)	해당 측정의 정상범위의 하한치
range_high	No	float (8,3)	해당 측정의 정상범위의 상한치
provider_id	No	integer	(외래키) 해당 측정에 대한 책임이 있는 의료 제공자 (provider 테이블 내)를 참조하기 위한 외래키
visit_occurrence_id	No	integer (8)	(외래키) 해당 측정이 기록된 방문(visit 테이블)을 참조하기 위한 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
measurement_source_value	No	varchar(50)	원천 데이터내에 표현되어 있는 측정 명칭
measurement_source_concept_id	No	integer (8)	(외래키) 원천 데이터에서 사용되고 있는 코드 체계를 지칭하는 컨셉에 대한 외래 키
unit_source_value	No	varchar(50)	원천 데이터에 표현되어 있는 그대로의 단위에 대한 코드 값
value_source_value	No	varchar(50)	숫자 및 컨셉으로 저장된 구조화된 값과 관련된 원천 데이터의 값



OBSERVATION

필드명	필수사항	데이터타입	설명 (한글)
observation_id	Yes	integer	각 측정(혈액 검사, 혈압 등)에 대한 고유한 식별자 (외래키) 관찰 기간의 대상이 되는 환자를 식별한 외래키. 관찰 대상 환자의 자세한 인구학적 정보가 person 테이블에 저장되어 있음
person_id	Yes	integer	
observation_concept_id	Yes	integer	(외래키) 관찰 관련된 표준 사전 내의 표준 컨셉 identifier를 참조
observation_date	No	date	관찰 날짜
observation_datetime	Yes	datetime	관찰 시간
observation_type_concept_id	Yes	integer	(외래키) 표준 사전 상에 미리 정의되어 있는 컨셉 중 관찰 기록이 저장되는 원천 데이터의 타입을 지칭하는 컨셉을 참조하는 외래키
value_as_number	No	float	숫자로 저장되는 측정값. 측정 결과가 숫자값으로 표현되는 측정에 적용 가능함
value_as_string	No	varchar(60)	문자로 저장되는 측정값. 측정 결과가 문자값으로 표현되는 측정에 적용 가능함
value_as_concept_id	No	Integer	(외래키) 컨셉 identifier로 저장되는 관찰 값. 관찰 결과가 표준 사전에 있는 표준 컨셉으로 표현되는 경우 적용 가능함 (예: 양성/음성, 존재/없음, 낮음/높음 등)



OBSERVATION

필드명	필수사항	데이터타입	설명 (한글)
qualifier_concept_id	No	integer	각 관찰(과거력, 가족력 등)에 대한 고유한 식별자 concept id (외래키) 표준 사전 내에 정의된 관찰 단위에 대한 표준 컨셉 identifier를 참조하는 외래키
unit_concept_id	No	integer	(외래키) 해당 측정에 대한 책임이 있는 의료 제공자 (provider 테이블 내)를 참조하기 위한 외래키
provider_id	No	integer	(외래키) 해당 측정이 기록된 방문(visit 테이블)을 참조하기 위한 외래키
visit_occurrence_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
visit_detail_id	No	integer	(외래키) 해당 상태가 결정 된 (진단을 받은) 방문 (visit_detail 테이블)을 참조하기 위한 외래키
observation_source_value	No	varchar(50)	원천 데이터내에 표현되어 있는 측정 명칭
observation_source_concep t_id	Yes	integer	(외래키) 원천 데이터에서 사용되고 있는 코드 체계를 지칭하는 컨셉에 대한 외래 키
unit_source_value	No	varchar(50)	원천 데이터에 표현되어 있는 그대로의 단위에 대한 코드 값 원천 데이터 내에서 사용되고 있는 관찰에 대한 수식어구. 원천
qualifier_source_value	No	varchar(50)	데이터에 기록되어 있는 그대로 표현 (예: 알러지, 흡연, Hepatitis 등)



Explore the AUSOM database

CDMPv1_achilles1.6

Dashboard

CDM Summary

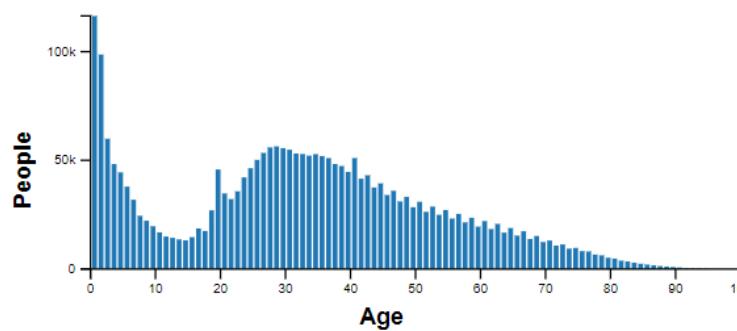
Source name: achilles_CDMPv1_table
Number of persons: 2.94M

