

ATLAS Ver 2.5.0

일산병원  
CDM 사용자 가이드

2019. 5.



국민건강보험 일산병원  
National Health Insurance Service Ilsan Hospital



**ATLAS Ver 2.5.0**

## **일산병원 CDM 사용자 가이드**

---

저 자 | 의료정보실 한창훈, 신동교, 최문영, 박성호, 김란희  
연구소 임현선

펴낸이 | 김성우

펴낸곳 | 국민건강보험 일산병원 의료정보실

주 소 | 경기도 고양시 일산동구 일산로 100

전 화 | 1577-0013

발행일 | 2019년 5월 30일



## 머 리 말

본 일산병원 CDM(Common Data Model) 사용자 가이드는 2018년 12월 우리병원에 구축한 CDM에 대하여 일산병원 사용자들의 접근성을 제고하고, 원활하게 사용할 수 있게 유도함으로써 궁극적으로 연구 성과를 극대화하고자 제작되었습니다.

현재까지 우리나라의 여러 병원들이 CDM을 구축했음에도, 완성도 있는 사용자 가이드를 제작하여 배포하거나 진료과별 교육을 적극적으로 실시한 병원은 없었습니다. 이런 점에서 우리병원 의료정보실 및 연구소의 우수성과 발전가능성을 엿볼 수 있습니다.

CDM 조기 활성화를 위한 각 진료과 교육만 해도 노고가 적지 않았을 텐데, 이에 그치지 않고 사용자 가이드를 제작한 일산병원 의료정보실의 의무기록팀을 비롯해 의료정보팀 및 연구소의 연구분석팀의 노고에 깊이 감사드리며, 이 지침서의 발간을 진심으로 축하드립니다.

앞으로도 일산병원 의료정보실과 연구소가 우리병원 진료정보의 숨은 가치를 발굴해내고, 우리나라 의료정보의 표준화를 이끌어가는 주역이 되기를 바랍니다.

2019년 5월  
일산병원장 김성우

## 감사의 글

세상에는 여러 가지 의사소통 방법이 있겠지만, 아카데믹 필드에서 가장 정확하고 효율적인 의사소통 방법은 ‘글’이라고 생각합니다. 일산병원 의료정보실은 2018년 12월 CDM 구축을 끝내고 조기 활성화에 대한 고민 끝에 교육계획을 수립하였습니다. 주어진 교육 시간은 매우 부족했으며, 연구자들에게 좀 더 많은 정보 전달이 필요하다는 결론에 도달하여 이 책을 제작하게 되었습니다.

이 가이드에는 CDM DB의 내용, CDM을 이용해 자료를 추출하고 연구를 수행할 수 있는 ATLAS에 대한 사용법은 물론 통계 검정을 통한 연구결과 도출과 해석, 그리고 타병원 자료까지 확대하여 연구를 수행하는 방법을 상세히 담고 있습니다.

이에, 실질적으로 CDM을 활용하여 연구를 수행하는 연구자들이 가까이 두고 지속적으로 참고할 수 있는 지침서로써 매우 유용할 것으로 보입니다. 그럼에도 불구하고 부족한 점이 있을 수 있으며, 사용자들의 고견을 통해 이 지침서는 더욱더 완성도와 가치를 높여나갈 수 있을 것입니다. 따라서, 이 책이나 CDM에 대한 개선 의견이 있다면 아낌없이 의료정보실로 전달해 주시면 감사하겠습니다.

제작을 위해 일산병원 의료정보실의 의무기록팀에서 수개월간 노력하였고, 의료정보팀과 연구소의 연구분석팀에서도 도움을 아끼지 않았습니다. 앞으로도 사용자 의견을 반영하여 지속적 개정과 CDM 자료의 가치를 창출해 나가는데 힘쓰겠습니다.

이 사용자 가이드는 ATLAS Ver 2.5.0 기준으로 작성되었으며, 일산병원 연구자를 위해 지속적으로 개정될 예정입니다. 감사합니다.

2019년 5월  
일산병원 의료정보실장 한창훈

# ATLAS Ver 2.5.0

## 일산병원 CDM 사용자 가이드

### 목 차

I. 일산병원 CDM 개요	1
1.1 CDM/OHDSI	1
1.2 일산병원 CDM	2
1.3 용어	3
II. ATLAS	4
2.1 개요	4
2.2 ATLAS 사용법	4
2.2.1 자료원(Data Sources)	4
2.2.2 용어–검색(Vocabulary–Search)	6
2.2.3 컨셉셋(Concept Sets)	9
2.2.4 코호트 편집(Editing Cohorts)	13
2.2.5 발생률(Incidence rates)	17
2.2.6 프로파일(Profiles)	17
2.2.7 추정(Estimation)	18
2.2.8 예측(Prediction)	19
2.2.9 작업현황(Jobs)	20
2.2.10 환경설정(Configuration)	20
2.2.11 환류(Feedback)	20
III. ATLAS를 이용한 Cohort 작성 사례	21
3.1 OHDSI가 관찰데이터로부터 만들어내고자 하는 근거	21
3.2 OHDSI가 연구자를 돋는 방법	21
3.3 코호트 정의를 위한 주요 기준	22
3.4 ATLAS를 이용한 코호트 생성 과정	23
3.5 ATLAS를 이용한 코호트 생성 예제	26

IV. Vocabulary .....	60
4.1 개요(Introduction) .....	60
4.1.1 개관(Overview, Background and Motivation) .....	60
4.1.2 배경(Background) .....	61
4.1.3 원칙(Principles) .....	62
4.1.4 정의 및 용어(Definitions and Terms) .....	64
4.2 일반적 구조와 사용(General Structure and Use) .....	65
4.2.1 도메인과 용어(Domains and Vocabularies) .....	65
4.2.2 표준, 분류 및 원천 개념(Standard, Classification & Source concepts) .....	66
4.2.3 개념의 맵핑(Mapping of Concepts) .....	68
4.2.4 자료 ETL(Data ETL) .....	74
4.3 도메인(Domain) .....	77
4.3.1 도메인 할당(Domain assignment) .....	77
4.3.2 병태 도메인(Condition domain) .....	80
4.3.3 장치 도메인(Device Domain) .....	83
4.3.4 약물 도메인(Drug Domain) .....	83
4.3.5 민족성 도메인과 용어(Ethnicity domain & Vocabulary) .....	90
4.3.6 성별 도메인과 용어(Gender domain & Vocabulary) .....	90
V. CDM 자료를 이용한 추정(Estimation)과 예측(Prediction) .....	92
5.1. ATLAS에서의 Estimation .....	92
5.2. ATLAS에서 Estimation 실행 방법 .....	92
5.2.1. Cohort 선정 .....	92
5.2.2. 모형 선정 .....	92
5.2.3. 추적 및 워싱아웃 기간 선정 .....	93
5.2.4. 환자 선정 .....	93
5.2.5. Propensity score matching .....	94



# ATLAS Ver 2.5.0

## 일산병원 CDM 사용자 가이드

5.2.6. 최종 모형 선정 .....	96
5.2.7. Negative control outcomes .....	97
5.2.8. 저장 .....	97
5.2.9. R코드 복사 .....	98
5.2.10. R Studio 실행 .....	98
5.3. 모형에 따른 결과 해석 .....	99
5.3.1 Logistic regression .....	99
5.3.2 Poisson regression .....	100
5.3.3 Cox regression .....	101
 VI. 시스템 구성현황 및 Table layout .....	102
6.1 시스템 개요 .....	102
6.1.1 시스템 구성현황 .....	102
6.1.2 분석서버 .....	102
6.1.3 DB서버 .....	102
6.2 테이블 목록 .....	103
6.3 테이블 상세 목록 .....	106
6.4 표준화된 2차 자료 .....	126
 VII. 다기관 공동 연구 .....	133
 참고자료 .....	135
 부록 : 1. R 환경설정 매뉴얼	
2. FeederNet 사용자 매뉴얼(일반사용자) v.1.1	



# I

# 일산병원 CDM 개요

## 1.1. CDM<sup>1)</sup> / OHDSI<sup>2)</sup>

OHDSI는 대규모 분석을 통해 보건의료 관찰 데이터(Observational Health Data)의 가치를 이끌어내기 위해 노력하는 여러 이해당사자 및 학제간의 협업 프로그램으로 모든 솔루션을 오픈소스로 제공하고 있음. 중앙조정본부(central coordinating center)는 미국 컬럼비아 대학에 있으며, 2018년 현재 OHDSI 연구 네트워크에 전 세계 6억 명이 넘는 환자 자료가 등록되어 있음.

OHDSI를 이용하기 위해서는 기본적으로 다음 3단계를 수행하여야 함.

- 1) 보유 DATA를 OMOP CDM<sup>3)</sup>으로 변환
- 2) OHDSI 도구를 활용하여 DATA 파악
- 3) OHDSI 연구 네트워크에 가입하여 2018년 현재 전 세계 6억5천만 명 이상의 환자를 대상으로 하는 연구에 협력

(OHDSI의 미션) 커뮤니티로 하여금 더 나은 보건의료 관련 의사 결정과 관리를 촉진하는 근거를 공동으로 생성할 수 있도록 하여 건강을 개선함.

(OHDSI의 비전) 관찰 연구를 통해 건강과 질병을 포괄적으로 이해함.

(OHDSI의 목적)

- 혁신(Innovation) : 관찰 연구는 혁신적인 사고로부터 큰 이익을 얻을 수 있는 분야이므로 새로운 방법론적 접근을 적극적으로 추구하고 장려함.
- 재현성(Reproducibility) : 건강 개선을 위해서는 정확하고, 재현 가능하며, 잘 보정된 근거가 필요함.
- 공동체(Community) : 환자, 건강 전문가, 연구원 또는 이 조직을 단순히 믿는 사람까지 모두가 OHDSI에 적극적으로 참여하는 것을 환영함.
- 협력(Collaboration) : 공동체 참여자들의 현실세계(real world)의 요구를 우선시하고 해결하기 위해 협력함.
- 개방성(Openness) : 방법, 도구, 그리고 생성하는 근거를 포함하여, 공동체의 모든 과정(proceeds)을 공개하고 공개적으로 접근 가능하게 하도록 노력함.
- 권익(Beneficence) : 항상 공동체 내의 개인과 조직의 권리를 보호하고자 함.

1) Common Data Model

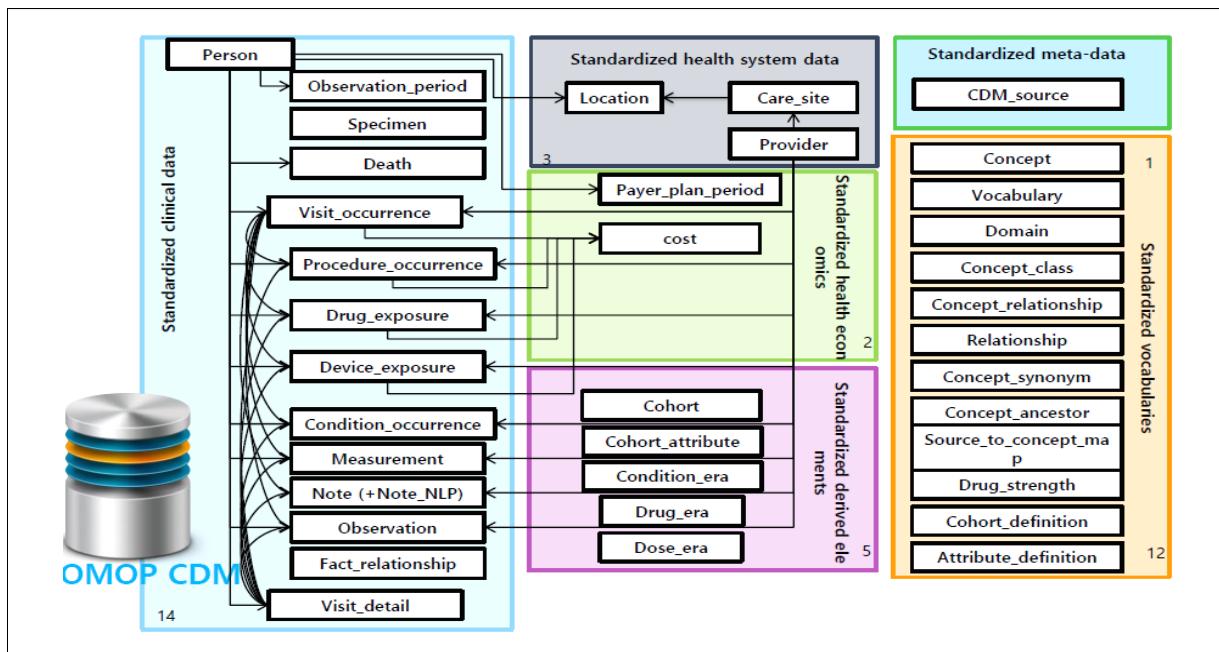
2) Observational Health Data Sciences and Informatics

3) Observational Medical Outcomes Partnership Common Data Model

## 1.2. 일산병원 CDM (<http://cdm.nhimc:8080/Atlas>)

우리병원은 개원이후 축적된 환자의 임상 DATA를 OMOP CDM으로 변환하였고 OHDSI 네트워크에 가입함.

- 적재 자료(개원~) : OMOP CDM [환자 중심으로 구축된 표준화된 임상 데이터 모델]
- 적재 주기 : 1회/1년 ※ 매 적재 시 개원부터 종료일까지 전수 재추출
- CDM Database : 병원 자료를 기반으로 하는 테이블과 이를 이용해 2차적으로 구축된 ‘Standardized Derived Tables’ , 그리고 표준 용어 테이블로 구성되어 있음.
- CDM Database 구성도



연구자는 우리병원의 ATLAS라는 툴을 이용하여 직접 CDM DATA를 추출하고 통계분석을 할 수 있으며, 필요 시 아래 담당자에게 문의할 수 있음.

- 코호트 작성(Concept set 정의 및 Vocabulary) : 의무기록팀 김란희(☎0805)
- 통계 분석 : 연구분석팀 임현선(☎6998), 박해용(☎6988)
- CDM DB 운영(추출 주기, 타 병원 자료 요청) : 의료정보팀 박성호(☎0398)
- IRB 승인 : 임상연구지원팀 김인순(☎0657)

### 1.3. 용어

- 병원 임상데이터 : 해당 요양기관에서 보유하는 CDM 변환 대상 데이터로써 환자기본정보 및 처치, 처방, 약 처방, 진단내역, 청구내역 등을 포함함.
- CDM(Common Data Model, 공통데이터모델) : 병원별로 상이한 임상데이터를 구조와 의미에 있어 동일한 규격으로 정의하는 모델
- CDM 데이터 : CDM으로 변환된 병원 임상데이터. 구체적인 데이터 목록은 본 가이드북의 ‘Ⅵ장’과 같으나, 추후 CDM 버전 개정에 따라 추가, 확대, 변경될 수 있음.
- OMOP CDM(Observational Medical Outcomes Partnership, OMOP) : 서로 다른 데이터베이스에 포함된 데이터를 공통 표현(용어, 어휘, 코드 체계)과 공통 형식(데이터 모델)으로 변환한 다음, 이를 기반으로 작성된 표준 라이브러리를 사용하여 체계적으로 분석하게 하는 데이터 표준 모델
- 분산형 바이오헬스 통합 데이터망 : 각 병원의 임상데이터를 CDM으로 변환 후, 데이터를 한 곳에 모으지 않고도 분석한 결과만을 모아 evidence를 제공하는 방식으로 다기관공동연구가 가능하게 하는 데이터망
- 코디네이팅 플랫폼 : 다수의 데이터 보유기관을 연계하고 이를 연구기관 및 데이터 수요처와 연결, 지원하는 코디네이팅 플랫폼으로, 병원 내 CDM 데이터를 보관, 관리하는 코디네이팅 플랫폼 Node와 각 Node를 통합, 운영 관리하는 코디네이팅 플랫폼 Portal로 구성됨.
- EMR(Electronic Medical Record, 전자의무기록) : 병력, 건강상태, 진찰, 입/퇴원기록 등 환자의 모든 정보를 전산화하여 입력, 관리, 저장하는 형태
- OCS(Order Communication System, 처방전달시스템) : 전산망을 통해 의사의 처방을 진료 지원부에 전달, 조회 할 수 있도록 하는 솔루션
- ERD(Entity Relationship Diagram) : 데이터와 데이터의 관계를 표현하여 개체 간 관계를 표시한 다이어그램

## 2.1. 개요

- ATLAS는 OMOP CDM을 이용해 연구자가 과학적 분석을 수행할 수 있도록 무료 이용이 가능한 웹기반 오픈소스 소프트웨어 도구임.
  - 연구자는 진료비 청구 자료(healthcare claims data)를 이용해 특정 약제에 대한 노출이나 특정 병태(condition)를 나타내는 진단명을 가진 사람들로 정의하는 방법으로 그룹을 설정하여 코호트를 생성할 수 있음.
  - ATLAS는 특정 질환을 가진 또는 특정 약제에 노출된 사람들을 구분할 수 있는 의료 관련 개념을 색인할 수 있는 ‘CDM 용어(CDM Vocabulary)’를 보유하고 있음.
  - 환자 정보는 특정 주제의 진료기록을 데이터로 형상화함으로써 특정 집단 안에 포함될 수 있게 함.
  - 해당 집단의 영향력 수준을 예측하는 연구는 두 개의 다른 코호트 집단을 비교할 수 있으며 R 패키지를 활용할 수 있음.
- ATLAS의 원활한 이용을 위해서는 구글 크롬 사용을 권장하며 셋업과 환경설정은 관련 웹사이트<sup>4)</sup>를 참조

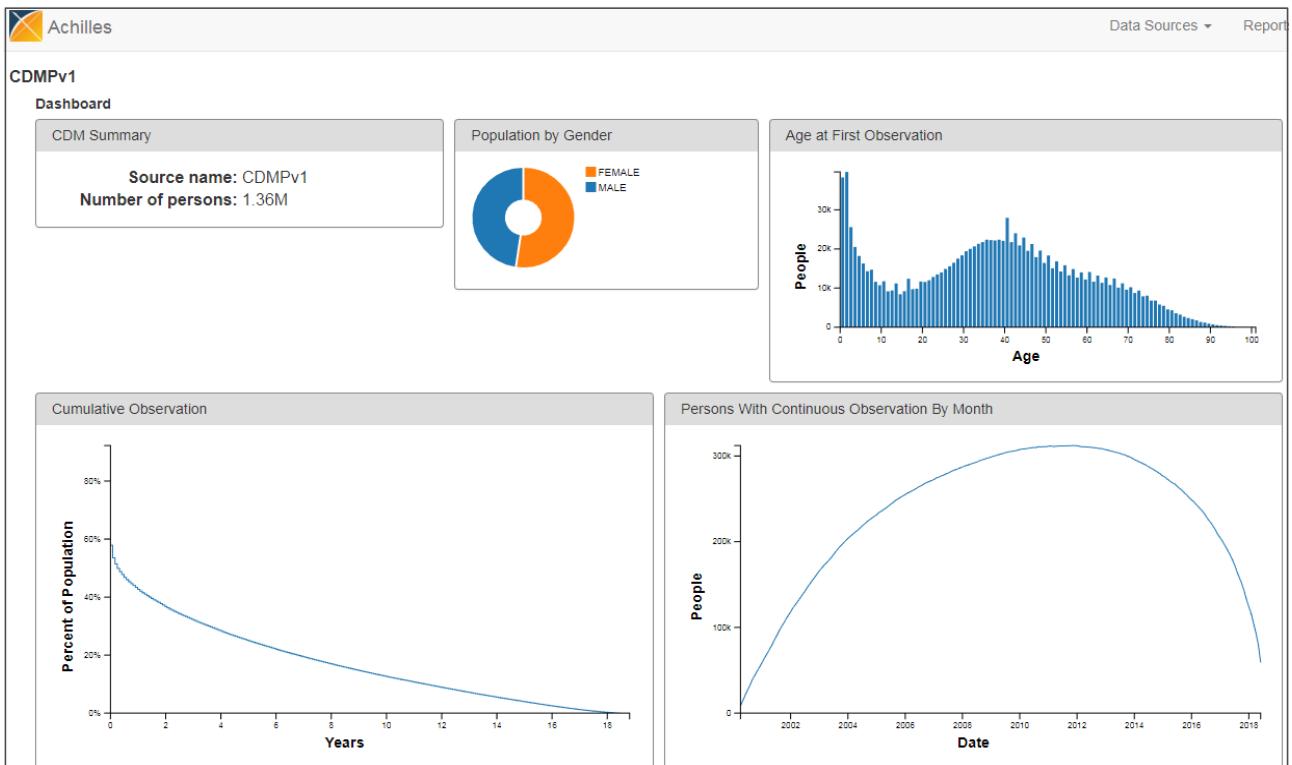
## 2.2. ATLAS 사용법

### 2.2.1. 자료원(Data Sources)

- ATLAS는 OMOP CDM에 대하여 기술 통계로 확인해 볼 수 있는 기능을 제공하며, 이 기능은 Achilles working group에서 처음으로 구축하였음.
- ATLAS는 자료를 추출하고 통계분석을 할 수 있는 툴이고, Achilles는 구축된 자료의 내용을 기 고안된 시각화 디자인에 따라 확인할 수 있는 툴임.
- 일산병원 CDM을 이용하기 위해서는 일산병원 내에서 구글 크롬을 이용해서 접속해야만 ATLAS를 원활하게 사용할 수 있음.

4) <http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:setup>

- Achilles : <http://192.168.1.53:8080/Achilles/#/CDMPv1/dashboard>



- ATLAS : <http://cdm.nhimc:8080/Atlas/#/home>

The screenshot shows the ATLAS application interface with the following sections:

- Left Sidebar:** Includes links for Home, Data Sources, Search, Concept Sets, Cohort Definitions, Incidence Rates, Profiles, Estimation, Prediction, Jobs, Configuration, and Feedback.
- Home Page:**
  - Home:** Welcome message and documentation link.
  - Getting Started:** Buttons for "Define a New Cohort" and "Search the Vocabulary".
  - Release Notes:** Links to ATLAS Version 2.5.0 Release Notes and WebAPI Version 2.5.0 Release Notes.
  - Known Issues:** A list of 19 feature enhancements and issue resolutions.

- 보고서 탐색하기 : Achilles
  - Atlas에서 Data sources 메뉴 선택
    - 우리 병원에 등록된 환자 수부터 각 CDM 테이블에 적재된 자료의 전반적인 현황을 파악할 수 있음. 시각화된 디자인 변경은 불가능하나 관심이 가는 부분에 커서를 올리는 등의 작동으로 관련 자료 찾기 등 확인 가능
    - 데모버전에서는 SynPUF의 데이터 자료를 활용할 수 있고, 이는 메디케어, 메디케이드 관련 보험 청구 자료 중 공공 이용 목적의 종합자료에 해당함 (<http://forums.ohdsi.org/t/synpuf/4936>).
  - 그 아래 드롭다운을 선택하여 Data source에 대한 보고서를 선택
  - 보고서(영문)별로 트리맵 또는 테이블 보기 옵션을 제공하는데 컨셉별로 드릴다운하여 접근할 수 있음.
    - 예를 들어 ‘Condition’ 보고서는 기본 값으로 트리맵 뷰를 제공함.
    - ‘Table’ 옵션을 선택하면 트리맵에서 볼 수 있는 컨셉을 표형식의 보기로 확인할 수 있음.
    - ‘Table’ 화면 오른쪽 상단에 있는 ‘Search’를 이용하여 보이는 결과의 필터링 가능
  - ※ 관심 있는 컨셉에 대해서는 트리맵/테이블 보기에서 해당 컨셉 줄 또는 셀을 드릴다운하게 되면 추가 상세정보가 보임.  
[\(http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:data-sources\)](http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:data-sources)

### 2.2.2. 용어-검색(Vocabulary-Search)

- Atlas는 OMOP 의 표준화된 용어를 검색하고 탐색할 수 있는 기능을 제공함. 이 기능은 원래 Hermes working group에서 개발하였고 현재는 Atlas에 완전히 통합되었음.
- Atlas의 왼쪽 메뉴에서 ‘Vocabulary’를 선택하면 두개의 탭(‘search’, ‘import’)이 나타남.

#### ■ Search

- 검색어를 입력하면 Atlas는 표준용어 전체를 검색하여 해당 용어와 일치하는 ‘concept’을 찾음. 와일드카드 검색<sup>5)</sup>을 수행하므로 ‘Dementia’라는

---

5) wildcard search : 전체 단어(full term)를 입력하지 않아도 해당 용어(term) 관련 개념(concept)을 검색할 수 있음. 다시 말해, dementia라는 용어를 검색하면 dementia가 포함된 모든 concept을 불러옴. 검색이 완료되면, 결과가 결과 탭에 확인됨.

용어를 검색하면 ‘Dementia’ 가 포함된 ‘concept’ 을 찾음. 검색이 완료되면 결과가 ‘Results’ 창에 표시됨.

The screenshot shows the ATLAS software interface. On the left is a sidebar with various navigation options: Home, Data Sources, Search (which is selected), Concept Sets, Cohort Definitions, Incidence Rates, Profiles, Estimation, Prediction, Jobs, Configuration, and Feedback. Below the sidebar, it says "Apache 2.0 open source software". The main area is titled "Search" and has tabs for "Search" (which is active) and "Import". A search bar contains the text "DEMENTIA". Below the search bar are buttons for "Column visibility", "Copy", "CSV", and "Show 15 entries". To the right of these are "Advanced Options" and a "Filter" input field. The main content area displays a table of search results. The table has columns: Id, Code, Name, Class, RC, DRC, Domain, and Vocabulary. The table shows 15 of 602 entries. The first entry is "4341802 N0000000914 Dementia". The last entry is "4218017 416780008 Primary degenerative dementia of the Alzheimer type, presenile onset". At the bottom of the table, there are buttons for "Previous" and "Next", and a legend for "Classification" (purple square), "Non-Standard" (red square), and "Standard" (blue square).

- Import를 클릭하면 Concept set으로 연결됨(Concept set 참조)

## ■ Advanced options

- 이 옵션은 ‘search’ 버튼 아래에 있으며 검색 쿼리를 수행하고, 그 결과를 특정 용어사전 및 도메인으로 필터링 할 수 있음.