Population by Gender

■ FEMALE
■ MALE



Age at First Observation

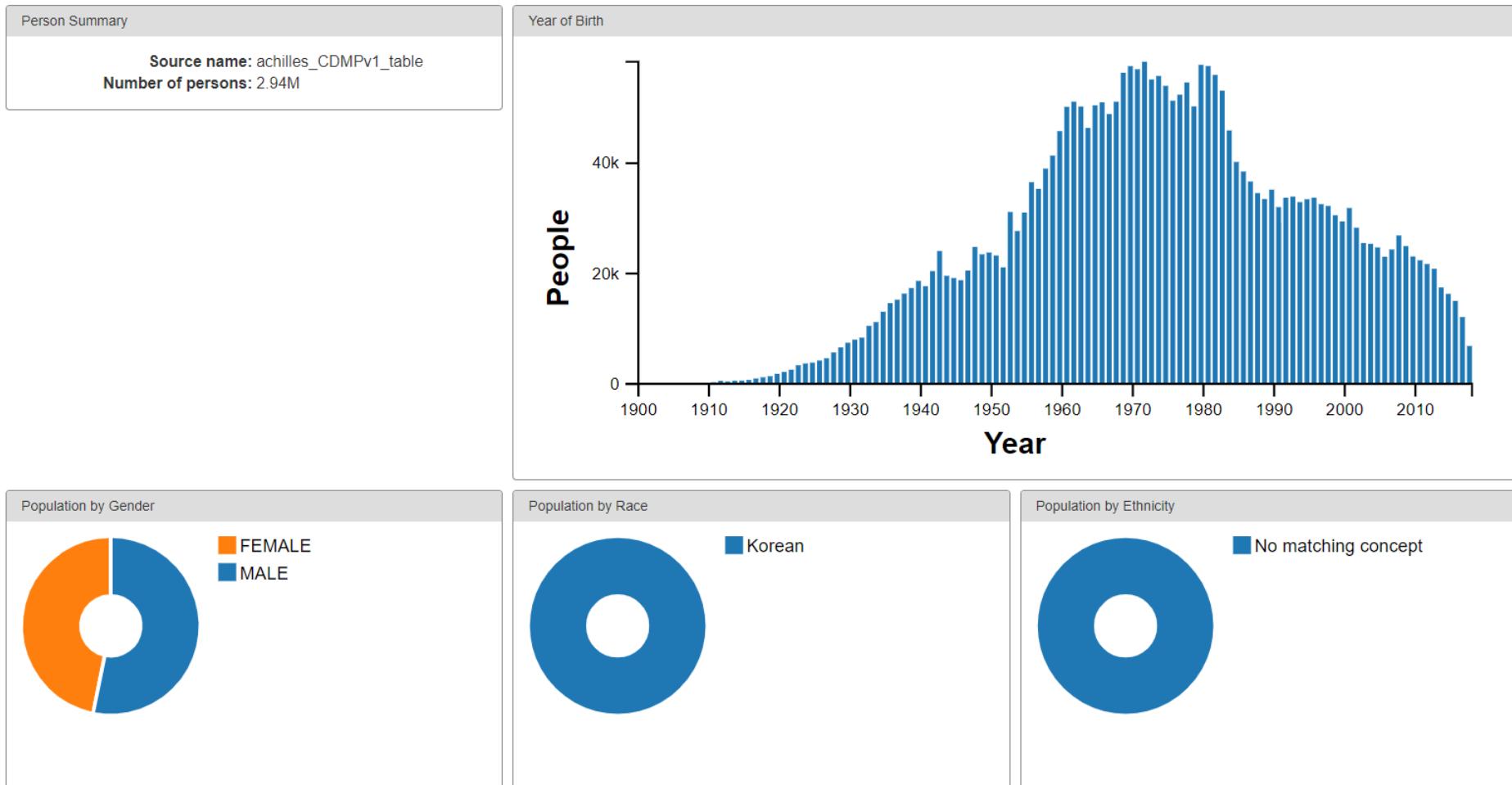


- Ajou University School Of Medicine (AUSOM)
- Patients: 2.94M
- Period: 05/1994-12/2017
- CDM version: v5.3