## ■ Results

- ‘Results’ 창은 기본 검색 또는 고급 검색의 결과가 표시됨.
- 화면 왼쪽에 있는 필터를 사용하여 검색 결과를 더 정제할 수 있으며 필터 목록 및 작동법은 다음과 같음.
  1. Vocabulary : 특정 용어사전을 필터링할 수 있다.
  2. Class : concept의 특정 Class(등급)을 필터링 할 수 있다.
  3. Domain : 특정 Domain을 필터링 할 수 있다.(Condition, Observation, Drug, ...)
  4. Standard Concept : 표준화된 concept 여부에 따라 필터링 할 수 있다 (Standard, Non-standard, ...)

5. Invalid Reason : 유효하지 않은 concept을 제외시킬 수 있다.
6. Has Records : CDM에 source daimon을 구성한 경우 당신의 CDM에 해당 concept의 data가 있는지를 확인해준다.
7. Has Descendant Records : ‘Has records’ 와 유사하게 CDM의 데이터가 포함 된 하위 개념의 concept을 필터링 해준다.

- ‘Results’ 에서는 Concept의 다음 정보를 포함함.
  - Concept ID : 용어사전에서 고유의 concept ID
  - Code : 근원에 해당하는 용어사전에서의 본래 코드
  - Name : concept의 이름. 글씨 색<sup>6)</sup>으로 구분
  - Class : 원천이 되는 용어사전에서의 concept의 분류/등급
  - RC : 레코드 개수. CDM에서 해당 컨셉으로 코드화된 레코드의 개수
  - DRC : 하위 레코드 개수. CDM에서 코드화된 하위 레코드의 합
  - Domain : 해당 컨셉의 도메인 (영역, 범위)
  - Vocabulary : 해당 컨셉과 관련된 용어사전

Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
4181852	100000000014	Dementia	Int / Cl	0	242862	Drug	MPAT
4182210	52440006	Dementia	Clinical Finding	1,364	93,141	Condition	SNOMED
4082747	279942005	Cerebral degeneration presenting primarily with dementia	Clinical Finding	0	32,384	Condition	SNOMED
21604792	N05D	ANTI-DEMENTIA DRUGS	ATC 3rd	0	31,836	Drug	ATC
374888	100318003	Dementia associated with another disease	Clinical Finding	18,242	30,961	Condition	SNOMED
4048875	15641003	Senile dementia	Clinical Finding	0	34,174	Condition	SNOMED
275705	100448005	Uncomplicated senile dementia	Clinical Finding	13,963	13,963	Condition	SNOMED
440103	429918004	Vascular dementia	Clinical Finding	0	10,351	Condition	SNOMED
378125	10048006	Presenile dementia	Clinical Finding	1,317	7,089	Condition	SNOMED
21604798	N05E	Other anti-dementia drugs	ATC 4th	0	6,341	Drug	ATC
379704	10049005	Senile dementia with depression	Clinical Finding	4,356	4,356	Condition	SNOMED
4101137	100497008	Senile dementia with depressive or paranoid features	Clinical Finding	0	4,356	Condition	SNOMED
430276	100410004	Uncomplicated arteriosclerotic dementia	Clinical Finding	4,330	4,330	Condition	SNOMED
380986	370234007	Senile dementia with delusion	Clinical Finding	3,888	3,888	Condition	SNOMED
376989	100493009	Uncomplicated presenile dementia	Clinical Finding	2,958	2,958	Condition	SNOMED

- 6) Green: 분류된 concept  
     Blue: 표준 concept(SNOMED-CT)  
     Red: 비표준 concept

- ‘Results’ 테이블의 특징
  - Paging : 한 페이지에 기본 15건이 나타남(‘Results’ 테이블 상단에 드롭다운 버튼을 이용하여 설정가능). ‘Results’ 테이블 오른쪽의 페이지 넘김 기능을 활용하여 데이터를 넘겨 볼 수 있음.
  - Filtering : ‘Results’ 테이블의 오른쪽 상단에 필터링할 용어를 입력하는 텍스트 상자가 있음. 이 박스에 용어를 타이핑하면 즉각적으로 결과가 필터링 됨.
  - Sorting : 각 열을 정렬할 수 있음.
  - Shopping Cart : ‘Shopping Cart’ 를 클릭함으로써 새로운 Creating concept sets에 사용할 컨셉을 고르게 됨. 아직 새로운 Concept sets를 생성하지 않고 ‘Shopping Cart’ 를 클릭한 경우, 자동적으로 새로운 Concept set이 구성 됨.

## ■ Import

- ‘Import’ 탭은 concept을 코호트 정의를 위해 가져오는 방법임.
  1. Concept identifiers : 쉼표로 구분된 concept\_ids 목록을 사용하여 Concept 을 Concept set으로 가져옴.
  2. Source Codes : 쉼표로 구분된 소스코드 값 목록을 사용하여 Concept을 concept set으로 가져옴.
  3. Concept set : Concept set을 나타내는 JSON object를 가져옴.

### 2.2.3. 컨셉셋(Concept Sets)

- Atlas를 이용하여 Concept Sets를 만들 수 있으며, 이는 연구의 표준화된 용어의 concept 목록으로 연구의 관심주제를 설명함.

## ■ Browsing concept sets

- Atlas에 메뉴에서 ‘Concept Sets’ 를 선택하면 이 시스템에서 이용가능한 모든 concept set를 보여줌.
- 이 테이블에는 ‘Vocabulary’ 메뉴의 Results table과 유사한 필터 및 정렬 기능이 있음.
- 이 화면에서 기존 Concept set을 선택하여 탐색하거나 ‘Add Concept Set’ 버튼을 클릭하여 새로운 Concept set를 생성할 수 있음.

## ■ Creating concept sets

- Concept Set을 만드는 방법
  - 기존 Concept set를 찾아 열거나 새로운 Concept set 추가
  - concept 추가 : ‘Vocabulary’ 를 탐색하여 새로운 Concept set을 구성하기 위해 사용할 Concept을 찾아 ‘shopping cart’ 아이콘 클릭
  - 이 경우 새로운 Concept set은 왼쪽 메뉴 가장 상단에 나타나며 이를 통해 현재 작업 중인 Concept set을 확인할 수 있고, 해당 Concept set으로 빠르게 돌아갈 수 있음.
- Concept set에 새로운 Concept 추가
  - ‘Vocabulary’에서 Concept을 검색하고 선택
  - 추가하고자 하는 Concept을 발견한 경우, ‘shopping cart’ 아이콘을 클릭하면 현재 Concept set에 해당 Concept이 추가됨.
  - 단, 한 번 더 ‘shopping cart’ 아이콘을 누르게 되면 현재 Concept set에서 해당 용어가 삭제됨.

## ■ Concept set expression

- 구성한 Concept set에 한 개 이상의 Concept이 있는 경우, Concept set expression 템을 통해 Concept set에 포함되어야 하는 개념을 추가로 정의 할 수 있음.
- Concept set expression 템에서 표 형식으로 선택한 Concept 목록을 확인할 수 있으며, Concept의 선택/비선택을 위한 3개의 체크박스가 있음.
  1. Exclude : 선택 시 Concept set에 해당 Concept이 사용되지 않음.
  2. Descendants : 자동적으로 해당 용어의 하위분류를 모두 선택함. 이 옵션과 Exclude 옵션을 함께 사용 시 해당 Concept은 물론 그 하위분류까지 모두 제외함.
  3. Mapped : 용어사전 내 관계를 사용하여 선택한 Concept에 연관되어 있는 모든 Concept을 자동으로 선택함.

Concept ID	Concept Code	Concept Name	Domain	Standard Concept Caption
1306216	290046	Lisinopril	Drug	Standard

- concept set expression을 정의할 때 다른 템(Included Concepts, Included Source Codes, Explore Evidence, Export, Compare)을 사용하여 Vocabulary에 대한 표현이 어떻게 평가되는지 탐색하여 개념을 선택할 수 있음.

### ■ Included concepts

- ‘Included concepts’ 템은 Concept set을 평가할 때 포함되는 모든 표준 concept을 표시하는데 사용됨.
- 먼저 Concept set의 일부로 포함되어 있는 표준 Concept을 검토하고 원편에 있는 필터를 사용하여 concept의 다양한 속성을 더 자세히 탐색할 수 있음.

### ■ Included source codes

- ‘Included source codes’ 템은 Concept set을 평가할 때 포함되는 모든 비표준 Concept들을 보여주는데 사용됨.
- Concept set을 평가할 때 포함된 모든 소스코드들을 검토할 수 있으며, 이는 ICD-9과 같은 소스코드에서 표준 개념들로 변환하여 정의가 완전함을 검증할 때 유용함. 또한, concept set을 다른 소스코드를 가진 데이터소스를 변환한 내용을 확인하는 데 유용함.

### ■ Explore evidence

- ‘evidence’ 템은 LAERTES<sup>7)</sup> Evidence base로 부터 개념별 근거(Concept specific evidence)를 보여주고 있음.
- Evidence base에는 출간된 문헌, 상품 표시 및 자발적 보고서가 있으며, 이를 사용하여 concept set에 포함된 Concept과 관련된 근거를 확인할 것

### ■ Export

- ‘Export’ 템은 Concept set expression이나 Concept ID 목록을 내보내는데 사용됨.
- concept set의 정의를 외부인과 공유하고자 할 때 유용함.
- 그런 다음 vocabulary area에서 import 기능을 사용하여 이 정의를 가져올 수 있음.

### ■ Compare

- ‘Compare’ 템은 두 Concept set expression을 비교하는데 사용되며 연구들 간의 Concept이 정의된 방식과 연구 결과들 간의 연관된 차이점을 보여줌.

---

7) Large-scale adverse effects related to treatment evidence standardization

## ■ Save

- ATLAS 내에서 Concept set을 계속 사용하기 위해서는 명명하고 저장해야 함.
- DB 내 저장 : Concept set 이름 박스 오른쪽 옆에 저장버튼 있으며 저장이 발생하면 이름 박스 오른쪽에 복사버튼이 활성화됨:

## ■ Close

- Concept set 이름박스의 오른쪽에 위치한 닫기 버튼을 클릭하여 Concept set 사용 종료
- 이렇게 하면 왼쪽 화면에서 참조가 제거됨.

## ■ Copy

- Concept set 저장 후 Concept set 이름박스 오른쪽의 ‘Copy’ 버튼 클릭하여 Concept set 복사

## ■ Optimize

- Concept set 생성 완료 후 Concept set 이름박스 오른쪽에 위치한 ‘optimize’ 버튼을 사용하여 주요 Concept에서 불필요하거나 누락된 concept을 볼 수 있음.
- 가장 적은 변수를 이용하여 Concept set을 구성할 수 있도록 concept을 추가하거나 삭제함으로써 Concept set을 수정할 수 있는 방법 제시
- 선택사항을 확인 후 현재 Concept set을 덮어 쓰거나 새로운 Concept set을 생성하기 위해 파란색 저장 옵션 박스를 선택하면 됨.

## ■ Delete

- Concept set 이름박스 오른쪽에 위치한 ‘Delete’ 버튼 클릭하여 더 이상 사용하지 않을 Concept set이나, 시험 중인 Concept set 삭제

## ■ Scope of concept sets

- 주어진 Concept set의 사용 범위는 한정적임. 즉, 새로운 코호트에서 Concept set을 복사하고 이를 삭제, 변경 등의 수정을 하면 원본에 해당하는 Concept set에 영향을 미치지 않고 대체 가능한 코호트로만 사용됨.

## ■ Domains of concept sets

- 주어진 Concept set안에서 일반적으로 다수의 영역(예를 들어, 진단, 검사, 처치)을 혼합해서 사용하는 것은 권장되지 않으나 필요할 때가 있음.

- 코호트에서 새로운 기준 추가를 선택하고 진단 기준 추가, 검사기준 추가, 치치기준 추가 등을 선택하는 방식으로 각 영역에서의 Concept set을 참조하여 하나의 코호트 내에서 참조할 수 있음.

### ■ A note on use of ‘at most’ vs. ‘exactly’ in definition of concept set occurrence

- 기준 정의 시 ‘at most’ 와 ‘exactly’ 는 주어진 concept set의 발생 횟수를 나타내는데 사용함.
- ‘at most 0’ 와 ‘exactly 0’ 은 실제 사용에 차이가 없으며 방식의 선호로 한 연구에서 한 가지 방식으로 사용하는 것이 가독성이 좋고 혼돈을 줄일 수 있음.

#### 2.2.4. 코호트 편집(Editing Cohorts)

- 분석 시작 시 일반적으로 첫 번째 단계는 연구하고자 하는 집단을 정의내리 는 것으로, 이를 코호트라 함.
- 원래 Circe working group의 한 부서에서 제작하였고, 현재 Circe는 ATLAS에 완전히 통합됨.

### ■ Browsing cohorts

- 기존 코호트는 ‘Cohort Definitions’ 메뉴의 옵션을 통해 코호트의 정의 및 기타 기능에 대한 탐색이 가능함. 필터 기능을 사용하여 검색문자 기반으로 목록을 필터링 할 수 있으며, 테이블 열의 정보를 정렬하여 원하는 코호트를 찾을 수 있음.
- 새로운 코호트는 화면 오른쪽 상단의 ‘New Cohort’ 클릭하여 생성

### ■ Cohort Definition

- 코호트는 특정 약제를 투약한 자, 특정 검사결과가 있는 자 또는 두 가지 조합을 포함하여 생성될 수 있음.
- 코호트 생성
  - ‘Cohort Definitions’에서 화면 오른쪽 상단의 ‘New Cohort’ (파란색) 클릭
  - Home 화면에서 ‘Define a New Cohort’ 버튼 클릭
- 코호트 제목 설정
  - 고유한 서술형 코호트 제목 설정 후 ‘Save’ 버튼 클릭
  - 만들고자 하는 코호트의 기본 첫 제목은 ‘New Cohort Definition’으로 화면 가장 위쪽의 텍스트 란에 설정되어 있음.

- 현재 작업 중인 코호트 확인
  - 생성한 코호트는 항상 현재 작업 중인 코호트로서 열려 있게 되고, ‘Close’ 버튼을 클릭하지 않는 이상 주요 콘텐츠 구성화면 왼쪽상단에 보임.
- 새로운 코호트를 생성하려면 기준에 열려있는 코호트를 닫아야 함.
- Cohort Definition(코호트 정의) : 코호트가 무엇인지 코호트를 정의하기 위해 필요한 기준은 무엇인지 등을 설명하는 것으로 다음과 같은 코호트 항목 기준을 선택하고 정의
  - Initial Event
  - Initial Event Inclusion Criteria
  - Additional Qualifying Inclusion Criteria(화면 오른쪽): 각 기준들을 추가적으로 정제 가능
- Cohort Exit<sup>8)</sup> Criteria 및 Cohort Censoring<sup>9)</sup> Events : Cohort Definition과 같은 방법으로 정의(필수 사항 아님)
- 용어사전 활용: 인구통계학적 기준 외에 대부분의 기준을 적절히 정의하기 위해서는 concept sets를 불러와 정의하고 이를 향후 재사용 할 수 있음
- ‘New Concept Set’ : 이전에 concept set을 생성해두지 않았다면, concept set 불러오기 화면에서 ‘New Concept Set’ 버튼 클릭을 통해 원하는 기준을 입력하고 코호트 생성

## ■ Concept Sets<sup>10)</sup>

- concept sets 탭에 코호트 내에서 정의한 concept set들이 나열됨.
  - 표에서 해당 concept set을 클릭하면 해당 concept set의 정의를 검토할 수 있으며, 형태는 Concept set 섹션의 양식과 동일함.
  - 열려 있는 코호트 concept set은 현재 concept set으로 왼쪽 메뉴에 나타남.

## ■ Generation : Cohort 대상자 추출을 위한 SQL code 생성

- 코호트 정의가 완료되면 ‘Generate’ 탭을 사용하여 Cohort 대상자 명단을 생성할 수 있음.
- CDM Data source table에서 아래 속성들에 대하여 확인 가능함.
  - Generate button : CDM에 SQL code가 실행되어 정의한 코호트 대상자를 선택할 수 있음.

8) Cohort exit : 일반적으로 cohort\_end\_date에 적용함

9) Censor : cohort\_start\_date나 cohort\_end\_date에 모두 적용 가능함.

10) [http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:concept\\_sets](http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:concept_sets)

- Source name : CDM의 source 명칭
- Generation status : cohort 생성 활동 상태를 나타냄.
- Distinct People : 코호트가 생성되었을 때, 코호트 내의 대상자 수를 중복 없이 보여줌.
- Generated : 코호트가 생성된 날짜와 시간
- Generation Duration : 코호트를 생성하는데 걸린 시간

## ■ Reporting

- 생성된 코호트에 대한 정의 내용을 좀 더 기술적으로 가독성 있게 설명하고 있으며, 연구자가 보고서 작성 시에도 활용할 수 있음.
- Reporting 템은 코호트의 요약 및 시각화 도구들을 제공함.
- 이 기능은 본래 Heracles working group의 한 부서에서 제작하였으며 현재 Heracles는 Atlas에 완전히 통합되었음.
- Reporting 템에서 이 기사<sup>11)</sup>의 Heracles 분석 리스트에서 정의한 보고서를 포함한 컬럼과 각 CDM source별로 정의된 컬럼을 볼 수 있음.
  - 그 컬럼의 표시를 체크해둠으로써 각 보고서를 선택할 수 있고,
  - Generate 버튼을 선택해서 보고서를 생성해낼 수 있음.
  - ‘Jobs’ 메뉴를 선택하면 보고자 하는 분석 작업이 시작됨.
  - 분석이 완료되면 각 보고서에 ‘view’ 버튼이 생성되며 이를 통해 그 결과를 확인해볼 수 있음.

## ■ Drilldown reports

- 일부 보고서에는 세부내용을 파악할 수 있는 최고 수준의 통계를 포함함. (Condition, Condition Eras, Drug Eras, Drug Exposure, Drugs by Index, Procedure and Procedures by Index).
- 이 보고서들은 트리맵<sup>12)</sup>과 표라는 두 가지 템의 초기 시각화를 제공함.
  - 트리맵 : 표 템에 표시되는 데이터 시각화
  - 표 템 : 각 보고서의 도표보기 제공
  - 트리맵 형식의 행을 클릭하거나 셀을 클릭하면 선택한 concept에 대한 세부 사항이 나타남.

11) <http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:heracles>

12) 입력된 수치를 면적으로 변환하여 보여주는 시각화 방식

## ■ Print Friendly

- 이 템은 코호트의 정의를 사용자가 읽을 수 있도록 설명을 제공
- 이는, 코호트의 정의를 이루는 concept set 및 source code와 함께 코호트를 정의하는데 사용되는 로직을 설명하는데 유용함.

## ■ How to copy an existing cohort/JSON

- JSON 템은 코호트의 정의 및 concept set을 지원하는 JavaScript Object Notation (JSON) 표현을 제공함.
- 코호트의 정의를 복사 및 붙여넣기 하여 OHDSI network의 다른 구성원에게 제공하며 연구의 일부 또는 같은 환경에서 동일한 코호트를 생성할 수 있음.
- 기존 코호트를 복사하기 위해 왼쪽메뉴의 ‘Cohort’ 버튼을 클릭하고 복사하고자 하는 코호트를 선택함.
- 특정 코호트의 JSON 코드 복사 절차
  - 관심 있는 코호트 내에서 Export 템 선택 → JSON button 클릭 → 아래 박스를 클릭 → Ctrl+A 클릭 → Ctrl+C 클릭
- 코호트의 로컬 복사본 생성
  - 로컬 Atlas 사이트에서 cohort 버튼(왼쪽 메뉴) 클릭 → 페이지 오른쪽 상단의 New Cohort 버튼 클릭 → export 템 클릭 → JSON 버튼 클릭 → 박스를 클릭 → delete키를 눌러 박스안의 텍스트 모두 제거 → Ctrl+V 클릭 → reload 버튼 클릭

## ■ Saving

- 코호트의 명칭을 붙이고 코호트 이름 박스 밖에 위치한 ‘Save’ 버튼<sup>13)</sup>을 클릭함.
- 코호트가 새롭게 생성되고 코호트의 정의, concept set 및 새로운 코호트에 포함되는 각종 요소를 확인해야 함.

## ■ Show SQL

- ‘Show SQL’ 버튼을 클릭하면 코호트 정의에 기반을 두어 추출된 SQL문장을 보여주는 대화상자가 제공됨.
- ATLAS로부터 SQL편집기 및 기타 프로그램으로 복사/붙여넣기 가능

---

13) <http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:cohorts>  
[http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:concept\\_sets](http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas:concept_sets)

- 최근 지원하고 있는 다른 SQL언어에 대한 SQL도 제공  
(예. MSSQL Server, MS APS, Oracle 등)

## ■ Close

- 현재 코호트 정의가 닫히며 저장되지 않은 변경사항이 있는 경우 저장하거나 변경사항을 취소하라는 메시지가 표시 됨.

## ■ Copy

- 코호트 복사 버튼()을 클릭하여 현재 정의하고 있는 코호트의 복사본을 생성

## ■ Delete Cohort

- 코호트 삭제 버튼() 클릭

### 2.2.5. 발생률(Incidence rates)

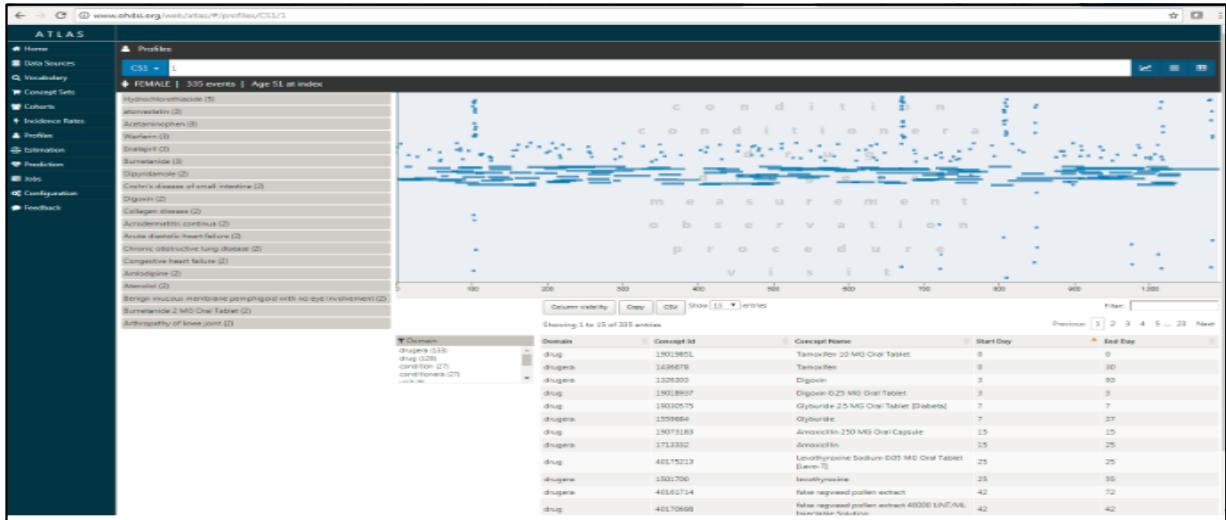
- 사용 가능한 모든 데이터 내에서 표적(Target) 코호트와 결과(Outcome) 코호트로 발생률(Incidence rates)을 생성하는 기능을 함.
- 점유율(Proportion)과 율(Rate)로 보고되며 대상, 사례 및 시간에 따른 위험도 등 제공

Source	Name	Persons	Cases	Proportion [+] per 10 persons	Time At Risk	Rate [+] per 10 years	Started	Duration
IPCT	ShlPUF 1%	670	62	92.54	285	210.17	2017-08-31, 08:48	08:05:51
CSI	ShlPUF 1K	7	1	542.86	5	200.00	2017-10-09, 18:15	08:00:02

### 2.2.6. 프로파일(Profiles)

- Atlas는 데이터베이스 내 개별 환자의 프로파일을 검색 및 탐색하는 기능
  - 환자 Profile 확인
    - 우선, 확인하고자 하는 데이터베이스 선택 후 환자 구분번호 입력
    - 왼쪽 메뉴에서 Conditions, Drugs, Procedure 등 각 도메인 필터링 가능

- 오른쪽 코너 하단의 table에서 환자 정보를 Domain, Concept IDs, Concept names, Start/End days로 구성하여 보여줌.



#### - 원하는 기간의 정보 확인

- 화면 중간의 그래픽은 지정된 시간(예: 처음 100일)을 확대되도록 크기 조정 가능
- 아래에 위치한 table에 지정한 시간대를 반영한 정보 내용으로 변경됨.

#### 2.2.7. 추정(Estimation)

- Atlas는 비교 코호트를 설계/활용하여 분석연구를 수행할 수 있는 기능 제공.
- ATLAS에서 제공하는 성향점수 분석(Propensity score methodology) 이용
  - 세 가지 주요 결과 모델 : Logistic regression(로지스틱 회귀분석), Poisson regression(포아송 회귀분석), Cox proportional hazards(콕스 비례 위험)

The screenshot shows the 'Population Level Effect Estimation' configuration page. The left sidebar has a red box around the 'Estimation' tab. The main area has sections for 'Specification' (with tabs for 'Unadjusted', 'Adjusted', and 'Stratified'), 'Statistical Model' (set to 'Logistic regression'), 'Time Period' (with dropdowns for 'start date', 'end date', and 'duration'), and 'Outcome' (with dropdowns for 'target cohort', 'comparator cohort', and 'outcome'). Below these are sections for 'Statistical Model' (checkboxes for 'remove patients who enter both cohorts', 'remove patients who have observed the outcome prior to cohort entry', and 'use propensity score adjustment as a confounding adjustment strategy for baseline covariates'), 'Baseline Covariates' (checkboxes for 'Demographic', 'Condition', and 'Treatment'), and 'Conditions' (checkboxes for 'age group', 'gender', 'ethnicity', 'education level', 'income', 'marital status', and 'religion'). A note at the bottom states 'At minimum required days off risk, applied to target and comparison cohorts.'

- 분석 기능에 대한 이해를 돋는 참고자료

- 비디오 튜토리얼 (유튜브) 참고

- ▶ OHDSI Estimation Methods Tutorial - Part 1 >> <https://youtu.be/jwHIVTxDPtY>
- ▶ OHDSI Estimation Methods Tutorial - Part 2 >> [https://youtu.be/\\_lSOli3TXtY](https://youtu.be/_lSOli3TXtY)
- ▶ OHDSI Estimation Methods Tutorial - Part 3 >> <https://youtu.be/qywULqu0lvM>
- ▶ OHDSI Estimation Methods Tutorial - Programmer Track (Part 4) >> <https://youtu.be/kgSxmhKptjk>
- ▶ OHDSI Estimation Methods Tutorial - Study Programmer(Part 4) >> <https://youtu.be/d4WjAkNWe8Y>
- ▶ OHDSI Estimation Methods Tutorial - Programmer Track (Part 4) >> <https://youtu.be/-mE9x-WO45Q>
- ▶ OHDSI Estimation Methods Tutorial - Part 5 >> <https://youtu.be/cjc9H1aWq40>

- 본 가이드의 VI 장. Estimation 설명 참고

### 2.2.8. 예측(Prediction)

- ATLAS는 정밀의료와 질병 차단을 위해 머신러닝 방법을 이용하여 다음과 같은 예측모델을 생성하는 기능을 내장(Regularized regression, Random forest, k-nearest neighbors)
- R패키지 patientLevelPrediction 사용
  - R패키지 patientLevelPrediction을 사용하여 CDM형식의 데이터로 환자 수준의 예측모델을 개발
- R패키지 patientLevelPrediction 의 특징
  - 코호트 및 관심있는 결과를 입력값으로 함.
  - OMOP CDM 형식의 데이터베이스 내에서 필요한 데이터를 추출
  - 모든 약, 진단, 처치, 나이, 중복이환 지수 등을 포함한 다양한 공변량들을 사용
  - 예측모델에 적합한 Large scale regularized regression(대규모의 정규화 회귀)
  - 예측모델을 평가하기 위한 기능 포함
  - 지원하는 결과모델 : Logistic, Poisson, Survival(time to event)
- Prediction 활용을 위한 참고 자료
  - 비디오 튜토리얼 (유튜브) : <https://youtu.be/BEukCbT8UoA>.
  - Github<sup>14)</sup> 의 평가 : <https://github.com/OHDSI/PatientLevelPrediction>

14) Git이라는 버전관리 시스템을 공유하여 소스코드를 발전시키고자 하는 웹 호스팅 서비스

## 2.2.9. 작업 현황(Jobs)

- 코호트를 생성하거나 코호트 보고서를 산출하는 등의 장시간의 프로세스로 Background에서 실행되고 있는 작업을 탐색할 수 있는 화면
- 화면 오른쪽 상단에 위치한 ‘Refresh jobs’ 버튼을 클릭하여 현재 실행되고 있는 작업이 있을 때 그 작업 상황을 새로 고칠 수 있음.

The screenshot shows the 'Jobs' section of the OHDSI ATLAS interface. On the left, there's a sidebar with navigation links like Home, Data Sources, Vocabulary, Concept Sets, Colors, Incidence Rates, Profiles, Estimation, Feedback, Jobs, Configuration, and Feedback. The 'Jobs' link is highlighted. The main area has a table with columns: Job Name, Status, Start Date, and End Date. A red box highlights the 'End Date' column header. Below the table, there's a detailed log of tasks:

Job Name	Status	Start Date	End Date
warming CEE cache	FAILED	8/6/2017 20:33 PM	8/6/2017 20:34 PM
warming LPCT cache	COMPLETED	8/6/2017 20:33 PM	8/6/2017 20:37 PM
REFRESH_CONTROLS_996	FAILED	8/6/2017 12:08:06 PM	8/6/2017 12:13:33 PM
warming CEE cache	COMPLETED	8/6/2017 11:33:41 AM	8/6/2017 11:36:06 AM
warming LPCT cache	COMPLETED	8/6/2017 11:33:41 AM	8/6/2017 11:36:30 AM
HERACLES_COHORT_9640_CES	COMPLETED	8/6/2017 11:33:41 AM	8/6/2017 11:32:04 AM
HERACLES_COHORT_9641_CES	COMPLETED	8/6/2017 11:28:11 AM	8/6/2017 11:31:49 AM
REFRESH_CONTROLS_886	FAILED	8/6/2017 11:34:25 AM	8/6/2017 11:34:27 AM
performing feasibility study on Index Population for Study: Feasibility of studying beta blocker use and risk of death in patients with ovarian cancer - OHDSI Cloud - 1K synthetic patients (CES)	COMPLETED	8/6/2016 9:11:35 AM	8/6/2016 9:11:31 AM
performing feasibility study on Index Population for Study: COPY OF MEDINFO tutorial - 120M patients for hypothesis study - OHDSI Cloud - 1K synthetic patients (CES)	FAILED	8/6/2016 9:11:51 AM	8/6/2016 9:11:03 AM
performing feasibility study on Index Population for Study: COPY OF MEDINFO tutorial - 120M patients for hypothesis study - OHDSI Cloud - 1K synthetic patients (CES)	FAILED	8/6/2016 9:12:04 AM	8/6/2016 9:10:45 AM
8/6/2016_141_Index_Population_for_Study_COPY_OF_COPY_OF_Feasibility_of_studying_beta_blocker_use_and_risk_of_death_in_patients_with_ovarian_cancer_CES	COMPLETED	8/6/2016 2:09:12 PM	8/6/2016 2:11:44 PM
generating cohort 802800 : DYNAPUF 1% (LPCT)	COMPLETED	8/7/2017 3:06:20 PM	8/7/2017 3:06:24 PM
generating cohort 802800 : SYNAPUF 1K (CES)	COMPLETED	8/7/2017 3:06:17 PM	8/7/2017 3:06:18 PM
generating cohort 802800 : SYNAPUF 1K (CES)	COMPLETED	8/7/2017 3:05:42 PM	8/7/2017 3:05:43 PM
generating cohort 802802 : DYNAPUF 1% (LPCT)	COMPLETED	8/7/2017 3:05:47 PM	8/7/2017 3:05:49 PM

## 2.2.10. 환경설정(Configuration)

- 환경설정 섹션에서 구성된 데이터소스를 검토하는 항목으로 여러 옵션들을 확인해볼 수 있으나 환경설정을 편집할 수는 없음. 원천 데이터의 수정은 데이터베이스 내에서 직접 이루어져야 함.

The screenshot shows the 'Configuration' section of the OHDSI ATLAS interface. The sidebar includes links for Home, Data Sources, Vocabulary, Concept Sets, Colors, Incidence Rates, Profiles, Estimation, Feedback, Jobs, Configuration, and Feedback. The 'Configuration' link is highlighted. The main area displays a table with columns: Source Name (Source Key), Table Qualifiers, Direct, Vocabulary Version, Evidence, Record Counts (IC / DRG), and Incidence. A red box highlights the 'Configuration' link in the sidebar.

Source Name (Source Key)	Table Qualifiers	Direct	Vocabulary Version	Evidence	Record Counts (IC / DRG)	Incidence
OHDSI_1 (http://opendatacommons.org/MEDINFO)	CEDB_synaptacit Vocabulary_unstructured Results_synaptacit	postgres	v1.0.23-029-17			
SHRIFT 1K (CES)	Vocabulary_unstructured Results_unstructured CEDB_unstructured Evidence_unstructured Evidence_structured Vocabulary_structured	postgres	v1.0.23-029-17			
LADIES (NUL)	Vocabulary_structured	postgres	v1.0.23-029-16	*		

## 2.2.11. 환류(Feedback)

- ‘issue log for ATLAS’ (<https://github.com/OHDSI/Atlas/issues>)에 접속하여 기존의 문제점을 검색해보거나 새로운 문제점에 대해 기록해 둘 수 있음. 새로운 기능 또는 질 향상을 위한 의견이 있다면, 이를 올려 공유함으로서 발전을 위한 커뮤니티로도 활용됨.

### III

## ATLAS를 이용한 Cohort 작성 사례

### 3.1. OHDSI가 관찰데이터로부터 만들어내고자 하는 근거

- 임상적 특성화(Clinical characterization)
  - 자연사(Natural history) : 당뇨에 걸린 사람이 누구인가? 그들 중 metformin을 먹은 사람은 누구인가?
  - 질 향상(Quality improvement) : 당뇨 환자 중 합병증을 경험한 비율은 얼마인가?
- 인구집단 수준 추정(Population level estimation)
  - 안전 감시(Safety surveillance) : metformin은 lactic acidosis를 유발하는가?
  - 비교효과(Comparative effectiveness) : metformin은 lactic acidosis를 glyburide 보다 더 많이 유발하는가? metformin과 glyburide의 치료 효과는 어떤 것이 더 좋은가?
- 환자 수준 예측(Patient level prediction)
  - 정밀의료(Precision medicine) : 나와 나의 medical history에 대하여 당신이 알고 있는 것을 모두 고려해 볼 때, 내가 metformin 투약을 시작하면 내년에 lactic acidosis를 경험할 확률은 얼마나 되는가?
  - 질병 차단(Disease interception) : 당신이 알고 있는 것을 모두 고려해 볼 때, 내가 당뇨에 걸릴 확률은 얼마나 되는가?

### 3.2. OHDSI가 연구자를 돋는 방법

- OHDSI는 다음과 같이 다양한 방법으로 연구자들이 실질적으로 연구를 진행 할 수 있도록 지원하고 있음.
  - OHDSI community를 통한 의사소통
  - Open-source knowledge base(LAERTES)
  - Open-source front-end<sup>15)</sup> web applications(ATLAS)
  - Open-source back-end<sup>16)</sup> statistical packages(R Methods Library)
  - OHDSI network studies

15) 사용자 인터페이스 방법으로 사용자들로부터 다양한 형태의 입력을 받아 프로세싱하는 방식

16) 사용자 인터페이스 방법으로 사용자들 눈에는 안보이지만 백그라운드 알고리즘으로 프로세싱하는 방법

### 3.3. 코호트 정의를 위한 주요 기준

- Cohort entry criteria
  - 초기 이벤트(Initial events)
    - 사람에 대한 약물 노출, 병태, 의료행위, 측정, 방문과 같은 이벤트가 시간이 명시되어(time stamped) 기록됨
    - 모든 이벤트는 시작일과 종료일이 있고, 어떤 이벤트(행위, 측정 등)는 시작과 종료일이 동일함.
  - 초기 이벤트 포함 기준(Initial event inclusion criteria) : 연구자가 정한 초기 포함 기준
- 부가적 적격 포함 기준 (Additional qualifying inclusion criteria)
  - qualifying cohort는 initial event가 발생했고, initial event의 포함 기준을 만족하며, 모든 부가적 적격 포함 기준을 만족하는 사람으로 정의됨.
- 코호트 종료 기준(Cohort exit criteria) : 대상이 코호트에서 빠져나가야 하는 기준
- 성향 점수(Propensity score) : 실험군과 대조군에 대한 배정 확률. 이를 실험군과 대조군 간에 유사하게 맞춤으로써 연구 결과에 따른 관심 변수의 효과를 Baseline characteristics의 차이로 인한 영향을 배제하고 해석 가능하게 함.
$$e(x) = \Pr(Z=1 | x)$$
  - Z: treatment assignment
  - x: a set of covariates at the time of treatment assignment
- 성향 점수(Propensity score) = 주어진 baseline covariates 하에 대상 코호트(target cohort)와 대조 코호트(comparator cohort)에 대한 확률
- 성향 점수(Propensity score)를 균형 점수(balancing score)로 사용할 수도 있음 : 두 개의 코호트가 유사한 Propensity score 분포를 가졌다면 공변량(covariates) 분포는 유사할 것임.
- 코호트 설계 요약(Cohort Design Summary)
  - 연구자가 실질적으로 코호트를 설계하는 것은 다음 5가지 요소를 구체화하여 선택하는 것으로 요약할 수 있음.

Input parameter	Design choice
Target cohort(T)	
Comparator cohort(C)	
Outcome cohort(O)	
Time at risk	
Model specification	

- CDM Specification을 숙지하여 다음 내용을 파악하면, ATLAS를 이용해 좀 더 적절하게 코호트를 생성하고 분석할 수 있음.
  - 원 자료의 소스코드가 무엇이었는가?
  - 테이블에 어떻게 들어가 있는가?
  - 매핑이 어떻게 되어있는가?
  - 도메인(Domain)의 의미
  - 용어(vocabulary)
- 즉, 의학 지식 뿐 아니라 CDM 데이터 생성 현황에 대해서도 잘 알면 좀 더 정확하게 연구 대상을 선택 가능함.

### 3.4. ATLAS를 이용한 코호트 생성 과정 요약

- 설계된 코호트를 ATLAS를 이용해 생성하기 위해서는 Concept set이 반드시 필요함. Concept set은 별도로 생성해서 등록해 놓고 필요 시마다 여러 코호트에서 재사용할 수도 있고, 특정 코호트 정의 중에 생성함으로써 해당 코호트에서만 사용할 수도 있음.

구분	Step1	Step2(ATLAS)	Step3(ATLAS)
방법 1	코호트 설계 (연구대상 정의)	필요한 Concept set 생성 (다른 코호트에서 재사용 가능)	코호트 생성 - Step 2의 Concept set 활용하여 조건 설정 - 필요시 Concept set 추가 생성 (다른 코호트에서 사용불가)
방법 2		코호트 생성 - Concept set 생성 (다른 코호트에서 사용불가) - 테이블 지정 - 상세 조건 설정	

- Concept set 생성 : 원하는 개념에 해당하는 코드 목록 확정
  - 코호트에서 원하는 대상을 단위 개념으로 각각 정의하는 것으로, Concept set을 정확하게 생성하는 것은 원하는 코호트를 생성하는데 매우 중요함.
  - OMOP CDM에서는 매우 방대한 용어가 사용되고 있으므로 적절한 옵션 설정으로 효율적으로 정확한 Concept set을 생성할 수 있음.
  - 다음과 같이 ‘NOAC’ Concept set을 생성할 수 있음.
    - **New Concept Set** → Naming : Concept\_NOAC → Search : NOAC(성분명으로 검색: dabigatran etexilate, apixaban, ribaroxaban, edoxaban) → filtering option 설정 → 선택 : → 저장 :
    - Filtering Option · Vocabulary : RxNorm
      - Class : ingredient
      - Domain : Drug
      - Standard Concept : Standard

#### ※ 주요 용어

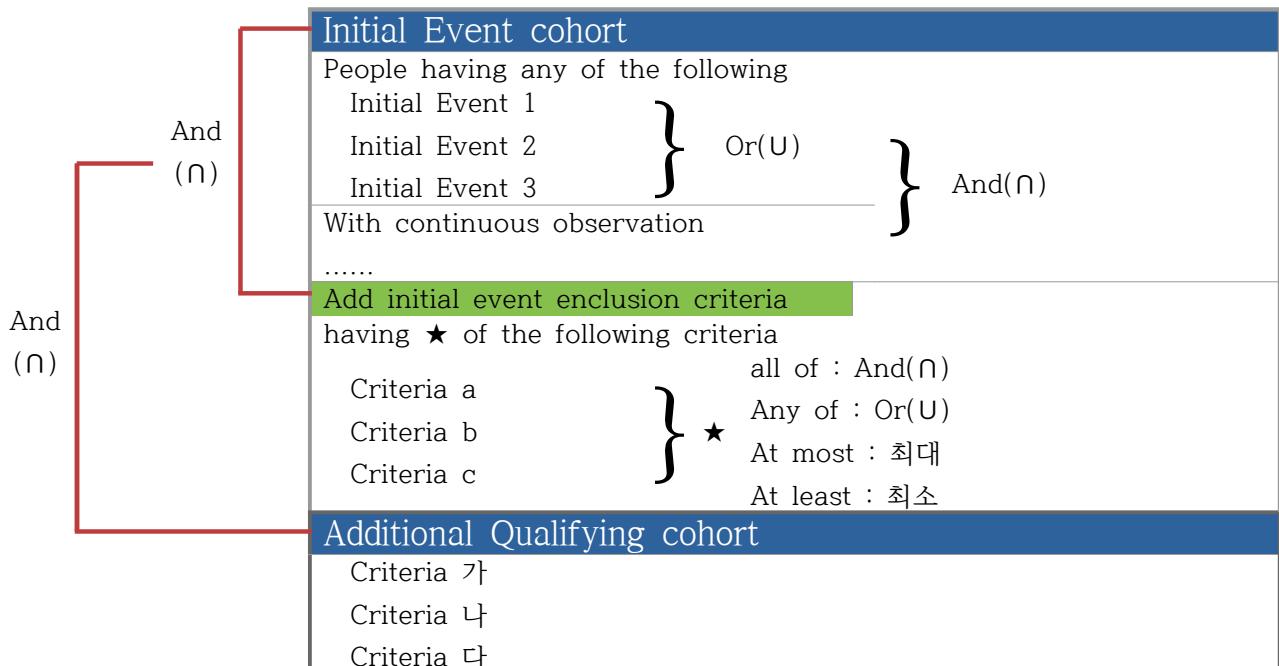
용어	의미
RC(Record Count)	해당 코드가 발생한 건수
DRC(Descendant Record Count)	해당 코드의 하위코드를 포함하여 발생한 건수
Filtering Option	Vocabulary
	사용된 용어 체계
	의미론적으로 유의한 분류단위 (Ingredient, Branded drug, Branded drug comp., ...)
	Domain
	자료 영역(Condition, Drug, Measurement, ...)
	Standard concept (사용된 용어 체계 범주)
	<ul style="list-style-type: none"> <li>· Standard: CDM 내에서 지정한 표준 용어(도메인별로 다름(Condition-SNOMED-CT, Drug-RxNorm, ...))</li> <li>· Classification: 분류체계이나 CDM 내 표준은 아님.</li> <li>· Non-Standard: 분류 체계가 아닌 분류</li> </ul>
	Invalid reason
Has Record	Valid(RC>0 or DRC>0), Invalid(RC=0 and DRC=0)
Has Descendant Record	False(RC=0), True(RC>0)
	False(DRC=0), True(DRC>0)

- Cohort 생성 : Concept set을 활용하여 원하는 연구대상을 한정지음.
  - Click or → 코호트 Naming :
    - Click → 저장:

- 코호트 정의(Cohort Definition) 화면과 부문별 기능

메뉴	코호트 정의 화면	부문별 기능
ATLAS	New Cohort Definition	← 코호트 이름 정하기
Home	Definition Concept Sets Generation Reporting Export Warnings	
Data Sources	enter a cohort definition description here	
Search		
Concept Sets		
Cohort Definitions	Initial Event Cohort People having any of the following: with continuous observation of at least [0] days before and [0] days after event index date Limit initial events to: earliest event per person. Add initial event inclusion criteria	← 코호트 설명
Incidence Rates		
Profiles		
Estimation		
Prediction		
Jobs		
Configuration	Additional Qualifying Inclusion Criteria New qualifying inclusion criteria Please select a qualifying inclusion criteria to edit.	Initial Event 지정 Index date setting
Feedback	Limit qualifying cohort to: earliest event per person.	← 제외 조건 설정
Apache 2.0 open source software provided by OHDSI join the journey	Cohort Exit Criteria Add a cohort exit criteria: • Based on a fixed time period relative to initial event start or end date • Based on the end of an era of persistent exposure to any drug within a defined concept set The minimum date from amongst the selected cohort exit criteria occurring after the cohort entry date will be selected as the end date for the person's episode.	← 코호트 종료 조건 (코호트 유효, 종료일)
	Censoring Events Exit Cohort based on the following: No censoring events selected. + Add Censoring Event	← 중도 절단 조건
	Cohort Collapse Strategy Specify era collapse gap size: [0] days This strategy collapses events into eras of a continuous period and eliminates the duplicate rows for the same subjects.	← 중복자료 처리 전략
	Cohort Censor Window Drops all persons, who do not fall inside the censor window. starting [YYYY-MM-DD] and ending [YYYY-MM-DD]	← 코호트 제외 기간

- 코호트 정의(Cohort Definition) 조건 적용 기본 구조



### 3.5. 코호트 생성 예제

[ 코호트 생성 예제 1<sup>17)</sup> ] : Entecavir vs Tenofovir against HBV

#### ○ Step1. 코호트 설계

- : B-viral hepatitis 치료를 위하여 Tenofovir(vs Entecavir) 초치료 받은 환자군
- Entecavir 180일 이상의 연속 처방 필요(20041230-20141229)
- 나이: 30~80세
- 제외 기준
  - HDV, HCV 등 기타 급성 간염 환자 제외
  - HIV 진단 환자 제외(전체 follow-up 기간 동안에도)
  - 간이식 환자 제외
  - 간암 환자 제외

Input parameter	Design choice	
	Initial	Qualifying(Exclusion)
Target cohort(T)	B-viral hepatitis +Entecavir(180일) +30~80세	<ul style="list-style-type: none"><li>· HDV, HCV 등 기타 급성 간염 환자</li><li>· HIV 진단 환자(전체 follow-up 기간 동안에도)</li><li>· 간이식 환자</li><li>· 간암 환자</li></ul>
Comparator cohort(C)	B-viral hepatitis +Tenofovir(180일) +30~80세	
Outcome cohort(O)	사망	
Time at risk	Entecavir/Tenofovir 복용	
Model specification	Entecavir/Tenofovir 복용이 생존/사망에 미치는 영향	

17) <https://www.youtube.com/watch?v=pMtJ3aBQ6sk>

## ○ Step2. Concept set 생성<sup>18)</sup>

: 코호트 정의 시 사용할 concept을 미리 생성해 둠.

- **Concept Sets** → Naming(저장) → Search(Filter-선택 ) → 저장

### 1) B-viral hepatitis Concept set 생성

- **Concept Sets** : B-viral hepatitis(Naming) → **Search** : hepatitis B
- Filter(Vocabulary, Class, Domain, Standard Concept) →

The screenshot shows a search interface for 'Concept Sets'. The search bar contains 'hepatitis B'. The results table displays 15 entries out of 43, with columns for Id, Code, Name, Class, RC, DRC, Domain, and Vocabulary. The results include various clinical findings related to hepatitis B, such as 'Chronic viral hepatitis B without delta-agent' and 'Viral hepatitis B without hepatic coma'. The left sidebar shows filters for Vocabulary, Class, Domain, and Standard Concept, with 'SNOMED (324)' selected under Vocabulary.

→ 추가 옵션(Exclude, Descendent, Mapped) 지정 → 저장

The screenshot shows the 'Concept Set' configuration interface for 'B-viral hepatitis'. It lists three included concepts: 'Viral hepatitis B without hepatic coma', 'Chronic viral hepatitis B without delta-agent', and 'Chronic viral hepatitis B with hepatitis D'. The interface includes tabs for 'Concept Set Expression', 'Included Concepts' (16), 'Included Source Codes', 'Explore Evidence', 'Export', and 'Compare'. Below the table, there are checkboxes for 'Exclude', 'Descendants', and 'Mapped', all of which are checked for the listed concepts.

18) 본 사례에서는 Concept set을 모두 코호트 정의 전에 생성하지만 Cohort Definition 중에도 생성 가능하고, 이 경우 해당 코호트에서만 사용가능함.

## 2) Entecavir Concept set 생성

The screenshot shows the SNOMED CT interface with the search term 'entecavir' entered. The results table displays one entry:

Vocabulary	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
RxNorm (22)	1711246	306266	entecavir	Ingredient	3,969	43,588	Drug	RxNorm

The left sidebar shows the classification hierarchy under 'Vocabulary' for RxNorm, which includes SNOMED (29), dm+d (9), VA Product (6), and GDN/SEMO (3). Other categories like Class, Domain, and Standard Concept are also listed.

The screenshot shows the Concept Set interface for the 'Entecavir' concept. The search bar contains 'Entecavir'. The results table displays one entry:

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
1711246	306266	entecavir	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Below the table, there are checkboxes for 'Classification' (purple), 'Non-Standard' (red), and 'Stand' (blue).

### 3) Entecavir 1MG Concept set 생성

; 용량까지 세분하여 조건으로 줄 경우 사용 - Drug Exposure table

The screenshot shows a search interface for 'entecavir'. The search bar contains 'entecavir'. Below the search bar are buttons for 'Column visibility', 'Copy', 'CSV', and 'Show 15 entries'. To the right is a 'Filter' input field. On the far right, there are 'Previous 1' and 'Next' buttons.

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
1711248	485434	entecavir 0.5 MG Oral Tablet	Clinical Drug	0	33,979	Drug	RxNorm	
1711250	485436	entecavir 1 MG Oral Tablet	Clinical Drug	19	5,640	Drug	RxNorm	
1711272	485536	entecavir 0.05 MG/ML Oral Solution	Clinical Drug	0	0	Drug	RxNorm	

On the left side, there is a sidebar with a tree view of vocabularies and classes. The 'RxNorm (22)' node is expanded, showing three entries. The 'Clinical Drug (9)' node under 'Class' is also expanded, showing three entries. The 'Drug (820)' node under 'Domain' is expanded, showing three entries. The 'Standard (519)' node under 'Standard Concept' is expanded, showing three entries.

The screenshot shows a 'Concept Set' creation interface. The title bar says 'Concept Set' and 'Entecavir 1MG'. Below the title are buttons for 'Optimize' and 'Delete'. The main area shows a table with one entry:

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
1711250	485436	entecavir 1 MG Oral Tablet	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

At the bottom, there are color-coded legends: purple for 'Classification', dark red for 'Non-Standard', and blue for 'Standard'.

#### 4) Tenocavir Concept set 생성

Screenshot of a search interface showing results for "tenofovir".