Explore the AUSOM database

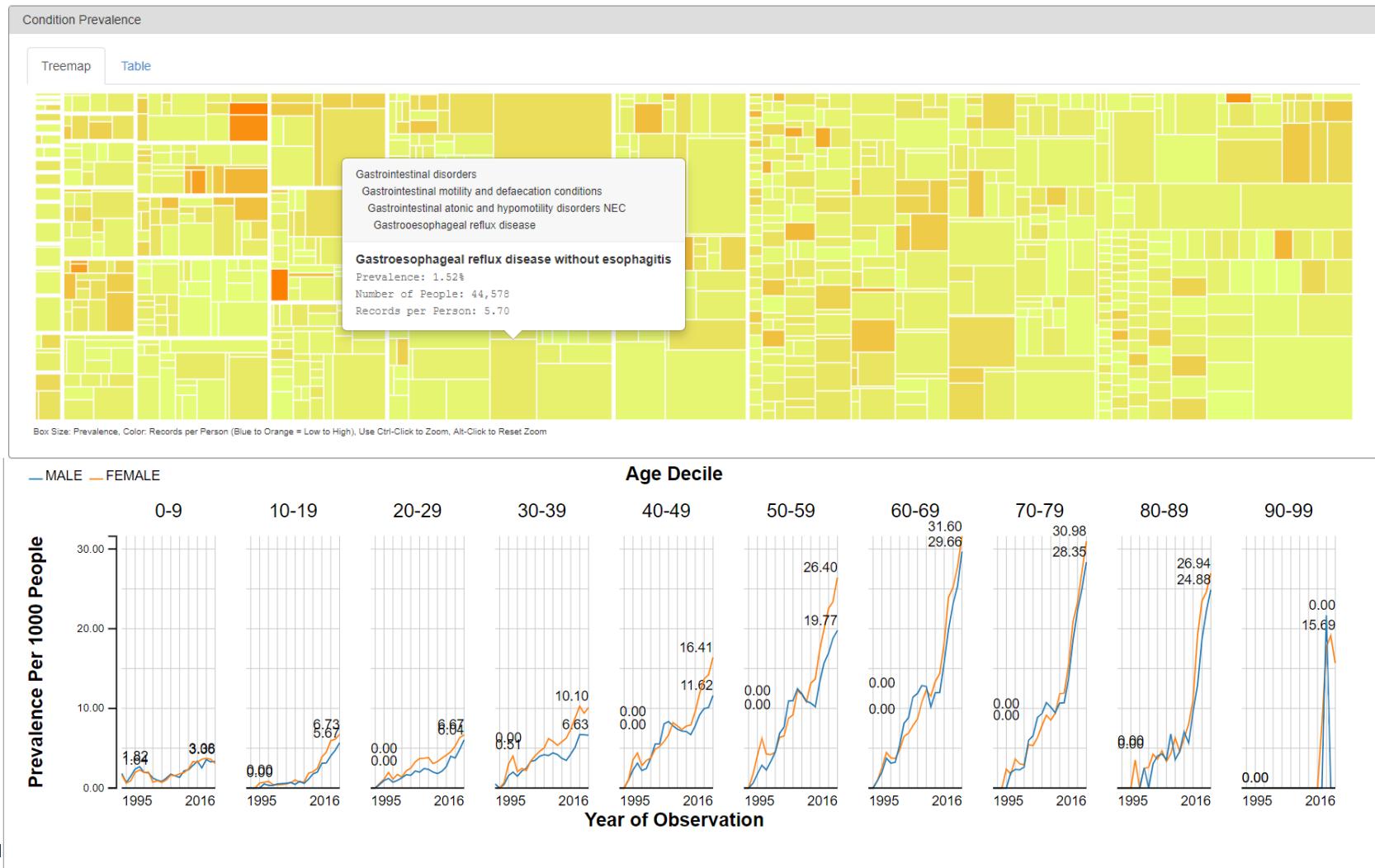
- Person table - 2,940,379 records





Explore the AUSOM database

- Condition_occurrence table - 32,544,907 records





Explore the AUSOM database

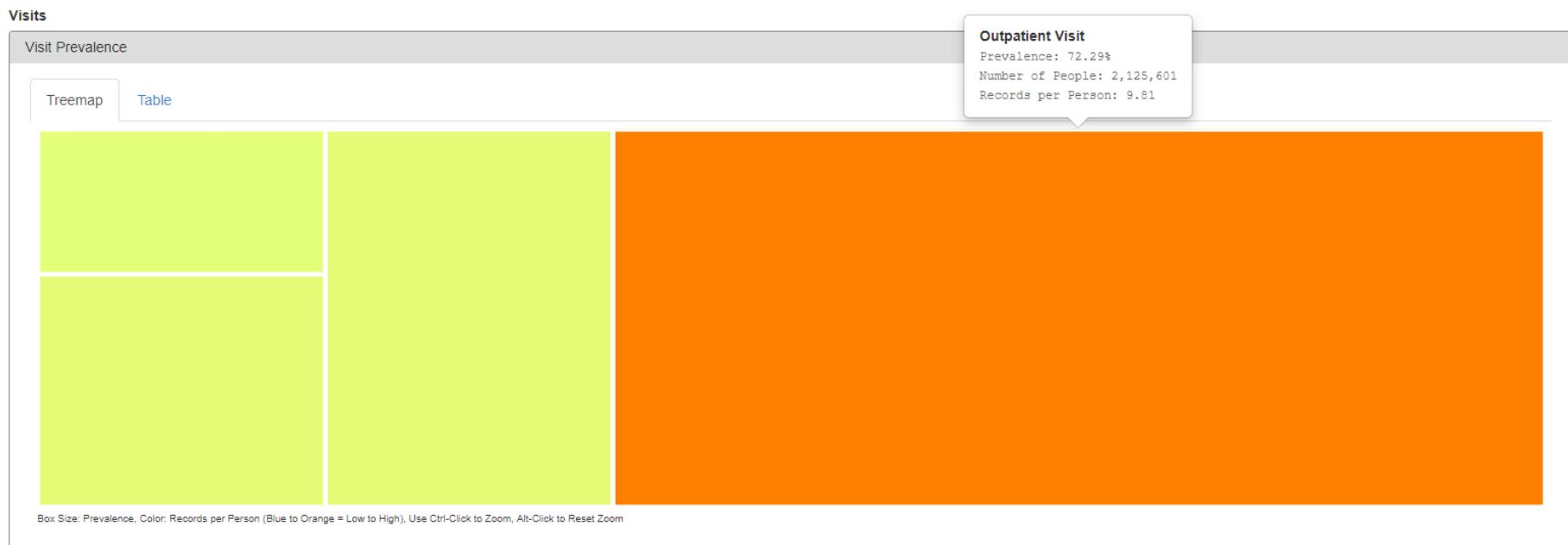
- Drug_exposure table - 122,339,472 records





Explore the AUSOM database

- Visit_occurrence table - 22,896,595 records





ECG-ViEW II (<http://ecgview.org/>)

Electrocardiogram Vigilance with Electronic data Warehouse
longitudinal observational QT database for surveillance on proarrhythmic potential of marketed drugs

Home | About ECG-ViEW | Data Request | Downloads | Publications | Contacts

Welcome to the Electrocardiogram Vigilance with Electronic data Warehouse II(ECG-ViEW II), the database comprising numeric parameters (RR interval, PR interval, QRS duration, QT interval, QTc interval, P axis, QRS axis, and T axis) of Electrocardiogram for surveillance of proarrhythmic potential of marketed drugs.

The database contains more than ~0.7 million of QT/QTc measurements and associated clinical data, including diagnoses(~7.7 million), drug prescriptions (~49 million) and selected laboratory-test results (~3.8 million), from ~461 thousands patients over the 19-years study period. The data is intended for researchers interested in surveillance on proarrhythmic potential of marketed drugs.

Scanned ECG printout → Cropped text area → Optical Character Recognition → Recognized ECG parameters → Text-parsing → Cropped text area (loop)

Background cleansing

Vent. rate 77 bpm
PR interval 168 ms
QRS duration 88 ms
QT/QTc 424/479 ms
P-R-T axes 67 32 153

Health examination Outpatient Emergency

Inpatient All

<Overall process to extract ECG parameters from scanned ECG> <The QTc distribution and size of study populations in ECG-ViEW>

좋아요 공유하기 39명이 좋아합니다. 친구들이 무엇을 좋아하는지 알아보려면 기입하기

This site is maintained by ABMI

Systems Biomedical Informatics
National Core Research Center



AUSOM dataset for Datathon 2019

- All CDM Table 생년월일, 방문일자, 사망일자 등의 날짜 DATA Shifting
- 민감 질병 해당 환자의 data 삭제
- 하위 0.5 percentile 이하 및 상위 99.5 percentile 이상의 값들을 각 하위 0.5 percentile 및 상위 99.5 percentile에 해당하는 값으로 대치
- CDM의 모든 테이블이 포함되지 않음 (13개 테이블만 포함)
- Local computer로 download 금지
- 이후 연구는 별도로 contact (www.ecgview.org / d.yoon.ajou@gmail.com)



Sample data

- <https://www.ohdsi.org/data-standardization/>