**Search Bar:** tenofovir

**Buttons:** Search, Import

**Table Headers:** Column visibility, Copy, CSV, Show 15 entries, Advanced Options, Filter.

**Table Data:**

Vocabulary	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
RxNorm (129)	1710312	349251	Tenofovir disoproxil fumarate 300 MG Oral Tablet	Clinical Drug	13,035	13,035	Drug	RxNorm
SNOMED (83)	1703093	476556	emtricitabine 200 MG / Tenofovir disoproxil fumarate 300 MG Oral Tablet	Clinical Drug	387	387	Drug	RxNorm
dm+d (28)	35605552	1721613	cobicistat 150 MG / elvitegravir 150 MG / emtricitabine 200 MG / tenofovir alafenamide 10 MG Oral Tablet	Clinical Drug	0	118	Drug	RxNorm
VA Product (22)	1718503	1858261	tenofovir alafenamide 25 MG Oral Tablet	Clinical Drug	0	110	Drug	RxNorm
Branded Drug Form (52)	42708096	1243227	Tenofovir disoproxil fumarate 250 MG Oral Tablet	Clinical Drug	0	60	Drug	RxNorm
Clinical Drug (48)	42708120	1243343	Tenofovir disoproxil fumarate 0.04 MG/MG Oral Powder	Clinical Drug	0	0	Drug	RxNorm
Branded Drug Box (47)	42708114	1243324	Tenofovir disoproxil fumarate 150 MG Oral Tablet	Clinical Drug	0	0	Drug	RxNorm
Clinical Drug Box (45)	42708116	1243328	Tenofovir disoproxil fumarate 200 MG Oral Tablet	Clinical	0	0	Drug	RxNorm
Clinical Drug Form (31)								

Screenshot of a "Concept Set" interface showing results for "tenofovir".

**Search Bar:** tenofovir

**Buttons:** Optimize, Save

**Table Headers:** Concept Set Expression, Included Concepts (333), Included Source Codes, Explore Evidence, Export, Compare.

**Table Data:**

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
1710312	349251	Tenofovir disoproxil fumarate 300 MG Oral Tablet	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
42708096	1243227	Tenofovir disoproxil fumarate 250 MG Oral Tablet	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
1718503	1858261	tenofovir alafenamide 25 MG Oral Tablet	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
1703093	476556	emtricitabine 200 MG / Tenofovir disoproxil fumarate 300 MG Oral Tablet	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
35605552	1721613	cobicistat 150 MG / elvitegravir 150 MG / emtricitabine 200 MG / tenofovir alafenamide 10 MG Oral Tablet	Drug	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Legend: Classification (purple), Non-Standard (red), Standard (blue).

## 5) Liver transplant(Procedure) Concept set 생성

; RC, DRC가 모두 0이므로 해당 코드로 생성된 레코드가 없음을 확인할 수 있음.

liver transplant																																																																																																																																																																																												
Search																																																																																																																																																																																												
Search		Import																																																																																																																																																																																										
liver transplant																																																																																																																																																																																												
<input type="button" value="Column visibility"/> <input type="button" value="Copy"/> <input type="button" value="CSV"/> Show 15 entries <input type="text" value="Filter:"/> <input type="button" value="Advanced Options"/>																																																																																																																																																																																												
Showing 1 to 9 of 9 entries																																																																																																																																																																																												
<table border="1"> <thead> <tr> <th>Vocabulary</th> <th>ID</th> <th>Code</th> <th>Name</th> <th>Class</th> <th>RC</th> <th>DRC</th> <th>Domain</th> <th>Vocabulary</th> </tr> </thead> <tbody> <tr> <td>SNOMED (35)</td> <td>4233677</td> <td>90651007</td> <td>Anesthesia for liver transplant, recipient</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Read (9)</td> <td>44808390</td> <td>858441000000104</td> <td>Discussion about liver transplantation</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>ICD10CM (8)</td> <td>4125319</td> <td>235458006</td> <td>Exploration of liver transplant</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>CIEL (7)</td> <td>4067459</td> <td>174426002</td> <td>Heterotopic liver transplant</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>LONIC (6)</td> <td>4099299</td> <td>27280000</td> <td>Liver transplant with recipient hepatectomy</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Clinical Finding (13)</td> <td>4101838</td> <td>28009009</td> <td>Liver transplant without recipient hepatectomy</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Procedure (12)</td> <td>44809584</td> <td>853761000000103</td> <td>Living donor liver transplantation</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Read (9)</td> <td>4067458</td> <td>174425003</td> <td>Orthotopic liver transplant</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Context-dependent (8)</td> <td>4068633</td> <td>174427006</td> <td>Replacement of previous liver transplant</td> <td>Procedure</td> <td>0</td> <td>0</td> <td>Procedure</td> <td>SNOMED</td> </tr> <tr> <td>Diagnosis (7)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Condition (31)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Procedure (22)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Observation (18)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Measurement (6)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Provider Specificity (2)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Standard Concept</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Non-Standard (38)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Standard (38)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Classification (4)</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>									Vocabulary	ID	Code	Name	Class	RC	DRC	Domain	Vocabulary	SNOMED (35)	4233677	90651007	Anesthesia for liver transplant, recipient	Procedure	0	0	Procedure	SNOMED	Read (9)	44808390	858441000000104	Discussion about liver transplantation	Procedure	0	0	Procedure	SNOMED	ICD10CM (8)	4125319	235458006	Exploration of liver transplant	Procedure	0	0	Procedure	SNOMED	CIEL (7)	4067459	174426002	Heterotopic liver transplant	Procedure	0	0	Procedure	SNOMED	LONIC (6)	4099299	27280000	Liver transplant with recipient hepatectomy	Procedure	0	0	Procedure	SNOMED	Clinical Finding (13)	4101838	28009009	Liver transplant without recipient hepatectomy	Procedure	0	0	Procedure	SNOMED	Procedure (12)	44809584	853761000000103	Living donor liver transplantation	Procedure	0	0	Procedure	SNOMED	Read (9)	4067458	174425003	Orthotopic liver transplant	Procedure	0	0	Procedure	SNOMED	Context-dependent (8)	4068633	174427006	Replacement of previous liver transplant	Procedure	0	0	Procedure	SNOMED	Diagnosis (7)									Condition (31)									Procedure (22)									Observation (18)									Measurement (6)									Provider Specificity (2)									Standard Concept									Non-Standard (38)									Standard (38)									Classification (4)								
Vocabulary	ID	Code	Name	Class	RC	DRC	Domain	Vocabulary																																																																																																																																																																																				
SNOMED (35)	4233677	90651007	Anesthesia for liver transplant, recipient	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Read (9)	44808390	858441000000104	Discussion about liver transplantation	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
ICD10CM (8)	4125319	235458006	Exploration of liver transplant	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
CIEL (7)	4067459	174426002	Heterotopic liver transplant	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
LONIC (6)	4099299	27280000	Liver transplant with recipient hepatectomy	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Clinical Finding (13)	4101838	28009009	Liver transplant without recipient hepatectomy	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Procedure (12)	44809584	853761000000103	Living donor liver transplantation	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Read (9)	4067458	174425003	Orthotopic liver transplant	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Context-dependent (8)	4068633	174427006	Replacement of previous liver transplant	Procedure	0	0	Procedure	SNOMED																																																																																																																																																																																				
Diagnosis (7)																																																																																																																																																																																												
Condition (31)																																																																																																																																																																																												
Procedure (22)																																																																																																																																																																																												
Observation (18)																																																																																																																																																																																												
Measurement (6)																																																																																																																																																																																												
Provider Specificity (2)																																																																																																																																																																																												
Standard Concept																																																																																																																																																																																												
Non-Standard (38)																																																																																																																																																																																												
Standard (38)																																																																																																																																																																																												
Classification (4)																																																																																																																																																																																												

Concept Set																																																
liver transplant																																																
Concept Set Expression		Included Concepts 12		Included Source Codes		Explore Evidence		Export																																								
<input type="button" value="Show 25 entries"/> <input type="text" value="Search:"/> <input type="button" value="Previous 1"/>																																																
Showing 1 to 4 of 4 entries																																																
<table border="1"> <thead> <tr> <th>Concept ID</th> <th>Concept Code</th> <th>Concept Name</th> <th>Domain</th> <th>Standard Concept Caption</th> <th>Exclude</th> <th>Descendants</th> <th>Mapped</th> </tr> </thead> <tbody> <tr> <td>4068633</td> <td>174427006</td> <td>Replacement of previous liver transplant</td> <td>Procedure</td> <td>Standard</td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>4101838</td> <td>28009009</td> <td>Liver transplant without recipient hepatectomy</td> <td>Procedure</td> <td>Standard</td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>4099299</td> <td>27280000</td> <td>Liver transplant with recipient hepatectomy</td> <td>Procedure</td> <td>Standard</td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>4067459</td> <td>174426002</td> <td>Heterotopic liver transplant</td> <td>Procedure</td> <td>Standard</td> <td><input type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> <td><input checked="" type="checkbox"/></td> </tr> </tbody> </table>									Concept ID	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped	4068633	174427006	Replacement of previous liver transplant	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	4101838	28009009	Liver transplant without recipient hepatectomy	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	4099299	27280000	Liver transplant with recipient hepatectomy	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	4067459	174426002	Heterotopic liver transplant	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Concept ID	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped																																									
4068633	174427006	Replacement of previous liver transplant	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>																																									
4101838	28009009	Liver transplant without recipient hepatectomy	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>																																									
4099299	27280000	Liver transplant with recipient hepatectomy	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>																																									
4067459	174426002	Heterotopic liver transplant	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>																																									
<input type="checkbox"/> Classification <input type="checkbox"/> Non-Standard <input type="checkbox"/> Standard																																																

## 6) Liver transplant(Condition) Concept set 생성

; Liver transplant(Procedure)와 같은 의미를 가진 개념이지만 조건을 줄 테이블이 다르므로 별도로 생성함.

The screenshot shows the SNOMED CT interface with the search term 'Liver transplant' entered. The results table displays 10 entries from a total of 10. The columns include Id, Code, Name, Class, RC, DRC, Domain, and Vocabulary. The results are as follows:

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
4332942	432777003	Accelerated rejection of liver transplant	Clinical Finding	0	0	Condition	SNOMED	
4331678	431222008	Acute rejection of liver transplant	Clinical Finding	0	0	Condition	SNOMED	
4198610	432908002	Chronic rejection of liver transplant	Clinical Finding	0	0	Condition	SNOMED	
4161644	432772009	Hyperacute rejection of liver transplant	Clinical Finding	0	0	Condition	SNOMED	
4340951	235910007	Liver transplant disorder	Clinical Finding	0	0	Condition	SNOMED	
4341658	235912004	Liver transplant failure	Clinical Finding	0	0	Condition	SNOMED	
4308395	213153001	Liver transplant failure and rejection	Clinical Finding	0	0	Condition	SNOMED	
42537742	737297006	Liver transplant recipient	Clinical Finding	0	0	Condition	SNOMED	
4341657	235911006	Liver transplant rejection	Clinical Finding	0	0	Condition	SNOMED	
760923	122521000119105	Lymphoproliferative disorder following liver transplant	Clinical Finding	0	0	Condition	SNOMED	

The screenshot shows the configuration of the 'liver transplant\_condition' concept set. The table lists 10 entries from a total of 10. The columns include Concept Id, Concept Code, Concept Name, Domain, Standard Concept Caption, Exclude, Descendants, and Mapped. The results are as follows:

	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
760923	122521000119105	Lymphoproliferative disorder following liver transplant	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4341657	235911006	Liver transplant rejection	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
42537742	737297006	Liver transplant recipient	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4308395	213153001	Liver transplant failure and rejection	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4341658	235912004	Liver transplant failure	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4340951	235910007	Liver transplant disorder	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4161644	432772009	Hyperacute rejection of liver transplant	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4198610	432908002	Chronic rejection of liver transplant	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4331678	431222008	Acute rejection of liver transplant	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4332942	432777003	Accelerated rejection of liver transplant	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	

- Step3 : 코호트 정의 및 생성

- 1) Target cohort

Initial Event Cohort 정의

- 사용할 테이블 결정

▶ 코호트 생성: Cohort Definitions → New Cohort

The screenshot shows the ATLAS Cohort Definitions interface. On the left is a sidebar with 'ATLAS' and links to Home, Data Sources, Search, Concept Sets, and Cohort Definitions. The main area is titled 'Cohort definitions' and shows a table with columns for Id, Name, Created, and Updated. One entry is highlighted: 'New Cohort Definition.phy.xls'. Below this is a large input field containing 's\_entecavir'.

▶ 코호트 이름 정하기: s\_entecavir

The screenshot shows the 'New Cohort' interface. At the top is a title bar with 'New Cohort' and a subtitle 'Patient use entecavir against B-viral hepatitis'. Below this is a navigation bar with tabs: Definition (selected), Concept Sets, Generation, Reporting, Export, and Warnings.

▶ 코호트 정의하기: + Add Initial Event ▾

The screenshot shows a dropdown menu titled '+ Add Initial Event ▾'. It lists several criteria types: 'Add Condition Era Criteria', 'Add Condition Occurrence Criteria', 'Add Death Criteria', 'Add Device Exposure Criteria', 'Add Dose Era Criteria', 'Add Drug Era Criteria', and 'Add First Exposure Criteria'. The 'Add Drug Era Criteria' option is currently selected.

The screenshot shows the configuration interface for 'Drug Era Criteria'. It includes a dropdown menu with 'a drug era of' and 'Any Drug' selected, and a 'Import Concept Set' button below it.

▶ 미리 정의해 놓은 Concept set 가져오기:

→ Import concept set → entecavir (조회) 선택

- 환자별 초기료를 대상으로 함

The screenshot shows a dropdown menu titled '+ Add criteria attribute... ▾'. It includes an option 'Add First Exposure Criteria' with the description 'Limit Drug Eras to first exposure in history.' To the right of this, there is a note: '- ✗ for the first time in the person's history'.

## - 관찰기간 설정

Add Era Start Date Criteria  
Filter Drug Eras by the Era Start Date.

**X era start is:** Between 2004-12-30 and 2014-12-29

## - 연속처방 조건

Add Era Length Criteria  
Filter Drug Eras by the Era duration.

**X with era length** Greater or Equal To 180 days

## - 환자 연령 구간 설정

Add Age at Era Start Criteria  
Filter Drug Eras by age at era start.

**X with age in years at era start** Between 30 and 80

## - washout 설정: Left censoring<sup>19)</sup> 방지

with continuous observation of at least 180 days before and 0 days after event index date

Limit initial events to: earliest event per person.

## - 추가 제약 조건 : 진단, 투약 이력 제한

- B viral hepatitis 진단 有
- Tenofovir 투약 이력 없음
- Lamibudin, Adefovir, Clebusudine ...등 투약 이력 없음

Initial Event Cohort

People having any of the following:

+ Add Initial Event ?

+ Add criteria attribute... Delete Criteria

a condition era of Entecavir

**X era start is:** Between 2004-12-30 and 2014-12-29

**X with era length** Greater or Equal To 180 days

**X with age in years at era start** Between 30 and 80

with continuous observation of at least 180 days before and 0 days after event index date

Limit initial events to: earliest event per person.

Initial event inclusion criteria: From among the initial events, include:

having all of the following criteria:

+ Add criteria to group... ?

+ Add criteria attribute... Delete Criteria

with at least 1 using all occurrences of:

a condition occurrence of B-viral hepatitis

where event starts between All days Before and 7 days After index start date add additional constraint

restrict to the same visit occurrence

and with at most 0 using all occurrences of:

a drug era of tenecavir

where event starts between 1095 days Before and 1 days After index start date add additional constraint

Limit cohort of initial events to: earliest event per person.

Remove initial event inclusion criteria

19) Right censoring: a data point is above a certain value but it is unknown by how much.

Left censoring: a data point is below a certain value but it is unknown by how much.

- 건수 확인 : Save → Generation

Source Name	Generation Status	People	Records	Generated	Generation Duration
▶ Generate	CDMPv1	COMPLETE	593	593 03/20/2019 1:33:48 PM	00:00:48

## ▶ Qualifying criteria 설정

Additional Qualifying Inclusion Criteria
?

New qualifying inclusion criteria
HCV&HDV제외
Copy
Delete

1. HCV&HDV제외

2. HIV제외

3. HCC제외

4. liver transplant 제외

having at most 0 of the following criteria:

with at least 1 using all occurrences of:  
a condition era of gi\_HDV&HCV  
where event starts between All days Before and All days After index start date add additional constraint

+ Add criteria attribute... Delete Criteria

Limit qualifying cohort to: earliest event per person.

Additional Qualifying Inclusion Criteria
?

New qualifying inclusion criteria
HIV제외
Copy
Delete

1. HCV&HDV제외

2. HIV제외

3. HCC제외

4. liver transplant 제외

having at most 0 of the following criteria:

with at least 1 using all occurrences of:  
a condition era of HIV  
where event starts between All days Before and All days After index start date add additional constraint

+ Add criteria attribute... Delete Criteria

Limit qualifying cohort to: earliest event per person.

Additional Qualifying Inclusion Criteria
?

New qualifying inclusion criteria
HCC제외
Copy
Delete

1. HCV&HDV제외

2. HIV제외

3. HCC제외

4. liver transplant 제외

having at most 0 of the following criteria:

with at least 1 using all occurrences of:  
a condition era of gi\_HCC  
where event starts between All days Before and All days After index start date add additional constraint

+ Add criteria attribute... Delete Criteria

Limit qualifying cohort to: earliest event per person.

Additional Qualifying Inclusion Criteria
?

New qualifying inclusion criteria
liver transplant 제외
Copy
Delete

1. HCV&HDV제외

2. HIV제외

3. HCC제외

4. liver transplant 제외

having at most 0 of the following criteria:

with at least 1 using all occurrences of:  
a procedure occurrence of liver\_transplant\_condition  
where event starts between All days Before and All days After index start date add additional constraint  
 restrict to the same visit occurrence

or with at least 1 using all occurrences of:  
a procedure occurrence of liver\_transplant  
where event starts between All days Before and All days After index start date add additional constraint  
 restrict to the same visit occurrence

+ Add criteria attribute... Delete Criteria

Limit qualifying cohort to: earliest event per person.

Additional Qualifying Inclusion Criteria

New qualifying inclusion criteria: Tenofovir 제외

enter an inclusion rule description: having at most 0 of the following criteria:

with at least 1 using all occurrences of: a drug era of gi\_Tenofovir

where event starts between All days Before and All days After index start date add additional constraint

Limit qualifying cohort to: earliest event per person.

- 건수 확인 : Save → Generation

Available CDM Sources

Source Name	Generation Status	People	Records	Generated	Generation Duration
CDMPv1	COMPLETE	340	340	03/20/2019 2:24:40 PM	00:01:21

Inclusion Report for CDMPv1

Inclusion Rule	Summary Statistics:			% Satisfied	% To-Gain
	Match Rate	Matches	Total		
1. HCV&HDV제외	74.89%	450	454	99.12%	0.66%
2. HIV제외		449		98.90%	0.88%
3. HCC제외		449		98.90%	1.10%
4. liver transplant 제외		454		100.00%	0.00%
5. Tenofovir 제외		352		77.53%	22.03%

Population Visualization: 100 people (22.03%), 4 criteria passed, 1 criteria failed.

2) Comparator cohort 생성: Tenofovir 복용 환자(나머지 동일 조건)

- S\_entecavir cohort를 copy(복사)

Cohort #84

s\_entecavir

→ 사본 생성→이름 변경(Tenofovir)→저장

Cohort #100

COPY OF: s\_entecavir

## → Tenofovir vs Entecavir 변경

**Initial Event Cohort**

People having any of the following:

- a drug era of **Entecavir**

**✗** for the first time in the person's history

**✗** era start is: Between **2004-12-30** and **2014-12-29**

**✗** with era length Greater or Equal To **180** days

**✗** with age in years at era start Between **30** and **80**

with continuous observation of at least **180** days before and **0** days after event index date

Limit initial events to: earliest event per person.

**Initial event inclusion criteria:** From among the initial events, include:

having **all** of the following criteria:

- with at least **1** occurrences of:
  - a condition occurrence of **B-viral hepatitis**
  - where **event starts** between **All** days **Before** and **7** days **After**, **index start date** [add additional constraint](#)
  - restrict to the same visit occurrence
- and with at most **0** occurrences of:
  - a drug era of **Tenofovir**
  - where **event starts** between **1095** days **Before** and **1** days **After**, **index start date** [add additional constraint](#)

Limit cohort of initial events to: earliest event per person.

**Remove initial event inclusion criteria**

**Additional Qualifying Inclusion Criteria**

New qualifying inclusion criteria **Entecavir 제외** **Copy** **Delete**

enter an inclusion rule description

having at most **0** of the following criteria:

with at least **1** occurrences of:
 

- a drug era of **gi.Entecavir**
- where **event starts** between **All** days **Before** and **All** days **After**, **index start date** [add additional constraint](#)

Limit qualifying cohort to: earliest event per person.

- 건수 확인 : Save → Generation

Available CDM Sources						
Source Name	Generation Status	People	Records	Generated	Generation Duration	
<b>Generate</b> CDMPv1	COMPLETE	134	134	03/20/2019 2:38:42 PM	00:01:08	<b>View Reports</b>
<b>Inclusion Report</b> <b>Cohort Features</b>						<b>By Events</b> <b>By Person</b>
<b>Inclusion Report for CDMPv1</b>						
Inclusion Rule		Summary Statistics:	Match Rate	Matches	Total	Population Visualization
1. HCV&HDV제외			57.02%	N	% Satisfied	% To Gain
2. HIV제외				234	99.57%	0.00%
3. HCC제외				234	99.57%	0.43%
4. liver_transplant 제외				234	99.57%	0.43%
5. Entecavir 제외				235	100.00%	0.00%
				136	57.87%	41.70%
						Switch to attrition view 98 people (41.70%), 4 criteria passed, 1 criteria failed.

- Text view: Export → Text View

; ATLAS를 이용해 정의한 코호트를 종합적으로 가독성 있는 텍스트로 보여줌.

s\_entecavir

Definition Concept Sets Generation Reporting Export Warnings 12

Text View Graphical View JSON SQL

**Copy To Clipboard**

**s\_entecavir**

Patient use entecavir against b-viral hepatitis

**Initial Event Cohort**

People having any of the following:

- a drug era of Entecavir<sup>3</sup>
  - for the first time in the person's history
  - era start is between 2004-12-30 and 2014-12-29 (inclusive)
  - with era length  $\geq 180$
  - with age in years at era start between 30 and 80 (inclusive)

with continuous observation of at least 180 days prior and 0 days after event index date, and limit initial events to: **earliest event per person**.

For people matching the Primary Events, include:

Having all of the following criteria:

- at least 1 occurrences of a condition occurrence of B-viral hepatitis<sup>2</sup>  
where event starts between all days Before and 7 days After index start date
- and at most 0 occurrences of a drug era of tenofovir<sup>10</sup>  
where event starts between 1095 days Before and 1 days After index start date

Limit cohort of initial events to: **earliest event per person**.

.....

Appendix 1: Concept Set Definitions

1. acute hepatitis

Concept Id	Concept Name	Domain	Vocabulary	Excluded	Descendants	Mapped
4243475	Acute hepatitis	Condition	SNOMED	NO	YES	YES
4055207	Acute hepatitis - non-infective	Condition	SNOMED	NO	YES	YES
439675	Acute hepatitis B with delta agent (coinfection) with hepatic coma	Condition	SNOMED	NO	YES	YES
439673	Acute hepatitis B with delta-agent (coinfection) without hepatic coma	Condition	SNOMED	NO	YES	YES
4341652	Acute hepatitis B with hepatitis D	Condition	SNOMED	NO	YES	YES
192242	Acute hepatitis C	Condition	SNOMED	NO	YES	YES
197490	Acute hepatitis E	Condition	SNOMED	NO	YES	YES
4314443	Hepatic coma due to acute hepatitis B	Condition	SNOMED	NO	YES	YES
46270037	Hepatic coma due to acute hepatitis C	Condition	SNOMED	NO	YES	YES
4059297	Toxic liver disease with acute hepatitis	Condition	SNOMED	NO	YES	YES

## - Graphic view

; ATLAS를 이용해 정의한 코호트를 종합적인 그래픽으로 보여줌.



## [ 코호트 생성 예제 2 ] : H.pylori and Stomach ca.

### ◦ Step1. 코호트 정의

- 코호트 개요: 2010년 이전에 위내시경 실시하여 아래 질병명이 있는 환자 중 2011년 이후 stomach cancer가 진단된 사람
- 추출 대상자
  - 2010년 이전에 위내시경 실시한 사람 전체
  - 환자들 중에서 아래 질병의 진단명이 있는 사람

#### ※ 질병 분류코드

- Helicobacter pylori 양성 (B980): C4650, B4151 ~ B4157
  - Atrophic gastritis (K294)
  - Peptic ulcer (K27)
  - 위궤양 (K25)
  - 십이지장궤양 (K26)
  - 결과: 위암 진단이 있는 환자
- \* 2011년 이후 - 현재 : 위암 발생자

Input parameter	Design choice	
	Initial	Qualifying(Exclusion)
Target cohort(T)	2010년 이전 위내시경 + 진단(+) 시점?	. 이전 Stomach ca
Comparator cohort(C)	2010년 이전 위내시경 + 진단(-) 시점?	
Outcome cohort(O)	위암	
Time at risk	내시경(진단) 시점	
Model specification	상기 진단이 위암 발생에 미치는 영향	

## ○ Step2. 필요한 Concept set 만들기

- New concept set → Naming : FM\_Hpylori → Search : Helicobacter pylori 검색

Filter · Vocabulary : SNOMED

- Class : Procedure
- Domain : Procedure
- Standard Concept : Standard

※ 자료 존재 여부

- RC(Record Count) : 레코드 존재
- DRC(Descendant Record Count) : 하위 레코드 존재

→ 확인 후 선택(Cart  ) … ※ 필요시 추가 검색 및 선택

Concept Set								
fm_gastroscopy				Optimize				
Concept Set Expression		Included Concepts (259)		Included Source Codes		Explore Evidence		Export
Show 25 ▾ entries								Search: <input type="text"/>
Showing 1 to 7 of 7 entries								Previous 1 Next
Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped	
2108899	43251	Esophagastroduodenoscopy, flexible, transoral; with removal of tumor(s), polyp(s), or other lesion(s) by snare technique	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
2108900	43255	Esophagastroduodenoscopy, flexible, transoral; with control of bleeding, any method	Procedure	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
2108892	43244	Esophagastroduodenoscopy, flexible, transoral; with band ligation of esophageal/gastric varices	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
45889086	1007260	Esophagastroduodenoscopy, flexible, transoral	Procedure	Classification	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
45889592	1021431	Esophagastroduodenoscopy Procedures	Procedure	Classification	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4294382	76009000	Esophagastroduodenoscopy	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
4297516	386831001	Endoscopy of stomach	Procedure	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	

 Classification  Non-Standard  Standard

- 같은 방법으로 Cohort definition에 필요한 Concept set을 모두 생성

Concept Sets					
New Concept Set					
Show 10 ▾ entries Filter Repository Concept Sets: fm					
ID	Title	Created	Modified	Author	
191	fm_stomach ca	04/14/2019 8:48:40 PM	04/14/2019 8:48:40 PM	anonymous	
190	fm_Hpylori	04/14/2019 8:41:38 PM	04/14/2019 8:42:02 PM	anonymous	
188	fm_ulcer	04/14/2019 8:34:16 PM	04/14/2019 8:39:21 PM	anonymous	
189	fm_atrophic gastritis	04/14/2019 8:35:09 PM	04/14/2019 8:35:20 PM	anonymous	
186	fm_gastroscopy	04/14/2019 8:19:51 PM	04/14/2019 8:24:23 PM	anonymous	

Showing 1 to 5 of 5 entries (filtered from 170 total entries)

Previous 1 Next

Export Mode

### ○ Step3. Target Cohort definition

- New cohort → Initial event 정의

**Initial Event Cohort**

People having any of the following:

- a procedure occurrence of **fm\_gastroscopy**
- occurrence start is: Before 2010-12-31

with continuous observation of at least 0 days before and 0 days after event index date

Limit initial events to: earliest event per person.

**Add initial event inclusion criteria**

- 건수 확인 : Save → Generation

Available CDM Sources						
Source Name	Generation Status	People	Records	Generated	Generation Duration	
► Generate cdmpv1	COMPLETE	111,822	111,822	04/14/2019 8:56:03 PM	00:03:13	

- Add initial event inclusion criteria

**Initial event inclusion criteria:** From among the initial events, include:

having any of the following criteria:

with at least 1 using all occurrences of:

a condition era of **fm\_Hpypori**

where event starts between 0 days Before and 30 days After index start date [add additional constraint](#)

or with at least 1 using all occurrences of:

a condition era of **fm\_atrophic gastritis**

where event starts between 0 days Before and 30 days After index start date [add additional constraint](#)

or with at least 1 using all occurrences of:

a condition era of **fm\_ulcer**

where event starts between 0 days Before and 30 days After index start date [add additional constraint](#)

Limit cohort of initial events to: earliest event per person.

**Remove initial event inclusion criteria**

- 건수 확인 : Save → Generation

Available CDM Sources						
Source Name	Generation Status	People	Records	Generated	Generation Duration	
► Generate cdmpv1	COMPLETE	8,885	8,885	04/14/2019 9:08:17 PM	00:03:19	<a href="#">View Report</a>
<a href="#">Inclusion Report</a>						
<a href="#">Cohort Features</a>						
						<a href="#">By Events</a> <a href="#">By Person</a>

Inclusion Report for cdmpv1			Match Rate	Matches	Total	Population Visualization	Switch to attrition vi
Summary Statistics:	100.00%		8,885	8,885	8,885		
Inclusion Rule	N	% Satisfied		% To-Gain			

- Additional Qualifying criteria 정의
  - 이전 Stomach ca 진단 환자 제외

Additional Qualifying Inclusion Criteria

New qualifying inclusion criteria  Copy Delete

1. 이전stomach ca제외

having   of the following criteria: + Add criteria to group...

with   using all occurrences of: Delete Criteria

a condition era of  + Add criteria attribute...

where  between  days  and  days   add additional constraint

Limit qualifying cohort to:  per person.

- 건수 확인 : Save → Generation

Available CDM Sources

Source Name	Generation Status	People	Records	Generated	Generation Duration
<input type="button" value="Generate"/> cdmpv1	COMPLETE	8,842	8,842	04/14/2019 9:33:45 PM	00:03:19

By Events

Inclusion Report for cdmpv1

Inclusion Rule	Summary Statistics:		Match Rate	Matches	Total
	N	% Satisfied		% To-Gain	
1. 이전stomach ca제외	8,842	99.52%		0.48%	

Population Visualization

Switch

## ○ Step4. Comparator Cohort definition

- Copy(  ) → Renaming → 조건 변경 → Save

**Initial Event Cohort**

People having any of the following:

a procedure occurrence of **fm\_gastroscopy** 

occurrence start is: **Before**  2010-12-31 

+ Add Initial Event 

with continuous observation of at least **0** days before and **0** days after event index date

Limit initial events to: **earliest event** per person.

**Initial event inclusion criteria:** From among the initial events, include:

having **at most** **0** of the following criteria:

+ Add criteria to group... 

with **at least** **1** **using all** occurrences of:

a condition era of **fm\_Hpypori** 

where **event starts** between **0** days **Before** and **30** days **After** **index start date** **add additional constraint** 

or with **at least** **1** **using all** occurrences of:

a condition era of **fm\_atrophic gastritis** 

where **event starts** between **0** days **Before** and **30** days **After** **index start date** **add additional constraint** 

or with **at least** **1** **using all** occurrences of:

a condition era of **fm\_ulcer** 

where **event starts** between **0** days **Before** and **30** days **After** **index start date** **add additional constraint** 

Limit cohort of initial events to: **earliest event** per person.

**Remove initial event inclusion criteria**

**Additional Qualifying Inclusion Criteria**

New qualifying inclusion criteria  Please select a qualifying inclusion criteria to edit.

1. 이전stomach ca제외

Limit qualifying cohort to: **earliest event** per person.

- 건수 확인 : Save → Generation

### Available CDM Sources

	Source Name	Generation Status	People	Records	Generated	Generation Duration
 	cdmpv1	COMPLETE	101,967	101,967	04/14/2019 9:42:11 PM	00:03:21
 						By Event

### Inclusion Report for cdmpv1

	Match Rate	Matches	Total
Summary Statistics:	99.06%	101,967	102,937
<b>Inclusion Rule</b>	<b>N</b>	<b>% Satisfied</b>	<b>% To-Gain</b>
1. 이전stomach ca제외	101,967	99.06%	0.94%

### Population Visualization

- Step5. Outcome Cohort definition
  - Stomach cancer

**Initial Event Cohort**

People having any of the following:

+ Add Initial Event ▾

+ Add criteria attribute... ▾

Delete Criteria

a condition era of fm\_stomach ca ▾

✗ for the first time in the person's history

✗ era start is: On or After ▾ 2010-12-31

with continuous observation of at least 0 ▾ days before and 0 ▾ days after event index date

Limit initial events to: earliest event ▾ per person.

Add initial event inclusion criteria

- Step6. Estimation

**Population Level Effect Estimation**

fm\_stomach ca

Specification Utilities

Choose your target cohort:  
fm\_EGD

Choose your comparator cohort:  
fm\_EGD\_comparator

Choose your outcome cohort:  
fm\_stomach ca

Specify the statistical model used to estimate the risk of outcome between target and comparator cohorts:  
Cox proportional hazards ▾

Define the time-at-risk window start, relative to target/comparator cohort entry:  
1 ▾ days from cohort start date

Define the time-at-risk window end:  
0 ▾ days from cohort end date ▾

Minimum washout period applied to target and comparator cohorts:  
0 ▾

Minimum required days at risk, applied to target and comparator cohorts:  
0 ▾

Remove patients who enter both cohorts?  
Yes ▾

Remove patients who have observed the outcome prior to cohort entry?  
Yes ▾

### [ 코호트 생성 예제 3 ] : Lumbar fusion and Reoperation

- Step1. 코호트 설계

- 2012년부터 A 집도의 lumbar fusion 환자(lumbar fusion 중에서 Middle lumbar interbody fusion)
- 그중에서 L-spine CT 처방 코드 GRT851, L-spine X-ray 처방 코드 GRG85000, GRG86000, GRG85002, GRG85013, GRG85001 L-spine MRI GRM816) ; 이 중에 하나라도 수술하고 2년 뒤에 시행한 환자
- 성별, 나이, 당뇨, 고혈압, BMD, BMI, 수술 날짜, 수술하고 검사 시행한 날짜, 재수술 여부

Input parameter	Design choice	
	Initial	Qualifying(Exclusion)
Target cohort(T)	lumbar fusion (Middle lumbar interbody fusion) + 조건 A	<ul style="list-style-type: none"> <li>· 2년 뒤 검사 시행</li> <li>- L spine CT(GRT851)</li> <li>- L spine X-ray(GRG85000, GRG86000 ,GRG85002, GRG85013, GRG85001)</li> <li>- L spine MRI(GRM816)</li> </ul>
Comparator cohort(C)	lumbar fusion (Middle lumbar interbody fusion) + 조건 B	
Outcome cohort(O)	재수술	
Time at risk	수술	
Model specification	조건 A/B 가 재수술에 미치는 영향	

## ○ Step2. 필요한 Concept set 만들기

### Concept set

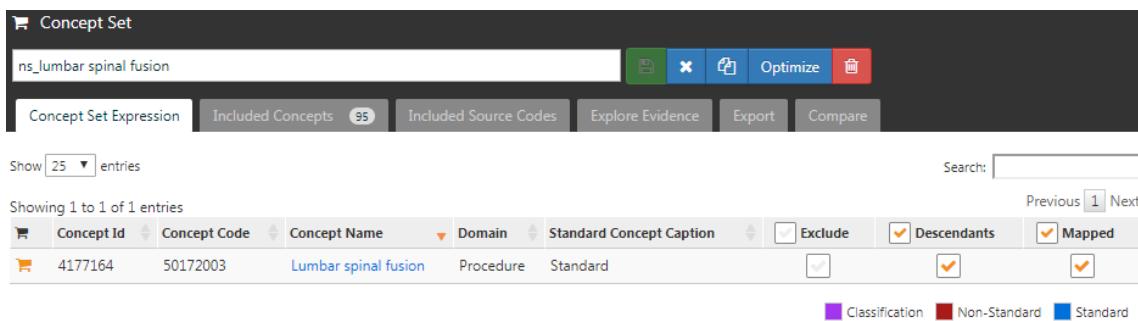
- Cohort Definition을 위해 필요한 개념에 대한 코드 목록
  - 만들어 놓은 Concept set은 다른 코호트 정의 시 사용이 가능함.
  - 단, Cohort Definition 내에서 만들 경우 해당 코호트에서만 사용 가능

- New concept set → Naming : NS\_Fusion → Search : Fusion 검색

→ Filter    · Vocabulary : SNOMED

- Class : Procedure
- Domain : Procedure
- Standard Concept : Standard

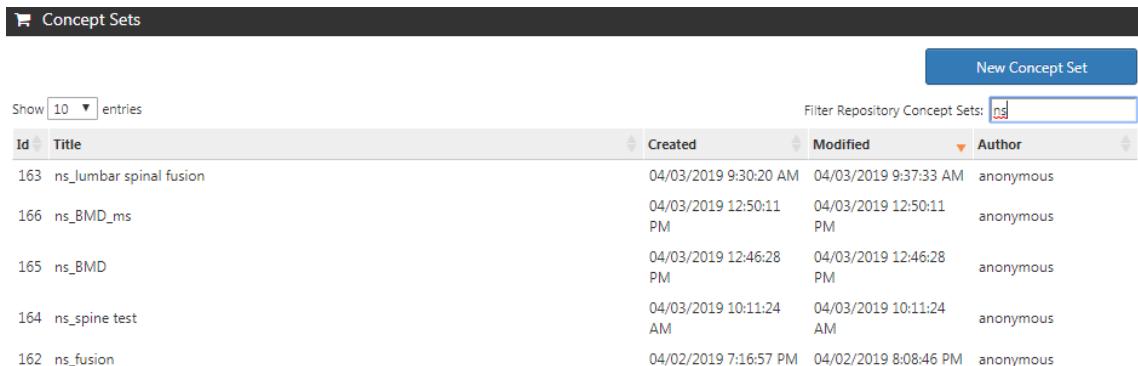
→ 확인 후 선택(Cart  ) … ※ 필요시 추가 검색 및 선택



Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
4177164	50172003	Lumbar spinal fusion	Procedure	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Classification Non-Standard Standard

- 같은 방법으로 Cohort definition에 필요한 Concept set을 모두 생성



ID	Title	Created	Modified	Author
163	ns_lumbar spinal fusion	04/03/2019 9:30:20 AM	04/03/2019 9:37:33 AM	anonymous
166	ns_BMD_ms	04/03/2019 12:50:11 PM	04/03/2019 12:50:11 PM	anonymous
165	ns_BMD	04/03/2019 12:46:28 PM	04/03/2019 12:46:28 PM	anonymous
164	ns_spine test	04/03/2019 10:11:24 AM	04/03/2019 10:11:24 AM	anonymous
162	ns_fusion	04/02/2019 7:16:57 PM	04/02/2019 8:08:46 PM	anonymous

### ○ Step3. Target Cohort definition

- New cohort → Initial event 정의

ns\_0403

Definition Concept Sets Generation Reporting Export Warnings

enter a cohort definition description here

**Initial Event Cohort**

People having any of the following:

a procedure occurrence of ns\_lumbar spinal fusion

occurrence start is: Between 2011-12-31 and 2016-12-30

+ Add Initial Event + Add criteria attribute... Delete Criteria

with continuous observation of at least 0 days before and 0 days after event index date

Limit initial events to: earliest event per person.

**Initial event inclusion criteria:** From among the initial events, include:

having any of the following criteria:

+ Add criteria to group...

Limit cohort of initial events to: earliest event per person.

Remove initial event inclusion criteria

- 건수 확인 : Save → Generation(전체: 1,213, 기간[‘12-‘16]:568)

Available CDM Sources						
Source Name	Generation Status	People	Records	Generated	Generation Duration	
► Generate cdmpv1	COMPLETE	568	568	04/03/2019 1:45:23 PM	00:04:25	

- Additional Qualifying criteria 정의

- 수술 후 2년이상 경과 후 영상 검사 시행

New qualifying inclusion criteria Lumbar spine test Copy Delete

**1. Lumbar spine test** enter an inclusion rule description

2. Reop having at least 1 of the following criteria:

+ Add criteria to group...

with at least 1 using all occurrences of:

a procedure occurrence of ns\_spine test

where event starts between 0 days Before and 730 days After index start date

[add additional constraint](#)

restrict to the same visit occurrence

Limit qualifying cohort to: earliest event per person.

· 수술 후 2년이상 경과 후 재수술

Additional Qualifying Inclusion Criteria

New qualifying inclusion criteria

Reop

Copy Delete

1. Lumbar spine test

2. Reop

enter an inclusion rule description

having at least 1 of the following criteria:

+ Add criteria to group... Delete Criteria

with at least 1 using all occurrences of:

a procedure occurrence of ns\_lumbar spinal fusion + Add criteria attribute... Add additional constraint

where event starts between 0 days Before and 730 days After index start date

restrict to the same visit occurrence

Limit qualifying cohort to: earliest event per person.

- 건수 확인 : Save → Generation

Definition Concept Sets Generation Reporting Export Warnings 4

Available CDM Sources

Source Name	Generation Status	People	Records	Generated	Generation Duration
cdmpv1	COMPLETE	33	33	04/03/2019 2:36:52 PM	00:05:21

Inclusion Report Cohort Features

By Events By Person

Inclusion Report for cdmpv1

Inclusion Rule	Summary Statistics:			Population Visualization			Switch to attric
	Match Rate	Matches	Total	N	% Satisfied	% To-Gain	
1. Lumbar spine test	10.06%	33	328	33	10.06%	89.94%	
2. Reop				328	100.00%	0.00%	

## [ 코호트 생성 예제 4 ] : Warfarin vs NOAC against AF

- Step1. 코호트 설계

1. Valvular replacement를 하지 않은 Atrial fibrillation 환자에게 있어 warfarin 사용 시와 NOAC 사용간의 의료이용과 의료비의 차이  
 ↗ (가설) warfarin 사용 시에는 추가 검사와 외래 내원 빈도가 높을 것이다. 반면에 NOAC은 추가 검사가 없기 때문에, 장기간 약물 처방을 할 것이다.
2. Valvular replacement를 하지 않은 Atrial fibrillation 환자에게 있어 warfarin 사용 시와 NOAC 사용 시 합병증과 뇌졸중 발생의 차이  
 ↗ (가설) warfarin 사용 시 합병증의 발생 (GI bleeding, ICH) 은 높으나 NOAC 사용 시 Stroke의 위험은 동일하거나, 적을 것이다.

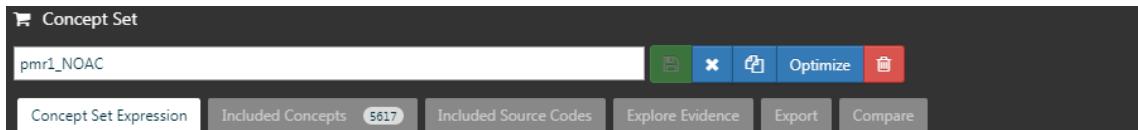
Input parameter	Design choice	
	Initial	Qualifying(Exclusion)
Target cohort(T)	Atrial fibrillation + warfarin	· Valvular replacement 환자 제외
Comparator cohort(C)	Atrial fibrillation + NOAC	
Outcome cohort(O)	의료이용 의료비 합병증의 발생 (GI bleeding, ICH)	
Time at risk	warfarin/NOAC 복용	
Model specification	warfarin 사용 시와 NOAC 사용이 의료이용/의료비에 미치는 영향	

- Step2. 필요한 Concept set 만들기

### Concept set

- Cohort Definition을 위해 필요한 개념에 대한 코드 목록
- 만들어 놓은 Concept set은 다른 코호트 정의 시 사용이 가능함.
- Cohort Definition 내에서 만들 경우 해당 코호트에서만 사용 가능

- New concept set → Naming : gi\_HBV → Search : NOAC 검색
    - ※ NOAC : dabigatran etexilate, apixaban, rivaroxaban, edoxaban
    - option Option · Vocabulary : RxNorm
      - Class : ingredient
      - Domain : Drug
      - Standard Concept : Standard
- 확인 후 선택(Cart  ) … ※ 필요시 추가 검색 및 선택



The screenshot shows the 'Concept Set' interface. At the top, there is a search bar containing 'pmr1\_NOAC'. Below the search bar are several buttons: 'Optimize' and 'Delete'. Underneath the search bar are tabs: 'Concept Set Expression' (selected), 'Included Concepts' (5617), 'Included Source Codes', 'Explore Evidence', 'Export', and 'Compare'. A dropdown menu 'Show 25 entries' is open. A search bar 'Search:' is also present. The main area displays a table of 12 entries, each with a delete icon, Concept ID, Concept Code, Concept Name, Domain, Standard Concept Caption, and checkboxes for Exclude, Descendants, and Mapped.

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
21600962	B01AA	Vitamin K antagonists	Drug	Classification	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
40244449	1232088	rivaroxaban 20 MG Oral Tablet [Xarelto]	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
40244445	1232084	rivaroxaban 15 MG Oral Tablet [Xarelto]	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
40241334	1114202	rivaroxaban 10 MG Oral Tablet [Xarelto]	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
40241331	1114195	rivaroxaban	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
21041627	OMOP344236	edoxaban 60 MG Oral Tablet [Lixiana]	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
21041628	OMOP344243	edoxaban 30 MG Oral Tablet [Lixiana]	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
45892847	1599538	edoxaban	Drug	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

- 같은 방법으로 Cohort definition에 필요한 Concept set을 모두 생성

Id	Title	Created	Modified	Author	
				Author	Author
260	pmr1_NOAC	04/20/2019 9:32:04 PM	04/23/2019 7:45:39 PM	anonymous	anonymous
259	pmr1_warfarin	04/20/2019 9:26:14 PM	04/23/2019 4:35:17 PM	anonymous	anonymous
261	pmr1_AF	04/20/2019 9:33:25 PM	04/23/2019 4:34:49 PM	anonymous	anonymous
262	pmr1_VR_procedure	04/20/2019 9:38:55 PM	04/23/2019 4:34:37 PM	anonymous	anonymous

- Step3. Target Cohort definition

- New cohort → Initial event 정의

pmr1\_warfarin

Definition Concept Sets Generation Reporting Export Warnings 4

enter a cohort definition description here

### Initial Event Cohort

People having any of the following:

+ Add Initial Event ▾

a condition era of pmr1\_AF ▾ + Add criteria attribute... ▾ Delete Criteria

with continuous observation of at least 0 days before and 0 days after event index date

Limit initial events to: earliest event per person.

**Initial event inclusion criteria:** From among the initial events, include:

having all of the following criteria:

+ Add criteria to group... ▾

with at least 1 using all occurrences of:

+ Add criteria attribute... ▾ Delete Criteria

a drug era of pmr1\_warfarin ▾

where event starts between All days Before and All days After index start date add additional constraint

Limit cohort of initial events to: earliest event per person.

Remove initial event inclusion criteria

- 건수 확인 : Save → Generation
- AF 진단

Available CDM Sources

Source Name	Generation Status	People	Records	Generated	Generation Duration
► Generate ▾ cdmpv1	COMPLETE	6,280	6,280	04/23/2019 7:48:37 PM	00:00:03

· & Warfarin 복용

Available CDM Sources

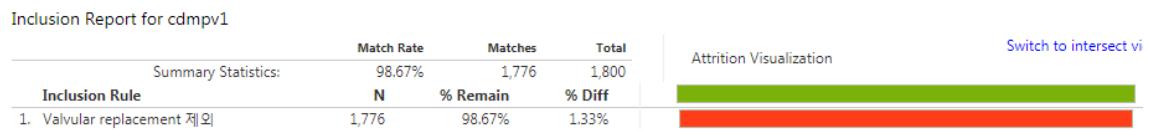
Source Name	Generation Status	People	Records	Generated	Generation Duration
► Generate ▾ cdmpv1	COMPLETE	1,800	1,800	04/23/2019 7:51:38 PM	00:00:06

- Additional Qualifying criteria 정의

The screenshot shows the 'Additional Qualifying Inclusion Criteria' interface. A rule named 'Valvular replacement 제외' is listed. The rule description is 'enter an inclusion rule description' and specifies 'having at most 0 of the following criteria:'. It includes a procedure occurrence of 'pmr1\_VR\_procedure' and a constraint where 'event starts' between 'All days Before' and 'All days After' the 'index start date'. There is also an option to 'restrict to the same visit occurrence'. A note at the bottom states 'Limit qualifying cohort to: earliest event per person.'

- 건수 확인 : Save → Generation

Available CDM Sources						
	Source Name	Generation Status	People	Records	Generated	Generation Duration
Generate	cdmpv1	COMPLETE	1,776	1,776	04/23/2019 8:02:27 PM	00:03:14
<a href="#">Inclusion Report</a> <a href="#">Cohort Features</a>						
						<a href="#">By Events</a> <a href="#">By Person</a>



- Step4. Comparator Cohort definition

- Copy(복사) → Renaming → 조건 변경 → Save

**Initial Event Cohort**

People having any of the following:

a condition era of **pmr1\_AF**

with continuous observation of at least **0** days before and **0** days after event index date

Limit initial events to: **earliest event** per person.

**Initial event inclusion criteria:** From among the initial events, include:

having **all** of the following criteria:

with **at least 1 using all** occurrences of:  
 a drug era of **pmr1\_NOAC**

where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)

Limit cohort of initial events to: **earliest event** per person.

**Remove initial event inclusion criteria**

**Additional Qualifying Inclusion Criteria**

**New qualifying inclusion criteria** **Valvular replacement 제외**

**1. Valvular replacement 제외** enter an inclusion rule description

having **at most 0** of the following criteria:

with **at least 1 using all** occurrences of:  
 a procedure occurrence of **pmr1\_VR\_procedure**

where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)

restrict to the same visit occurrence

Limit qualifying cohort to: **earliest event** per person.

- 건수 확인 : Save → Generation

Available CDM Sources						
Source Name	Generation Status	People	Records	Generated	Generation Duration	
<b>Generate</b> cdmpv1	COMPLETE	2,316	2,316	04/23/2019 8:21:47 PM	00:04:10	<input type="button" value="View Reports"/>
<b>Inclusion Report</b> <b>Cohort Features</b>						
						<input type="button" value="By Events"/> <input type="button" value="By Person"/>

Inclusion Report for cdmpv1						
Inclusion Rule	Match Rate		Matches		Total	
	Summary Statistics:		98.97%	2,316	2,340	Population Visualization
	N	% Satisfied		% To-Gain		<a href="#">Switch to attrition view</a>
1. Valvular replacement 제외	2,316	98.97%		1.03%		

- Step5. Outcome Cohort definition

**Initial Event Cohort**

People having any of the following:

a visit occurrence of **pmr1\_OPD visit**

with continuous observation of at least **0** days before and **0** days after event index date

Limit initial events to: **earliest event** per person.

**Add initial event inclusion criteria**

## [ 코호트 생성 예제 5 ] : Type II DM vs sight-threatening DR

### ○ Step1. 코호트 설계

- : Type II DM 환자에서 sight-threatening DR의 유병률 및 치료 행태
- Type II DM : E11-14 상병명(주진단명, 부진단명 포함(1-4))으로 당뇨병 용제(리스트 별첨)를 연간 1회 이상 처방받은 환자
- 환자 중 아래 레이저, 수술, 주사시술을 한번이라도 시행 받은 환자
- EDI codes : S5160 (레이저 범망막 광응고술), S5121 (유리체절제술)
- 제외 : 주진단명 H3531 (V201) exudative macular degeneration
  - H34 retinal vascular occlusion
  - H330 rhegmatogenous retinal detachment
  - H352 proliferative vitreoretinopathy
  - H3570 central serous chorioretinopathy으로 시술받은 환자
- outcome - 심부전으로 인한 사망률과 재입원
- ※ 중복 처방되거나 동시 시행 환자들도 있음.  
 1회 이상 수술 또는 약품처방 건수 있는 경우 포함(동시 시행 또는 여러 약제 병용 투여는 추가 분석이 필요할 것으로 보임.)  
 아바스틴의 경우 항암제로도 많이 처방되는 약품으로 안과 시행건수로 제한하거나 또는 주진단명에 cancer code (C code)가 있는 경우 제외

Input parameter	Design choice	
	Initial	Qualifying (Exclusion)
Target cohort(T)	<ul style="list-style-type: none"> <li>- Type II DM</li> <li>- 당뇨병 용제 연간 1회이상 처방</li> <li>- 레이저, 수술, 주사시술을 한번이라도 시행</li> <li>- 약품청구(안과 처방)           <ul style="list-style-type: none"> <li>· 653600320 (ranibizumab, Lucentis)</li> <li>· 641105780 (afibercept, Eylea)</li> <li>· 645000360 (bevacizumab, Avastin)</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- 주진단명           <ul style="list-style-type: none"> <li>· H3531(V201) exudative macular degeneration</li> <li>· H34 retinal vascular occlusion</li> <li>· H330 rhegmatogenous retinal detachment</li> <li>· H352 proliferative vitreoretinopathy</li> </ul> </li> <li>- H3570 central serous chorioretinopathy으로 시술받은 환자</li> <li>- 아바스틴 복용 &amp; Cancer 환자</li> </ul>
Comparator cohort(C)		
Outcome cohort(O)	<p>심부전으로 인한 사망</p> <p>심부전으로 인한 재입원</p>	

- Step2. 필요한 Concept set 만들기

- New concept set → Naming(DM) → Search(Diabetes)

→ Filter · Vocabulary : SNOMED

- Class : Clinical finding

- Domain : Condition

- Standard Concept : Standard

→ 확인 후 선택(Cart  ) … ※ 필요시 추가 검색 및 선택

Showing 1 to 24 of 24 entries

	Concept Id	Concept Code	Concept Name
	4193704	313436004	Type 2 diabetes mellitus without complication
	201826	44054006	Type 2 diabetes mellitus
	443731	420279001	Renal disorder due to type 2 diabetes mellitus
	4129519	237627000	Pregnancy and type 2 diabetes mellitus
	43531010	609567009	Pre-existing type 2 diabetes mellitus in pregnancy

같은 방법으로 Cohort definition에 필요한 Concept set을 모두 생성한다.

 Id	Title
123	eye_drug_Lucentis_Eylea_Avastin
122	eye_vitrectomy
129	eye_Type2DM
128	eye_HF
127	eye_visit
126	eye_exclusive condition_H3570
125	eye_exclusive condition
124	eye_drug_Avastin
120	eye_DMdrug
118	eye_DR

- Step3. Target Cohort definition
  - New cohort → Initial event 정의

**Initial Event Cohort**

People having any of the following:

a condition era of **eye\_Type2DM**

+ Add Initial Event Delete Criteria

with continuous observation of at least **0** days before and **0** days after event index date

Limit initial events to: earliest event per person.

Initial event inclusion criteria: From among the initial events, include:

having **all** of the following criteria:

+ Add criteria to group... Delete Criteria

with **at least** **1** **using all** occurrences of:

a drug era of **eye\_DMdrug**

where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)

- 당뇨약 복용한 당뇨 진단 환자수 : 40,109명

Source Name	Generation Status	People
<b>Generate</b> CDMPv1	COMPLETE	40,109

- 환자 중 아래 레이저, 수술, 주사시술을 한번이라도 시행 받은 환자

- KEDI codes

S5160 (레이저 범망막 광응고술)

S5121 (유리체 절제술)

and with **at least** **1** **using all** occurrences of:

a procedure occurrence of **eye\_vitrectomy**

+ Add criteria attribute... Delete Criteria

where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)

restrict to the same visit occurrence

Source Name	Generation Status	People
<b>Generate</b> CDMPv1	COMPLETE	263

## - Additional Qualifying Inclusion criteria

Additional Qualifying Inclusion Criteria

New qualifying inclusion criteria

**Excluded condition**

H3531 (V201) exudative macular degeneration, H34 retinal vascular occlusion, H330 rhegmatogenous retinal detachment,H352 proliferative vitreoretinopathy

having [at most ▾] 0 ▾ of the following criteria:

+ Add criteria to group...

**1. Excluded condition**  
H3531 (V201) exudative macular degeneration, H34 retinal vascular occlusion, H330 rhegmatogenous retinal detachment,H352 proliferative vitreoretinopathy

**2. exclusive condition&procedure**  
H3570 central serous chorioretinopathy.오늘  
A/는

with [at least ▾] 1 ▾ [using all] occurrences of:  
a condition era of eye\_exclusive condition ▾

+ Add criteria attribute... ▾

where event starts between All ▾ days Before ▾ and All ▾ days After ▾ index start date add additional constraint

Delete Criteria

Limit qualifying cohort to: earliest event ▾ per person.

vitreoretinopathy

**2. exclusive condition&procedure**  
H3570 central serous chorioretinopathy.오늘  
A/는

with [at least ▾] 1 ▾ [using all] occurrences of:  
a condition era of eye\_exclusive condition\_H3570 ▾

+ Add criteria attribute... ▾

where event starts between All ▾ days Before ▾ and All ▾ days After ▾ index start date add additional constraint

and with [at most ▾] 0 ▾ [using all] occurrences of:  
a procedure occurrence of Any Procedure ▾

+ Add criteria attribute... ▾

where event starts between All ▾ days Before ▾ and All ▾ days After ▾ index start date add additional constraint

restrict to the same visit occurrence

Delete Criteria

Limit qualifying cohort to: all events ▾ per person.

## - 선택된 코호트 : 0명

Available CDM Sources

Source Name	Generation Status	People	Records	Generated	Generation Duration
Generate	CDMPv1	COMPLETE	0	0 03/18/2019 10:25:23 AM	00:01:21

Inclusion Report Cohort Features

Inclusion Report for CDMPv1

Inclusion Rule	Match Rate			Matches			Total		
	Summary Statistics:			N	% Satisfied	% To-Gain			
1. Excluded condition	0.00%			209	79.47%	0.00%			
2. exclusive condition&procedure				0	0.00%	79.47%			

Population Visualization

Switch to attrition view

By Events By Person

### III. Target 코호트 생성 - Another approach

#### 1. Initial Event cohort

Type II DM 환자에서 sight-threatening DR

**Initial Event Cohort**

People having any of the following:

- a procedure occurrence of **eye\_vitrectomy**
- a condition era of **eye\_drug\_Lucentis Eylea Avastin**

with continuous observation of at least **0** days before and **0** days after event index date  
 Limit initial events to: **earliest event** per person.

**Initial event inclusion criteria:** From among the initial events, include:  
 having **all** of the following criteria:

- with **at least** **1** **using all** occurrences of:  
 a condition era of **eye\_Type2DM**
- where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)
- and with **at least** **1** **using all** occurrences of:  
 a drug era of **eye\_DMdrug**
- where **event starts** between **All** days **Before** and **All** days **After** **index start date** [add additional constraint](#)

#### - Vitrectomy, Photocoagulation 환자

Source Name	Generation Status	People
<b>► Generate</b>	COMPLETE	875

- 약품청구 : 653600320 (ranibizumab, Lucentis), 641105780 (afibercept, Eylea), 645000360 (bevacizumab, Avastin)

Source Name	Generation Status	People
<b>► Generate</b>	COMPLETE	875

#### - 당뇨 약제 & Type II 당뇨

Source Name	Generation Status	People
<b>► Generate</b>	COMPLETE	411

## 4.1. 개요(Introduction)

### 4.1.1. 개관(Overview, Background and Motivation)

- 표준 용어는 OMOP CDM의 일부로 OHDSI 협업을 통해 유지 관리됨.
- 표준 용어는 다음과 같은 경우에 필요한 모든 code sets, terminologies, vocabularies, nomenclatures, lexicons, thesauri, ontologies, taxonomies, classifications, abstractions 및 기타 데이터가 포함됨.
  - 원시 데이터를 CDM으로 변환된(i.e., 표준화) 데이터의 생성
  - 변환된 데이터의 검색, 쿼리 및 추출, 변환된 데이터에 내재된 클래스 및 개념의 계층구조 탐색
  - 데이터의 의미 해석
- OMOP CDM에서 사용 가능한 환자 수준의 데이터는 표준화된 용어의 개념을 사용하는 모든 임상적 사실과 사건을 명시적으로 나타내야 함.
- 극히 일부를 제외하고는, CDM 테이블에는 글자 그대로의 정보는 없음. 따라서 표준화된 용어는 OMOP CDM의 필수 요소임.
- OMOP CDM에서 이용된 표준용어는 누구나 웹페이지에서 다운로드 받아 무료로 이용할 수 있음.
- 일반적으로 매우 소수의 상업적 어휘라고 명시된 이외의 OMOP CDM의 모든 부분과 구성 요소들은 오픈 소스임. 자세한 내용은 Apache 라이센스 V2.0 참조할 것.
- 표준 용어는 다음과 같은 임상 영역 데이터의 표준화된 표현을 제공함.
  - Demographics : Gender, Ethnicity, Race
  - Condition
  - Drug
  - Procedure
  - Measurement
  - Observation

- Note
- Device
- Specimen
- Unit
- Visit
- Death
- Provider

#### **4.1.2. 배경(Background)**

- 전자 형태로 대규모 보건의료 데이터베이스를 사용할 수 있게 되면서 환자에게 의료 서비스를 적용하는 방법에 대한 체계적이고 대규모의 근거와 통찰력을 얻을 수 있게 되었음.
- 이런 분야를 관찰 결과 연구(Observational Outcome Research)라 하며, 질병의 발생과 다른 임상적 사건의 영향을 설명하고 이해하기 위해 환자 수준의 종적인 임상 자료를 사용하고, 질병의 진행에 대한 치료 개입을 설명함.
- 이 연구는 일반적으로 연구 이외의 목적으로 수집되는 데이터(보험 청구와 전자 건강 기록[EHR] 또는 전자 의무 기록[EMR]과 같은 관리 데이터를 연구에 추가적으로 사용하는 것임.
- 일차적 목적으로 자료를 수집하기 때문에 자료의 형식과 표현은 일차적 용도를 따르고 있고, 이에 따른 영향이나 편향이 생길 수 있음. 또한 모든 원천 데이터 세트는 형식과 콘텐츠 표현에서 서로 다름. 보건의료 시스템이 국가마다 다르기 때문에 이 문제는 국제적 연구 수행을 더욱 더 어렵게 함.
- 이 모든 것이 강력하고 재현 가능하고 자동화된 연구를 중요한 도전과제로 만들었음.
- 해결책은 데이터와 표현의 표준화임.
  - 표준화는 서로 다른 원천데이터에 대하여 방법론과 도구를 사용할 수 있도록 해주어 분석가가 특정 데이터의 특이성을 분석하여 연구에 적합하도록 조작해야하는 일을 할 필요가 없게 될 뿐만 아니라,
  - 하나의 데이터 세트에 대한 분석 방법을 개발하여 CDM 형식의 다른 데이터 세트에 적용할 수도 있게 됨.

- OMOP의 표준 용어는 OMOP 연구 수행을 위한 OMOP 공통 데이터 모델과 함께 초기에 구축되었음.
- 따라서 이 용어들의 설계는 매우 간단하고 이러한 시험을 수행하는 데 필요한 최소 기능을 가능하게 하는 동시에, 오픈 소스 프로젝트의 매우 제한된 자원을 사용하여 이와 같은 대규모 자원을 유지하고 개발할 수 있도록 함.
- 2013년 11월 OMOP가 종료된 이후 OMOP Standard Vocabularies는 Odysseus에서 운영하는 OHDSI의 지원을 받고 있음.

### 4.1.3. 원칙(Principles)

- 표준 용어는 몇 가지 원칙에 의해 만들어지며 모든 원칙이 완벽하게 수행되지는 않지만, 지속적인 개선 및 개발 과정의 일반적인 동기와 방향을 나타냄.
  - 표준화 : 관찰 데이터에 사용되는 여러 용어는 공통된 형식으로 통합됨. 이것은 연구자들이 원래의 용어들의 다양한 형식과 life cycle 규칙을 이해하고 다루어야 하는 짐을 덜어줌.
  - 단일 표준 개념 : 각 임상 실체에 대해 그것을 대표하는 하나의 개념, 즉 표준 개념만 있음. 다른 동등하거나 유사한 개념은 비표준으로 지정되고 표준 개념에 매핑됨.
  - 도메인: 각 개념에는 도메인이 할당되어 있음. ‘Dirty(더러운)’이라는 비표준 개념도 하나 이상의 도메인에 속할 수 있음. 이것은 또한 쿼리 시 임상 실체를 어떤 CDM 테이블에 배치하거나 조회해야 하는지를 정의함.
  - 포괄적 적용 범위 : 환자의 의료 경험(예: Condition, Procedure, 의약품에 대한 노출 등) 및 의료 시스템의 행정적 산출물(예: Visits, Care Sites 등)과 관련된 모든 사건은 도메인 개념이 적용됨.
  - 계층구조(Hierarchy) : 도메인 내에서 모든 개념은 계층 구조를 이루고 있음. 이를 통해 상위 개념(예: 약물 클래스)에 따라 계층적으로 포함되는 모든 개념(예: 의약품)을 쿼리할 수 있음. 이것은 두 가지 문제 해결을 의미함.
    - 각 개념에는 하나 이상의 분류가 있어야 함(상향식).
    - 각 분류에는 모든 관련 개념(하향식)이 포함되어야 함.
  - 용어 내·간 관계 및 비표준 개념에서 표준 개념으로 매핑

- 수명 주기(Life cycle)가 데이터 표현을 최신 상태로 유지하지만 사용종료 및 개정된 개념의 처리를 지원함.
- 이 기준들은 관찰 연구에 기여하는 목적이 있다는 것에 주목하는 것이 중요함.
  - 그런 점에서 표준 용어는 전체 생물 의학 문헌의 색인화와 검색을 지원하는 UMLS와 같이, 개념을 동등하게 매핑하고 있는 대규모 수집과는 다름.
  - UMLS 자원은 많은 표준화된 어휘 요소들을 구성하기 위한 기반으로 매우 많이 사용되고 있지만, 프레임워크를 조정하기 위한 상당한 추가 노력이 이루어짐.
    - 주로 메타데이터를 위한 추가 용어가 구축됨.
    - 포괄적인 커버리지 달성을 위해 매핑과 관계가 추가되고 있음. 동등성을 달성할 수 없는 경우, 보다 세분화된 비표준에서 상위 수준의 표준 개념에 이르는 ‘상향적(uphill)’ 관계가 형성됨.
    - 포괄적 도메인 구조가 구축되고 각 개념에는 도메인(또는 도메인의 조합)이 할당되었음.
    - 의학 및 임상 진료에서 사용되는 분류를 나타내기 위해 도메인 내의 계층적 트리가 구축됨.
- 그러나 모든 도메인의 모든 기준을 달성하기 위해 중요한 작업이 수행되어야 함. 현재 복잡한 도메인 및 비관리 도메인의 경우 다음과 같은 컴플라이언스를 달성할 수 있음.

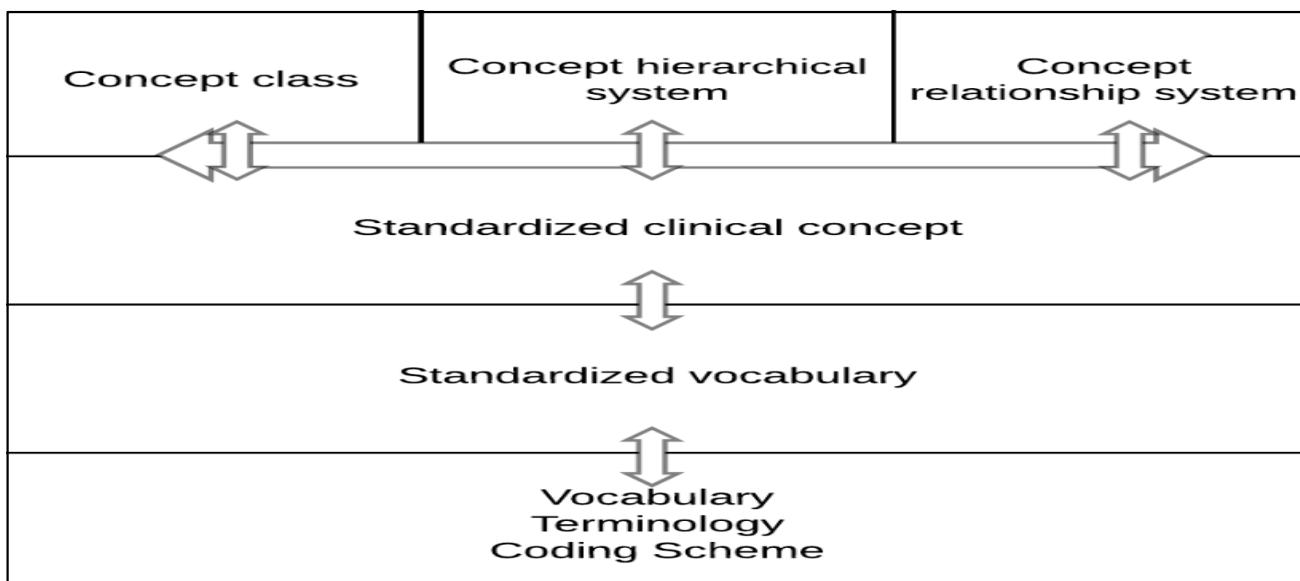
Domain	Standard-ization	Unique Concepts	Reliable Domains	Comprehensive Coverage	Hierarchy	Mapping
Drug	×	x	In US, other countries in process	×	×	×
Condition	×	x	mostly	×	×	mostly
Procedure	×	heavily overlapping	×	×		×
Measurement	×	somewhat	mostly	×	minimal	
Device			mostly			
Unit	×	x	×	×		

- 모든 개념에 대해 Life Cycle이 구현되며, 그 규칙은 CONCEPT table 설명 및 개별 용어에 대한 논의(특정 규칙이 있는 경우, 그러나 이것은 드문 경우)에서 설명됨.

#### 4.1.4. 정의 및 용어(Definitions and Terms)

- OMOP CDM(Common Data Model)의 목적상 다음과 같은 용어가 사용됨.

용어	설명
표준 용어 (Standardized Vocabularies)	공통 형식으로 통합되고 CDM 테이블에 저장되는 Vocabularies, Classification, Domains 및 Concepts의 체계를 포함함.
용어(Vocabulary)	사용 가능한 개념의 계층구조, 종양학 또는 분류학을 포함하여 그것들 간 사용 가능한 관계를 포함하는 코드나 개념의 집합 많은 어휘들이 ICD-9-CM, SNOMED-CT, RxNorm, Read와 같은 국가 또는 국제 기구에서 채택됨.
Terminology	Vocabulary와 유사하며 종종 동의어로 사용되지만 이 문서에서는 사용되지 않음.
코딩 체계(Coding scheme)	Vocabulary와 유사하며 종종 동의어로 사용되지만 이 문서에서는 사용되지 않는다.
분류(Classification)	약물에 대한 화학 구조와 같이 의미론적으로 유용한 클래스를 정의하는 개념 및 개념 관계의 계층적 시스템
도메인(Domain)	표준화된 모든 개념에 대해 정의된 약물, 병태(Condition), 행위(Procedure) 등과 같은 임상적 의미 범주
개념(Concept)	각 용어에 정의된 정보의 기본 단위
개념 클래스(Concept Class)	용어 내에서 분류를 특징짓는 개념의 속성 분류(Classification)와의 차이점은 개념 클래스는 계층구조가 없는 단일 속성이라는 것임.

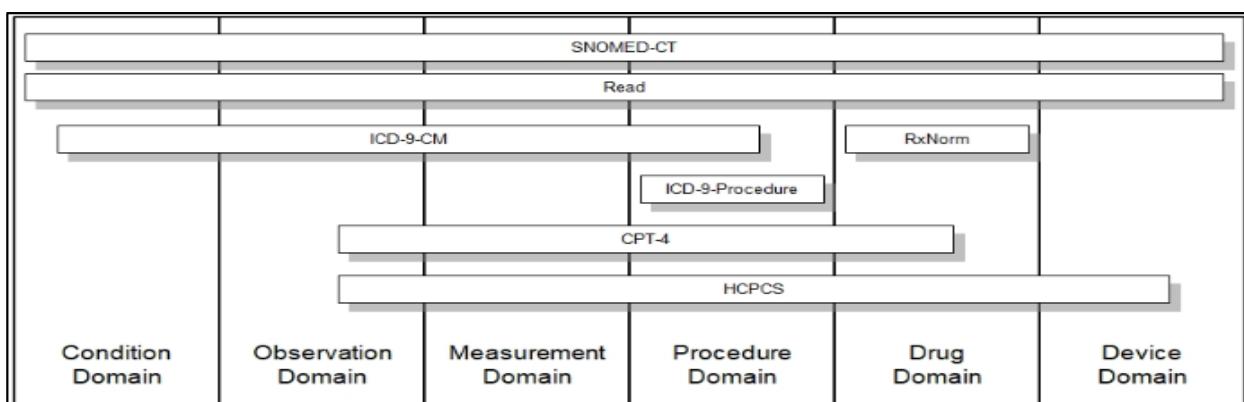


## 4.2. 일반적 구조와 사용(General Structure and Use)

- OMOP 표준 용어는 의료 정보를 기록하는 각기 다른 측면에서 사용되고 있는 많은 다양한 용어들을 결합하고 있음. 이러한 서로 다른 목적은 각각의 용어들이 다른 형식, 질, 포괄성 및 범위, 그리고 life cycle을 갖도록 함. 이러한 가변성에 어떤 질서를 부여하기 위해, 다수의 구조 요소가 용어들에 도입되었으며, 본 절에서 설명될 것임.
  - 참고 : 용어의 내용이 그대로 남아 있더라도, 새로운 형식과 구조적 요소들은 각 어휘를 고유 형태와 다르게 표현하게 될 것임. 원래의 형식으로 용어가 필요한 경우 OMOP 표준 용어를 사용하지 말 것.

### 4.2.1. 도메인과 용어(Domains and Vocabularies)

- 표준 용어는 도메인과 용어로 구성되어 있음. 도메인은 임상적 개체(clinical entity)의 특성 또는 유형을 의미함. 그것은 또한 데이터 레코드를 저장할 CDM 데이터 테이블을 정의함. 용어는 국외 또는 국제적으로 존재하는 표준에서 가져온 개념 세트이거나 적절한 표준을 사용할 수 없는 경우 표준 용어 팀에 의해 만들어진 개념 세트임.
- 참고 : 도메인과 용어 사이에 일대일 관계는 없음. SNOMED 또는 Read와 같은 일부 용어는 매우 광범위하며 모든 의료 도메인의 개념을 포함하고 있음. 약물에 대한 RxNorm 또는 행위에 대한 ICD9Proc과 같은 다른 용어들은 특정 도메인에 한정됨. 많은 경우에서 용어는 단일 도메인이라고 가정되지만 실제로는 그렇지 않음. 예를 들어 CPT4와 HCPCS는 이름에 의해 행위 코드만 포함할 것으로 예상되지만 실제로는 관찰, 조건, 기기 및 약물 개념을 포함하고 있음.



#### 4.2.2. 표준, 분류 및 원천 개념(Standard, Classification and Source Concepts)

- 도메인 내에서 코드는 다수의 용어로부터 파생되며, 코드는 종종 동일하거나 중복되는 의미를 가짐. 이러한 상황을 정리하기 위해, 각 코드는 다음 세 가지 중 하나로 지정됨.

- 표준 개념(standard concepts, standard\_concept = 'S' )

표준 개념은 표준화된 임상 데이터 테이블의 고유한 임상 개체를 나타내는데 사용되는 ‘공식’ 개념임. 그 개념 ID는 각각의 \*\_concept\_id 필드임. 일반적으로 표준개념은 도메인에 대한 포괄적인 적용범위의 잘 확립된 용어에서 비롯되며, 그 개념은 잘 정의됨. 예를 들어, Condition 도메인에서는 SNOMED 어휘를 통해 달성됨. 장치 도메인과 같이 특정 도메인에서 사용 가능한 엔티티의 포괄적 목록이 없는 경우 표준 개념은 다양한 Vocabularies에서 가져온 것임. 행위 도메인에서도 마찬가지임.

- 분류 개념(Classification Concepts, standard\_concept = 'C' )

분류 개념은 표준 개념과 계층적 관계를 가지므로 CONCEPT\_ANCESTOR 테이블의 기록을 사용하여 표준 개념을 쿼리하는데 사용할 수 있음. 그러나 이러한 데이터 자체는 데이터 테이블에 표시될 수 없음. 예를 들어, ‘COPD’에 대한 MedDRA 개념은 이 병의 모든 형태인 표준 SNOMED-CT 개념과 계층적 관계를 가짐. 마찬가지로, ‘VA class’라는 용어의 개념 4283987 ‘ANTICOAGULANTS’는 DRUG\_EXPOSURE 또는 DRUG\_ERA 테이블에 표시될 수 없지만, ‘Ingredient’ , ‘Clinical Drug’ 또는 ‘Branded Drug’과 같은 개념 클래스를 가진 그 Descendant Concepts은 가능함.

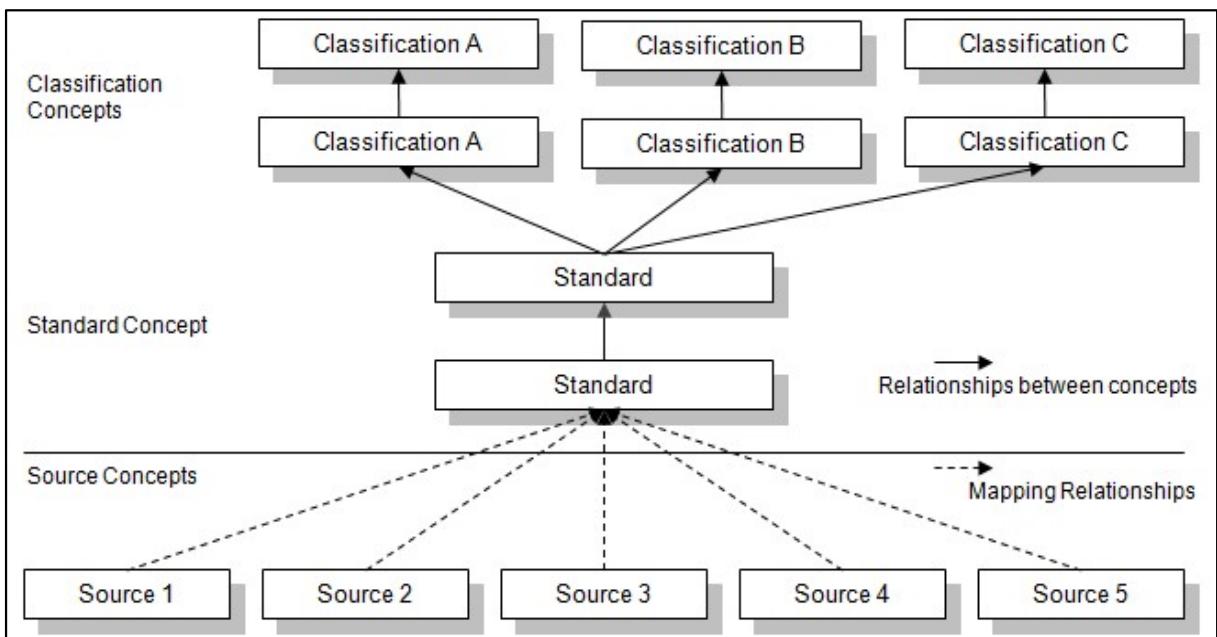
분류개념을 표준개념과는 다른 vocabularies에서 도출할 수도 있음. 분류 개념은 고유하지 않다는 점에 유의해야 함. 예를 들어, NDF-RT, VA Class, ETC 및 ATC Vocabularies에서 도출한 약물 클래스 ‘항응고제’에 대한 개념이 있음. 또한 멤버십은 용어에 따라 다름. 대부분의 경우 동등한 분류개념의 목록(membership list)은 유사하거나 동일하지만, 의학은 일반적으로 이 세 가지 분류들의 표준적 정의에 합의하지 않았음.

- 원천 개념(Source Concepts, standard\_concept = NULL)

원천 개념들은 모두 Standard나 Classification Concepts이 아닌 나머지 개념들임. 이 개념들은 시간에 따라 변경될 수 있음에 유의할 것: 사용이 종료된 경우(valid\_end\_date가 2099-12-31보다 작고 invalid\_reason = ‘D’ 또는 ‘U’ 인

경우), 이전의 표준 또는 분류 개념은 공식적으로 원천 개념으로 변경됨. 원천 개념은 데이터 테이블의 \*\_source\_concept\_id 필드에만 나타날 수 있음. 그들은 원천 데이터에서 코드를 나타냄. 각 원천 개념은 ETL 프로세스 중에 하나 이상의 표준 개념에 매핑됨. 매핑을 사용할 수 없는 경우 ‘the Standard Concept with the concept\_id = 0’이 \*\_concept\_id 필드에 작성됨. 자세한 내용은 데이터 ETL 섹션을 참조할 것.

원천 개념은 도메인의 모든 개념에 대해 다음과 같은 논리적 구조를 생성함.



- Source Concepts은 표준 개념을 통해 매핑되며, 이것들은 서로 다양한 의미론적 관계를 가지고 있음. 또한 그들은 서로 계층적 관계와 분류 개념을 가지고 있는데, 이 경우에는 용어 A, B, C에서 도출됨. 일반적으로 분류 개념 사이의 계층적 관계는 동일한 용어의 개념 간에만 일어남.
- 모든 개념은 CONCEPT 테이블에, 모든 관계는 CONCEPT\_RELATIONSHIP 테이블에, 그리고 모든 세대간 계층적 관계는 CONCEPT\_ANCESTOR 테이블에 저장됨. 후자는 표준 개념과 분류 개념 사이에만 정의되며, 원천 개념은 이 계층적 트리 구조에 있지 않음.
- 용어의 모든 요소는 CONCEPT 테이블에 개념으로 표현됨. 이는 다음과 같은 세 가지 범주의 개념으로 귀결됨.
  - 표준 개념: 표준 개념은 고유한 임상적 개체(entity)를 나타내기 위해 사용되어야 하는 ‘공식적’ 개념으로 standard\_concept = ‘S’로 지정됨. 표준 개념

만일 CDM의 표준화된 임상 데이터 테이블의 \*\_concept\_id나 \*\_type\_concept\_id 필드에 나타낼 수 있음.

- 분류개념(standard\_concept = 'C'로 지정)은 표준개념과 계층적 관계를 가지므로 CONCEPT\_ANCESTOR 테이블의 기록을 사용하여 표준개념을 쿼리하는 데 사용할 수 있음. 그러나 이러한 그 데이터 자체는 데이터 테이블에 표시될 수 없음. 예를 들어, 용어 'VA class'의 개념 4283987 'ANTICOAGULANTS'는 DRUG\_EXPOSURE 또는 DRUG\_ERA 테이블에 표시될 수 없지만, 개념 클래스 'Ingredient', 'Clinical Drug' 또는 'Branded Drug'과 같은 그 하위 개념들은 나타낼 수 있음.
- 원천 개념은 데이터 테이블의 \*\_source\_concept\_id 필드에만 나타날 수 있음. 이는 원천 데이터의 코드를 나타냄. 각 원천 개념은 ETL 프로세스 중에 하나 이상의 표준 개념에 매핑됨. 매핑을 사용할 수 없는 경우 concept\_id = 0인 Standard Concept는 \*\_concept\_id 필드에 작성됨. 자세한 내용은 데이터 ETL 섹션을 참조할 것.
- 개념을 표준(Standard), 분류(Classification) 또는 소스(Source)로 지정하는 것은 용어에 따라 달라짐. 자세한 내용은 각 용어별 상세내용을 참조할 것.

#### 4.2.3. 개념의 매핑(Mapping of Concepts)

- 매핑은 하나의 개념을 다른 개념으로 변형하는 과정임.
- CDM의 임상 데이터 테이블은 표준 개념만 허용함.
- 원천데이터베이스에 사용된 모든 코드는 표준개념으로 변환되어야 함.
- 매핑은 CONCEPT\_RELATIONSHIP 테이블에 기록함으로써 이루어짐. 그들은 많은 특정 관계번호(relationship\_id)를 통해 각 개념을 표준 개념에 연결함:

Relationship ID	Purpose
Maps to	개념을 *_concept_id에 배치할 표준 개념으로 매핑
Maps to value	MEASUREMENT 및 OBSERVATION 테이블의 value_as_concept 필드에 넣을 개념과 표준 개념 간 특수 매핑

#### 4.2.3.1. 'Maps to' 관계

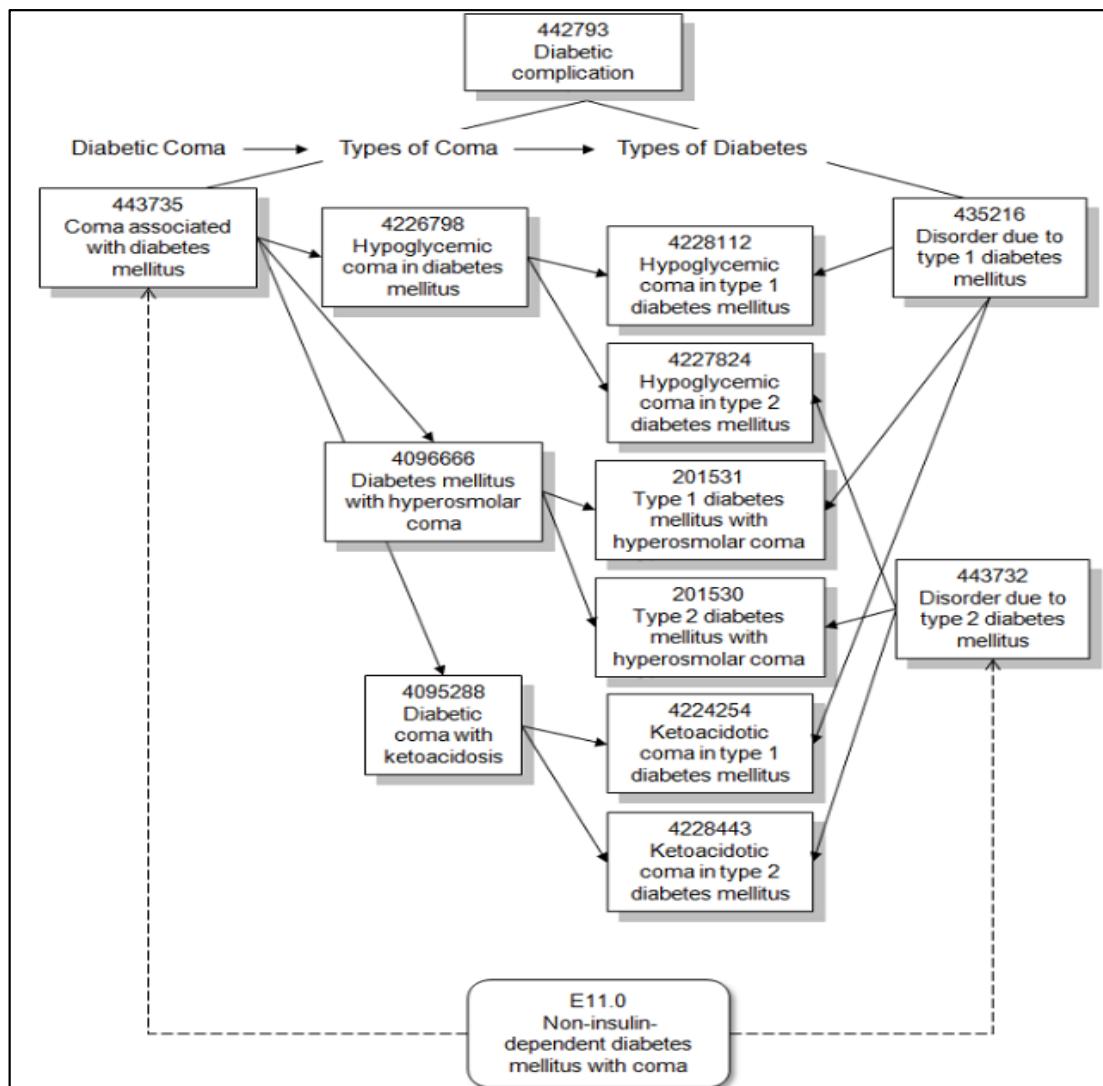
- 'Maps to' 매핑이 필요한 개념은 원천 개념과 표준 개념임. 매핑은 동등한 (equivalent) 대상 개념에 연결하기 위하여 시도함. '동등한(Equivalent)' 이란 동일하다는 의미를 가지며, 중요한 것은 계층(있는 경우)의 하위도 동등하거나 동일한 의미론적 범주를 포괄한다는 것임. 동등한 개념이 없는 경우, 매핑은 더 넓은 개념에 대한 매치를 시도함. 이렇게 하면 대상 용어의 쿼리가 원래의 소스 용어에서 쿼리된 것과 동일한 레코드를 검색할 수 있음.
- 일반적으로 원천 개념과 표준 개념은 다음과 같이 매핑 됨.
  - 원천 개념은 하나 또는 몇 개의 표준 개념에 매핑됨. 만일 그것들이 둘 이상의 표준 개념에 매핑 되었다면, 그 결과로 생긴 CDM 테이블에서는 소스의 각 레코드에 대해 둘 이상의 레코드가 작성됨.
  - 또한 표준 개념은 표준 개념에 매핑되며, 일반적으로 이것은 그 자체로 매핑임.
  - 분류개념(standard\_concept= 'S')은 표준개념에 대한 매핑을 가지고 있지 않음.

#### 4.2.3.2. 'Maps to value' 관계

- 이러한 관계는 관찰 및 측정과 그 결과를 구별하도록 설계되었음. 예를 들어, ICD9CM V12.71 concept\_id 44820383 'Personal history of peptic ulcer disease' 는 SNOMED 4214956 'History of clinical finding in subject' (Observation Domain) 와 SNOMED 4027663 'Peptic ulcer(Condition Domain)' 와 또 다른 'Maps to value' 관계를 가지고 있음.
- 매핑으로 인한 '손실(Losses)'  
한 용어에서 다른 용어로의 매핑이 데이터의 품질에 미치는 영향과 주어진 포함 및 제외 기준에 대한 코호트의 환자를 신뢰성 있게 식별할 수 있는지에 대한 중대한 우려가 있음. 이러한 효과는 다음과 같은 다양한 이유가 있을 수 있음.
  - 매핑이 누락된 경우: 매핑이 추가될 수 있도록 OHDSI 포럼에 보고할 것.
  - 소스코드가 잘못 정의됨(예: ICD9CM 799 'Other ill-defined and unknown causes of morbidity and mortality')
  - 네거티브 기록(예: ICD9CM V64.0 'Vaccination not carried out' )
  - 환자와 무관한 코드(예: ICD9CM V65 'Other persons seeking consultation' )

- 소스 및 대상 계층 구조가 불일치

- 후자의 효과는 매우 pre-coordinated Concepts(다른 차원의 조합인 복잡하고 정교한 개념)에서 드물지 않으며, 조정은 개념이 매핑되는 용어의 위상의 구조(structure of topology)에 따라 달라짐.
- 만약 그것들이 동등하지 않다면, 직접적인 매핑을 설정할 수 없음. 그러나 계층적 개념을 사용하여 검색할 때 이러한 개념을 검색하는 기능을 유지하기 위해 대신 두 가지 이상의 매핑을 제공함. 예를 들어 ICD10 Concept 45755355 ‘Non-insulin-dependent diabetes mellitus with coma’ (코드 E10.0)은 SNOMED에 직접 매핑할 수 없음:



- SNOMED 대상의 위계구조는 ICD10의 조직 구조(organizational structure)와 다른 topology를 가짐. 여기서 제1형(인슐린 의존성, E10)과 제2형(인슐린 비의존성, E11) 당뇨병이 주요 특징임. 제2형 당뇨병과 혼수상태 E11.0의 조합은 합병증의 혼수상태와 함께 제2형 당뇨병을 정의함. SNOMED에서 당뇨 합병증은 그 자체의 계층을 가지고 있는데, 이것은 혼수상태의 당뇨나 1, 2형 당뇨병에 합병증을 동반하는 개념으로 나눠짐. 다음 단계에서는 혼수상태의 형태(hypoglycemic, hyperosmolar and ketoacidotic)가 구별됨. 그러나 제2형 당뇨병과 혼수상태의 조합은 없음.
- 해결책은 E11.0을 당뇨병과 관련된 44054006 Type 2 당뇨병과 420662003 코마 둘 다에 매핑하는 것임. SNOMED에는 뚜렷한 등가물이 없지만, 혼수상태와 함께 2형 당뇨에 대한 계층적 질의가 정확한 기록을 찾아냄.

#### **4.2.3.3. 병태 매핑 규칙(Rules for Condition Mapping)**

- 일반적으로 병태(Condition)란 진단, 징후 또는 증상으로 언급된 질환이나 질병의 존재를 나타내는 레코드임. 그러나 인체의 생물학적 기능에서 비롯되지 않았지만, 여전히 의학적 치료가 필요한 병태가 많이 있음. 또는 Condition이 맞지만, 기록당시(History of)에 유효하지 않거나 해당 주체의 병태가 아닌 경우(Family history of)도 있음. 이 때 매핑 규칙은 다음과 같음.
  - 시술 후 관리(Aftercare after procedures). 그것들은 단일 개념 413467001 ‘Aftercare’에 매핑된 후, 이차적으로 ‘Maps to Value’ 관계로 적절한 행위에 매핑됨. 예를 들어 ICD10CM Z47.1 ‘Aftercare following joint replacement surgery’는 413467001 ‘Aftercare’ 및 4189532 ‘Implantation of joint prosthesis’ 행위에 모두 매핑됨. ‘Aftercare’은 그 자체가 병태가 아니라 관찰 도메인에 귀속된다는 점에 유의할 것.
  - 다른 병태의 후유증(Late effects or sequelae of other conditions). 가능하면, 이것들은 후유증이라고 설명하는 단일 개념에 매핑됨. 예를 들어 ICD10CM S82.874S concept\_id 45589204 ‘Nondisplaced pilon fracture of right tibia, sequela’는 197150 ‘Late effect of fracture of lower extremities’로 매핑됨. 도메인은 병태(Condition)임.
  - 병력(History of a condition). 이는 condition\_start\_date 시점에 있지 않은 병태를 말함. 따라서, 그들은 Observation Domain에 있는 개념 4214956

‘History of clinical finding in subject’에 매핑됨. 병태 자체는 ‘Maps to Value’ 관계를 통해 기록됨. 예를 들어 ICD10CM Z87.820 concept\_id ‘Personal history of traumatic brain injury’는 4214956 ‘History of clinical finding in subject’과 ‘Maps to value’ 4132546 ‘Trainical braining’에 매핑됨.

- 가족력(Family history of a condition). 이 역시 4167217년 ‘Family history of clinical finding’와 실제 병태를 ‘Maps to Value’ 관계를 통해 이중으로 매핑함. 예를 들어 ICD10/ICD10CM Z80.0 concept\_id 45542462 ‘Family history of malignant neoplasm of digestive organs’는 4167217 ‘Family history of clinical finding’와 443568 ‘Malignant neoplasm of gastrointestinal tract’에 매핑됨.
- 치료력(History of medical treatment). 마찬가지로, 이것들은 Observation Concept 4207283 ‘History of drug therapy’와 실제 치료와의 ‘Maps to Value’ 링크를 통해 매핑됨. 예를 들어, ICD10 Z92.0 concept\_id 45605174 ‘Personal history of contraception’는 4207283 ‘History of drug therapy’과 4027509 ‘Contraception’으로 매핑함. 정확한 성분 또는 임상/브랜드 약물이 알려진 경우, 이러한 약물 개념으로 직접 매핑되겠지만 전형적으로 그렇진 않음.
- 약물 부작용(Adverse effect of medication). 가능하다면, 동등한 것에 직접 매핑되나, 대개 그런 경우는 존재하지 않고, 위의 치료력처럼 처리됨. 예를 들어, ICD10CM T36.8X5A concept\_id 45551127 ‘Adverse effect of other systemic antibiotics, initial encounter’와 T36.8X5D ‘Adverse effect of other systemic antibiotics, subsequent encounter’는 437191 ‘Antibacterial drug adverse reaction’에 매핑되고, ICD10CM T36.8X5S concept\_id 45560654의 ‘Adverse effect of other systemic antibiotics, sequela’는 4207283 ‘History of drug therapy’에 모두 매핑됨.
- 약물 과소 투여(Underdosing of medication). 이 정보는 부작용과 유사하게 취급되며, 일반적으로 이에 정확하게 부합하는 SNOMED 개념은 존재하지 않음. 예를 들어 ICD10CM T36.0X6A concept\_id 45565479 ‘Underdosing of penicillins, initial encounter’ 및 T36.0X6D concept\_id 45565480 ‘Underdosing of penicillins, subsequent encounter’는 40488434

‘Medication dose too low’로 매핑하고, 반면에 T36.0X6S concept\_id 45565481 ‘Underdosing of penicillins, sequela’는 4207283 ‘History of drug therapy’로 매핑함.

- 장기 부재 또는 이식/인공삽입물의 존재(Status of organ absence or transplant /prosthetic presence). 장기 부재는 장기를 제거하는 시술로 발생함(선천적인 병태 제외, 이 경우 이에 상응하게 매핑). 따라서, 4215685 ‘Past history of procedure’와 ‘Maps to Value’ 링크를 통해 기관을 제거하는 적절한 시술에 매핑됨. 예를 들어 ICD10CM Z94.0 concept\_id 35225404 ‘Kidney transplant status’는 4215685 ‘Past history of procedure’ 및 4322471 ‘Transplant of kidney’에 매핑됨. Z95.5 concept\_id 35225418 ‘Presence of coronary angioplasty implant and graft’는 4215685 ‘Past history of procedure’ 및 4184832 ‘Coronary angioplasty’에 매핑됨.
- AND나 OR을 통해 두 개 이상의 의미 구성요소를 나열하는 사전 조합된 개념(Pre-coordinated Concepts). 이러한 개념은 다음과 같은 우선순위로 처리됨.
  - 계층적으로도 잘 연결되는 등가 조합 개념으로 매핑
  - 두 구성 요소 모두 개별적으로 매핑
  - 가장 일반적인 상위개념으로 매핑
- 위 세 가지 경우에 대한 예는 다음과 같음.
  - ICD10 A01 concept\_id 45576225 ‘Typhoid and paratyphoid fevers’는 4022808 ‘Human typhoid AND/OR paratyphoid fever’과 단일 ‘Maps to’ 관계를 가짐.
  - IC10CM F12.22 concept\_id 45591098 ‘Cannabis dependence with intoxication’은 4052690 ‘Cannabis intoxication’ 와 440387 ‘Cannabis dependence’에 두 개의 ‘Maps to’ 관계를 가지고 있음.
  - ICD10 L02.0 concept\_id 45596354 ‘Cutaneous abscess, furuncle and carbuncle of face’는 400082007 ‘Disorder of skin of head’에 단일 ‘Maps to’ relationship\_id를 가짐.
- 모성 관리(Maternal care). 많은 병태(conditions)가 임신한 여성의 병태가 아니라 태아의 병태 때문에 관리를 요하지만, 이런 모든 병태들은 엄마에게 매핑되고 있음. 예를 들어, ICD10 O35.6 concept\_id 45567927 ‘Maternal

care for (suspected) damage to fetus by radiation’ 는 1992553006 ‘Fetus with radiation damage’ 와 289908002 ‘pregnancy’ 에 두 개의 ‘Maps to’ 관계를 가지고 있고, 두 가지 병태 모두 엄마에게 기록됨.

- 면역 필요(Need for immunization). 이 개념은 비면역(Immunity gap)을 나타내는 관찰(Observation)에 매핑됨. 그 다음에 relationship\_id ‘Maps to value’ 를 이용하여 예방접종이 필요한 condition(SNOMED 개념으로 표현)에 두 번째 매핑이 됨. 여기서 백신 자체(RxNorm Concept로 표현될 수 있는)에 매핑되지 않았다는 점에 유의해야 함. 예를 들어 ICD10 Z23 concept\_id 45556822 ‘Need for immunization against single bacterial diseases’ 는 170536002 ‘Vaccination required’ 에 ‘maps to’ 로, 87628006 ‘Bacterial infectious disease’ 에 ‘maps to value’ 로 매핑됨.
- 비정상적인 검사 결과를 나타내는 병태(Conditions indicating abnormal levels of a test). 이것들은 측정값(Measurement)과 결과(result) Concepts으로 나눠짐. 예를 들어 ICD10 R77.1 concept\_id 45553745 ‘Abnormality of globulin’ 은 측정 43510 ‘Globulin measurement’ 과 ‘Maps to value’ 관계로 4135493 ‘Abnormal’ 로 매핑됨.
- 장치 사용에서 기인된 병태(Conditions resulting from the use of Devices). 장치에 대한 노출의 부정적 또는 기타 결과는 473023007 ‘Complication associated with device’ 에 매핑되고, 가능한 경우 ‘Maps to value’ 관계로 장치 자체에 대하여 매핑됨. 예를 들어 ICD10 Y73.0 concept\_id 45590664 ‘Gastroenterology and urology devices associated with adverse incidents, diagnostic and monitoring devices’ 는 45767866 ‘Gastro-urological device’ 에 ‘Maps to value’ 링크를 가짐.

#### 4.2.4. 자료 ETL(Data ETL)

- 원시 데이터부터 CDM 형식의 데이터에 이르기까지 표준화된 용어가 ETL 프로세스에 미치는 가장 중요한 영향은 각 개념의 도메인임. 기록이 어떤 원본 테이블 또는 어떤 코딩 체계를 가지고 있는지에 관계없이 대상 테이블은 각 개념의 domain\_id에 의해 결정됨. 모든 ETL은 원천 데이터의 모든 기록을 처리하기 위해 다음과 같은 논리를 따라야 함:
  - 레코드에서 소스코드를 검색할 것(Retrieve the source code from the

record). 소스코드에 해당하는 개념을 찾을 것. 대부분의 경우, 이것은 CONCEPT 테이블의 vocabulary\_id 필드의 올바른 내용을 가진 concept\_code 필드의 소스코드를 조회함으로써 이루어짐. 때때로, 소스코드는 정확한 위치를 찾기 위해 조작될 필요가 있음. 예를 들어 ICD-9-CM 코드에는 두 번째 또는 세 번째 문자 뒤에 점이 있지만 원천 데이터에서는 점 없이 저장되는 경우가 많음. 또는 NDC 코드는 다양한 형식(-/\*를 동반 또는 미동반, 9 또는 11자리)으로 제공되므로 매핑 프로세스가 더 복잡해짐. 검색 과정의 특정 권장사항에 대해서는 각 용어별 내용을 참조할 것.

- CONCEPT\_RELATIONSHIP 테이블에서 모든 활성(invalid\_reason 필드가 NULL 이어야 함) 레코드를 검색하여 표준 개념(standard\_concept= ‘S’ )에 매핑할 것. 매핑하고자 하는 개념을 concept\_id\_1으로 사용하고, 매핑 대상 개념을 concept\_id\_2로 relationship\_id= ‘Maps to’ 와 함께 사용할 것. 소스코드가 로컬 코드이고 SOURCE\_TO\_CONCEPT\_MAP 테이블이 있는 경우 target\_concept\_id 필드에서 대상 개념을 확인할 것. 대상 개념이 표준 개념이나 다른 용어의 개념일 경우 개념 자체가 될 수 있음. 대부분의 경우, 소스코드는 단일 대상 표준 개념에 매핑되지만, 어떤 경우에는 두 개나 세 개 개념에 매핑될 수도 있음.
- domain\_id 필드의 내용을 기준으로 각 대상 표준 개념에 해당하는 CDM 테이블에 레코드를 기록할 것. 표준 개념의 ID를 <domain>\_concept\_id에, 원천 개념을 <domain>\_source\_concept\_id에, 소스코드를 <domain>\_source\_value 필드에 넣을 것. 대부분의 데이터 테이블은 start\_date와 type concept을 필요로 하며, 일부 테이블은 더 많은 필드를 고려해야 함. CDM Specifications에서 각 표의 세부 정보를 참조할 것. 각 도메인에 해당하는 테이블/필드 조합은 다음과 같음

CDM table	Field	Comment
Any	Any	일반 개념(Generic Concepts)은 concept_id로 끝나는 모든 필드에 있을 수 있음.
	gender_concept_id	
	race_concept_id	
	ethnicity_concept_id	
	visit_concept_id	
	procedure_concept_id	

CDM table	Field	Comment
MEASUREMENT or OBSERVATION or SPECIMEN	modifier_concept_id	
	drug_concept_id	
	route_concept_id	
	unit_concept_id	Units은 다른 맥락에서 사용됨.
	device_concept_id	
	condition_concept_id	
	measurement_concept_id	
	operator_concept_id	
	value_as_concept_id	
	observation_concept_id	
MEASUREMENT or OBSERVATION or SPECIMEN	relationship_concept_id	
	place_of_service_concept_id	
	specialty_concept_id	
	currency_concept_id	Currency values는 모든 *_COST 테이블에 있음.
	revenue_code_concept_id	
	specimen_concept_id	
SPECIMEN or MEASUREMENT or OBSERVATION	anatomic_site_concept_id or value_as_concept_id or value_as_concept_id	Anatomical Site Concepts은 Specimen의 origin을 설명하기 위해 사용되는 동시에 측정이나 관찰 결과로도 사용됨.
	disease_status_concept_id	

- \* 잠재적 대상 테이블이 두 개 이상인 경우, ETL은 표준 개념이 사용되는 상황을 식별하여 정확한 테이블을 선택해야 함.
  - 일부 원천 개념의 경우 표준 개념에 대한 매핑이 없는데(CONCEPT\_RELATIONSHIP 또는 SOURCE\_TO\_CONCEPT\_MAP 테이블에 기록이 없음), 일반적으로 원천 개념이 너무 일반적이거나 잘못 정의되었기 때문임. 이 경우, 원천 개념 자체의 domain\_id 필드를 사용하여 기록을 정확한 테이블에 배정해야 함. 일부 원천 개념은 ‘장치/행위’ 와 같은 결합 도메인이 있음. 이 경우, 각 결합 도메인 테이블에 레코드를 기록할 것(이 경우 DEVICE\_OCCURRENCE 및

PROCEDURE\_OCCURRENCE 테이블에 기록). 표준 개념으로 개념 0(concept\_id=0) 을 각 <domain>\_concept\_id 필드에 쓸 것. 매핑 프로세스에서 ‘도메인’이나 ‘메타데이터’ 도메인의 개념을 검색하여 매핑 테이블에 오류가 발생한 경우에는 사례를 CDM-Builder에 보고할 것.

- ‘Obs Period Type’, ‘Death Type’, ‘Visit Type’, ‘Procedure Type’ 등과 같은 Type Concept 도메인이 생성된 경우에도 마찬가지임. 이러한 유형의 개념은 유효하고 각 CDM 테이블의 <domain>\_type\_concept\_id 필드에 배치되어야 하지만 매핑 프로세스의 결과가 될 수는 없음. 그들은 기록의 기원을 나타내며, Type 개념의 선택은 ETL 과정에 견고하게 연결되어있어야 함.

### 4.3. 도메인(Domain)

- 도메인은 표준화된 용어의 각 개념에 대해 정의된 OMOP 정의 임상 개념 주임. 도메인 목록은 DOMAIN 테이블에 저장됨. CONCEPT 테이블에서 각 개념 레코드의 domain\_id 필드는 개념이 귀속된 도메인을 결정함.
- 모든 표준 및 분류 개념(standard\_concept = ‘S’ 또는 ‘C’)은 고유한 도메인 을 가지고 있으며, 비표준 원천 개념(standard\_concept = NULL)은 도메인의 조합을 나타내는 domain\_id 항목을 가질 수 있음.

#### 4.3.1. 도메인 할당(Domain assignment)

- 도메인 할당은 각 용어마다 다른 경험치를 사용하여 수행됨. 대부분의 어휘는 소수 도메인 중 하나에 대한 개념을 가지고 있지만, HCPCS나 Read와 같은 일부 용어는 매우 광범위함. 할당은 각 도메인의 정의에 따르지만(아래 참조), 일부 범주의 경우 할당이 모호할 수 있음. 이러한 경우에 OMOP 표준화 용어에서 우리는 다음과 같은 규칙을 사용함:

Category	Example	Domain
Behavioral abnormality	ICD10CM Z72.810 Child and adolescent antisocial behavior	Condition
Blood bank product	CPT4 86930 Frozen blood, each unit; freezing (includes preparation)	Device
Contrast material for imaging	HCPCS A9581 Injection, gadoxetate disodium, 1 ml	Device

Category	Example	Domain
Food Supplement	SPL 304fdee6-0290-4717-be3d-b367bec7e411 Ferrous Gluconate	Device
Graft (alive or artificial)	HCPCS L8670 Vascular graft material, synthetic, implant	Device
Implant (not drug eluding)	HCPCS L8610 Ocular implant	Device
Medical supply	HCPCS A4336 Incontinence supply, urethral insert, any type, each	Device
Nutrition (enteral)	HCPCS B4152 Enteral formula, nutritionally complete, calorically dense (equal to or greater than 1.5 kcal/ml) with intact nutrients, includes proteins, fats, carbohydrates, vitamins and minerals, may include fiber, administered through an enteral feeding tube, 100 cal...	Device
Nutrition (parenteral)	HCPCS B5200 Parenteral nutrition solution: compounded amino acid and carbohydrates with electrolytes, trace elements, and vitamins, including preparation, any strength, stress - branch chain amino acids - premix	Device
Prosthesis	HCPCS L6687 Upper extremity addition, frame type socket, below elbow or wrist disarticulation	Device
Radioactive material for imaging	HCPCS C1204 Technetium tc 99m tilmanocept, diagnostic, up to 0.5 millicuries	Device
Radiological procedure supply, radiopharmaceutical)	HCPCS C2644 Brachytherapy source, cesium-131 chloride solution, per millicurie	Device
Animal Drug	SPL 3cc46410-1360-43ba-af4f-fd33c9bc3533 THYROID SOOTHE - chamomilla, iodum, lycopus, nux vom, zingiber spray	Drug
Cellular Therapy	RxNorm 1000596 autologous cultured chondrocytes 30000000 CELLS/ML Injectable Solution [Carticel]	Drug
Implant (drug eluding)	HCPCS J7311 Fluocinolone acetonide, intravitreal implant	Drug
Procedure Drug	HCPSC J2704 Injection, propofol, 10 mg	Drug
Cytogenetic study	CPT4 88271 Molecular cytogenetics; DNA probe, each (eg, FISH)	Measurement
Cytopathology	CPT4 1012386 Cytopathology smears, cervical or vaginal	Measurement
Drug testing	HCPCS G6030 Amitriptyline (Assay of amitriptyline)	Measurement
Electrocardiogram	HCPCS G0403 Electrocardiogram, routine ecg with 12 leads; performed as a screening for the initial	Measurement

Category	Example	Domain
	preventive physical examination with interpretation and report	
Hearing test	CPT4 92563 Hearing test using earphones	Measurement
Immunological test procedure	CPT4 95065 Direct nasal mucous membrane test	Measurement
Semen analysis	HCPCS G0027 Semen analysis; presence and/or motility of sperm excluding huhner	Measurement
Cosmetic	SPL 8491b3d2-1edc-4694-adf7-b9649edb00eb ONWARD CHAPSTICK - lipstick	Observation
Evaluation or assessment	HCPCS G8930 Assessment of depression severity at the initial evaluation	Observation
Lifestyle choice	ICD10 Z58.7 Exposure to tobacco smoke	Observation
Pathology report	HCPCS G9428 Pathology report includes the pt category and a statement on thickness and ulceration and for pt1, mitotic raten	Observation
Service without exact procedure	HCPCS H0030 Behavioral health hotline service, HCPCS H2016 Comprehensive community support services, per diem	Observation
Visit without procedure	HCPCS S9127 Social work visit, in the home, per diem	Observation
Anything called *ectomy	HCPCS S2350 Diskectomy, anterior, with decompression of spinal cord and/or nerve root(s), including osteophytectomy; lumbar, single interspace except some devices having *ectomy in description like HCPCS L8020 Breast prosthesis, mastectomy form	Procedure
Anything called *graphy	HCPCS C9733 Non-ophthalmic fluorescent vascular angiography	Procedure
Anything called *scopy	HCPCS S2070 Cystourethroscopy, with ureteroscopy and/or pyeloscopy; with endoscopic laser treatment of ureteral calculi (includes ureteral catheterization)	Procedure
Attention	ICD10 Z43.5 Attention to Cystostomy	Procedure
Brachytherapy	HCPCS G0458 Low dose rate (ldr) prostate brachytherapy services, composite rate	Procedure
Counseling	HCPCS H0005 Alcohol and/or drug services; group counseling by a clinician	Procedure
Detoxification	HCPCS H0009 Alcohol and/or drug services; acute detoxification (hospital inpatient)	Procedure
Examination	ICD10 Z01.0 Examination of eyes and vision	Procedure
Imaging	CPT4 91110 Gastrointestinal tract imaging, intraluminal	Procedure

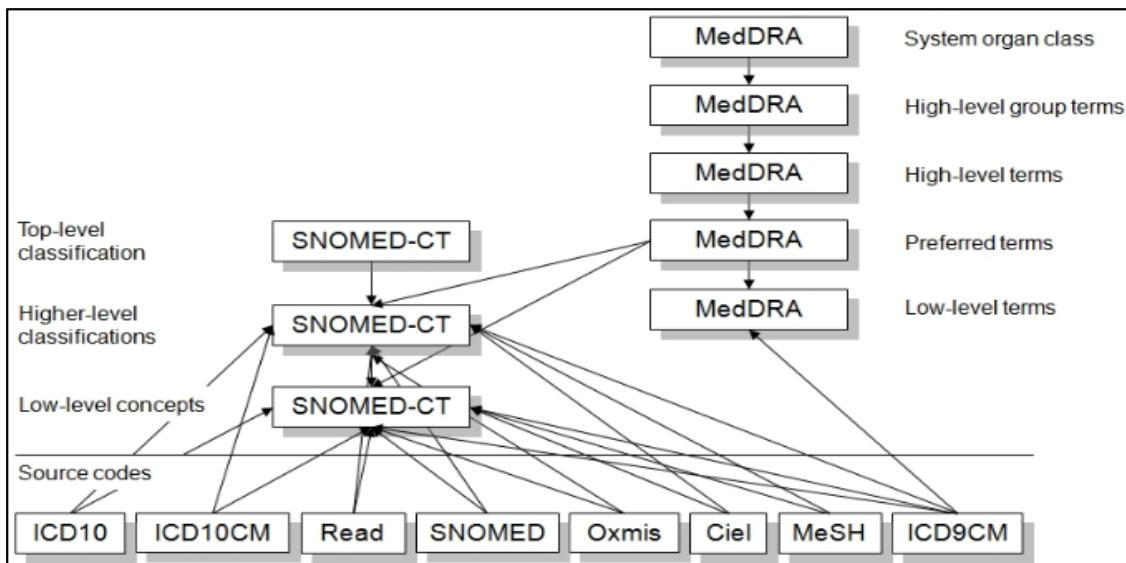
Category	Example	Domain
	(eg, capsule endoscopy), esophagus through ileum, with interpretation and report	
Immunotherapy	CPT4 95115 Professional services for allergen immunotherapy not including provision of allergenic extracts; single injection	Procedure
Management	ICD10 Z30.9 Contraceptive management, unspecified	Procedure
Monitoring	HCPCS 59050 Fetal monitoring during labor by consulting physician (ie, non-attending physician) with written report; supervision and interpretation	Procedure
Physician consult	CPT4 1014276 Pathology consultation during surgery	Procedure
Procedure Drug (without exact drug information and no RxNorm mapping)	HCPCS J8498 Antiemetic drug, rectal/suppository, not otherwise specified	Procedure
Service with exact procedure	HCPCS H0005 Alcohol and/or drug services; group counseling by a clinician	Procedure
Surgical pathology (gross or microscopic)	CPT4 88300 Level I - Surgical pathology, gross examination only	Procedure
Visit with exact procedure	HCPCS S9097 Home visit for wound care	Procedure

### 4.3.2. 병태 도메인(Condition domain)

- Condition이란 어떤 사람에 대하여 의료 제공자가 관찰하거나 환자가 보고한 진단, 징후 또는 증상으로 명시된 질병이나 의학적 병태의 존재를 알려주는 기록임.
- Condition domain은 단순한 증상의 관찰과 완전하게 진단을 확정을 하는 것 또는 그 중간정도 단계를 뚜렷하게 구별하지 않는다는 점에 유의해야 함.
- Condition domain Concepts은 CONDITION\_OCCURRENCE 및 CONDITION \_ERA 테이블의 Condition\_concept\_id 필드와 OBSERVATION 테이블의 value\_as\_concept\_id 필드(observations의 경우 ‘History of’, ‘Family History of’ 등)에서만 사용해야 함.

#### 4.3.2.1. 전체 구조(Overall Structure)

- Condition Domain 의 전체 구조



- 표준 개념은 SNOMED 용어에서 가져옴. Condition Domain의 모든 유효한 (invalid\_reason is Null) SNOMED 개념은 표준 개념이므로, CONDITION\_OCCURRENCE 및 CONDITION\_ERA 테이블의 Condition\_concept\_id에서 사용할 수 있음.
- SNOMED는 위계구조를 가진 용어임. 따라서 SNOMED 개념을 클래스 개념으로 사용할 수도 있음. 즉, CONCEPT\_ANCESTOR 테이블에 있는 모든 SNOMED 개념의 하위 항목이 쿼리에서 의미론적으로 정확하게 매치될 것으로 예상할 수 있음.
- MEDRA는 추가 분류체계로 사용됨. MedDRA와 SNOMED 사이의 연결은 동등한 SNOMED 개념에 대한 MedDRA의 선호 용어(preferred term) 개념 수준에서 이루어짐.
  - 따라서 MedDRA 개념도 데이터를 쿼리하는데 사용될 수 있으며, CONCEPT\_ANCESTOR 테이블에 있는 MedDRA의 개념 역시 SNOMED 하위 항목에 정확하게 의미론적으로 일치할 것을 예상할 수 있음.
  - 그러나 MedDRA의 하위 개념은 SNOMED에 매핑되지 않았음.
  - 이러한 선택적 매핑으로 인해 MedDRA 선호 용어 수준보다 일반적인 SNOMED 개념은 CONCEPT\_ANCESTOR 쿼리에서 검색할 수 없음.

### 4.3.2.2. 관계(Relationships)

- 모든 소스 vocabularies는 SNOMED 표준 개념에 매핑되며, 이는 용어 소스 (Read, CIEL, MeSH)에 의해 제공되거나 용어팀에 의해 수행됨.

#### ■ ICD10, ICD10CM에서 SNOMED로의 매핑

- ICD10CM과 SNOMED 간의 매핑을 위한 입력으로 사용된 다양한 소스 (UMLS, NHS)는 용어 팀에 의해 수동으로 완성됨.
- ICD10에는 주로 Condition 개념이 포함되지만 행위(Procedure), 관찰 (Observation), 장치(Device), 측정(Measurement) 또는 혼합 개념(mixed Concepts)도 포함되어 있다는 점에 유의할 것.
- 또한, 개념\_코드가 동일한 ICD10 및 ICD10CM 개념은 동일한 의미론적 정의를 가질 수도 있고 가지지 않을 수도 있기 때문에 ICD10CM(US 버전) 개념에는 이 매핑을 사용하면 안 된다는 점에 유의할 것.

#### ■ Read에서 SNOMED로의 매핑

- Read는 SNOMED에 대한 매핑 정보를 포함하여 배포함. 이 매핑을 사용하며, 사용되지 않는 SNOMED 대상 개념은 유효한 업데이트로 대체됨.  
참고: SNOMED-Int에 적절한 개념이 없거나 중복이 제거되지 않은 경우 많은 Read 코드가 SNOMED-UK 개념에 매핑됨.

#### ■ OXMIS에서 SNOMED로의 매핑

- 이 매핑은 용어팀에서 수행됨.

#### ■ CIEL에서 SNOMED로의 매핑

- CIEL은 SNOMED에 대한 매핑 정보를 포함하여 배포함. 이 매핑이 사용되며, 사용되지 않는 SNOMED 대상 개념은 유효한 업데이트로 대체됨.

#### ■ MeSH에서 SNOMED로의 매핑

- UMLS 배포에 포함된 MeSH에서 SNOMED로의 매핑이 사용되었음.

#### ■ ICD9CM에서 SNOMED로의 매핑

- ICD9CM과 SNOMED 간의 매핑을 위한 입력으로 사용된 다양한 소스(UMLS, HLI))는 용어 팀에 의해 수동으로 완성됨.
- ICD9에는 주로 Condition 개념이 포함되지만 행위(Procedure), 관찰 (Observation), 장치(Device), 측정(Measurement) 또는 혼합 개념(mixed Concepts)도 포함되어 있다는 점에 유의할 것.

## ■ ICD9CM에서 MedDRA로의 매핑

- ICD9CM과 MedDRA 간의 매핑을 위한 입력으로 사용된 다양한 소스(UMLS, MSSO)는 용어 팀에 의해 수동으로 완성됨.  
참고: 이 매핑은 역사적 이유와 편의를 위해 제공되지만 Condition 영역의 계층 구조에는 참여하지 않음.

## ■ Internal relationships

- SNOMED 및 MedDRA가 제공하는 모든 관계가 로드되지만 relationship\_id 형식 지정 규칙을 따르고 양방향 매핑을 허용하기 위해 이름을 변경함. SNOMED 및 MedDRA를 참조할 것.

### 4.3.3. 장치 도메인(Device Domain)

- 장치는 질병을 진단, 예방 또는 치료하기 위해 사용되는 장치, 이식물, 시약 등을 가리키며, 신체의 내·외부에 화학적 작용을 통해 목적을 달성하는 것은 장치가 아님(이것은 약물).
- 장치 도메인 개념은 DEVICE 테이블의 device\_concept\_id 필드에서만 사용해야 함.

#### 4.3.3.1. 전체 구조(Overall Structure)

- 장치 도메인의 개념은 Read, SNOMED, HCPCS 및 CIEL 용어에서 찾을 수 있음. SNOMED 또는 HCPCS 디바이스만을 표준 개념으로 보며, Read 및 CIEL의 장치는 SNOMED에 매핑됨.

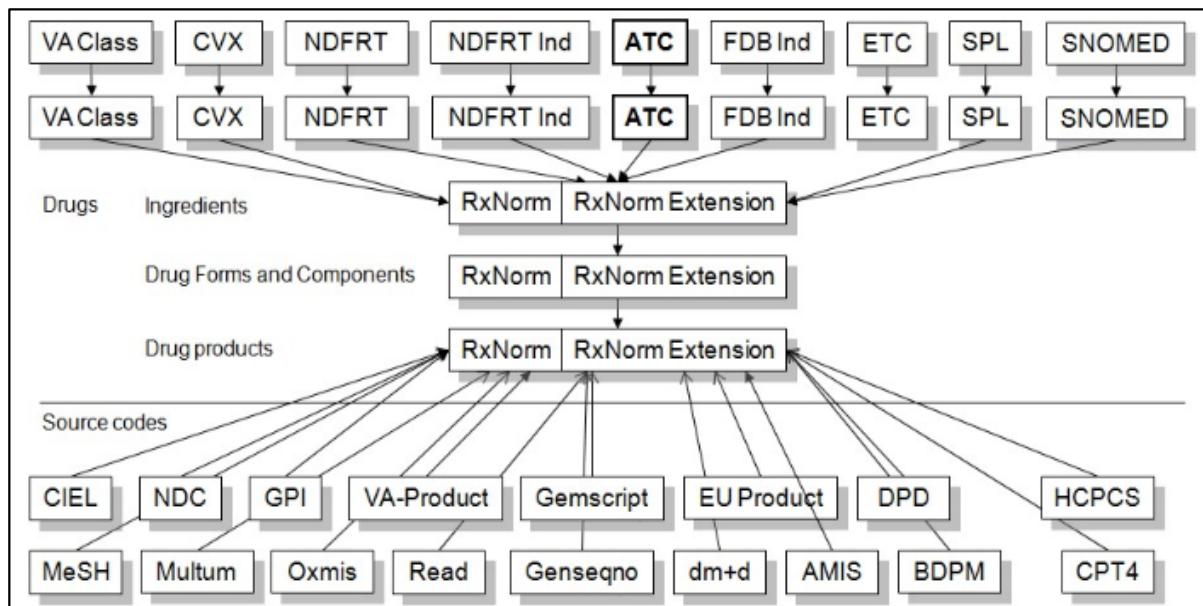
### 4.3.4. 약물 도메인(Drug Domain)

- 약물 노출 도메인 개념은 섭취하거나 체내에 다른 방법으로 투여할 때 약물의 이용에 관한 정보를 획득함. 약물은 사람에게 투여할 때 특정한 생화학 효과를 발휘하도록 만들어진 생화학적 물질임. 다음 제품은 약물이 아니라 장치로 간주됨:
  - 방사성 의약품
  - 영상 조영제
  - 유아용 유동식을 포함한 영양제 및 보충제

- ; 실제로 약간 임의적이며 확신하기 어려운 경우가 있는데,
- ▶ parental use를 위한 소금 용액(환자에게 수분을 공급하고 이온 균형을 유지)은 약물로 간주하지만,
  - ▶ feeding patients에게 글루코스나 바이타민을 추가하는 것은 장치로 간주함.
  - 혈액에서 직접 파생된 제품(예: erythrocytes 또는 plasma)
  - 약물 도메인 개념은 DRUG\_EXPOSURE의 drug\_concept\_id, DRUG\_ERA 및 DOSE\_ERA(두 가지 모두 성분 수준) 테이블 또는 OBSERVATION 또는 MEASUREMENT 테이블의 value\_as\_concept\_id 필드(예: ‘Plasma 수준’과 같은 측정)에서 사용되어야 함.

#### 4.3.4.1. 전체 구조(Overall Structure)

- 분류(classifications)



- 약물 도메인은 원천 개념(below the horizontal line), 약물(표준 개념) 및 분류(분류 개념)로 구성됨. 원천 개념(Source Concept)은 다양한 데이터베이스에서 코딩 체계로 사용되는 비표준 개념임. 따라서, 그것들은 표준 개념에 매핑됨.
- 표준 개념은 RxNorm 및 RxNorm Extension 용어를 기반으로 함. 모든 표준 RxNorm 및 RxNorm 확장 개념은 drug\_concept\_id 필드에서 사용됨.
- RxNorm은 결합된 도메인 공간의 기반임. 모든 유효한 RxNorm 개념

(invalid\_reason is Null)은 표준 개념으로 사용됨. 그러나 RxNorm은 미국에서 판매되는 제품에 대한 약물 개념만 포함하고 있음. 국제적 제품에서 파생된 모든 개념은 RxNorm Extension 용어에서 사용할 수 있음.

- 다른 vocabulary\_id와는 별개로, RxNorm과 RxNorm Extension은 관계와 개념의 계보를 위한 동일한 규칙을 가진 복합 개념 공간을 형성함.
- RxNorm과 RxNorm 확장 개념은 위계구조를 가지고 있음. 따라서 개념들은 클래스 개념으로도 사용될 수 있음. 즉, CONCEPT\_ANCESTOR 테이블에 있는 표준 개념의 하위 항목이 쿼리에서 정확한 의미론적 매치될 것으로 예상 할 수 있음. 예를 들어 성분(ingredient)의 하위 항목에 대한 쿼리는 해당 성 분을 포함하는 모든 약품(임상 약품 또는 브랜드 약품)을 결과로 제시해야 함.
- 이 결합된 RxNorm/RxNorm Extension Drug Domain은 약물 분류 시스템임. 이것들은 다양한 용어에서 파생되었음. 참고: 분류 개념은 다른 분류에 존재할 수 있으므로 고유하다고 간주할 수 없음. 예를 들어, NSAID Analogs 는 ETC, VA Class 및 SNOMED Class와 매우 유사한 멤버십을 가지고 있음. 또한 참고: ATC에서만 RxNorm 및 RxNorm Extension 약물에 대해 정의되었으며, 따라서 모든 약물에 대한 범용 약물 분류 시스템으로써의 역할이 가능함. 그 외 분류는 RxNorm과 RxNorm Extension의 하위 세트를 위한 향 후 추가에 대해서만 정의됨.

#### 4.3.4.2. 개념 클래스(Concept Classes)

- Source and Classification Concepts는 source 용어에 한정됨(각 source 용어 참조). 약물 개념에 대한 개념 클래스는 RxNorm의 하위 집합을 기반으로 함.
- RxNorm 용어 유형(RxNorm Term Types)<sup>20)</sup>

Term Type (TTY)	Name	Concept Class
IN	Ingredient	Ingredient
PIN	Precise Ingredient	Ingredient(standard_concept = NULL)
MIN	Multiple Ingredients	Not implemented
BN	Brand Name	Brand Name

20)For more information: <https://www.nlm.nih.gov/research/umls/rxnorm/docs/2015/appendix5.html>

Term Type (TTY)	Name	Concept Class
SCDC	Semantic Clinical Drug Component	Clinical Drug Component
SCDF	Semantic Clinical Drug Form	Clinical Drug Form
SCDG	Semantic Clinical Dose Form Group	Not implemented
SCD	Semantic Clinical Drug	Clinical Drug or Quantified Clinical Drug
SBDC	Semantic Branded Drug Component	Branded Drug Component
SBDF	Semantic Branded Drug Form	Branded Drug Form
SBDG	Semantic Branded Dose Form Group	Not implemented
SBD	Semantic Branded Drug	Branded Drug or Quantified Branded Drug
GPCK	Generic Pack	Clinical Pack
BPCK	Brand Name Pack	Branded Pack
DF	Dose Form	Dose Form
DFG	Dose Form Group	Not implemented
PSN	Prescribable Name	Not implemented
SY	Synonym	Implemented in the CONCEPT_SYNONYM table
TMSY	Tall Man Lettering Synonym	Not implemented
ET	Dose Form Entry Term	Not implemented

- 액체 제제의 개념 클래스 ‘Clinical Drug’ 과 ‘Branded Drug’ (SCD 및 SBD)의 개념은 제공된 총 양을 포함하거나 포함하지 않고 해당 약물 강도를 농도로 제공함. 그것은 그들 중 많은 수가 쌍으로 존재한다는 것을 의미함. RxNorm은 다른 용어 유형을 사용하여 이러한 구분을 나타내지 않지만, OMOP은 정량 인자가 제공되는 약물에 대해 추가 개념 클래스로 ‘양이 산정된 임상 약물’ 과 ‘양이 산정된 브랜드 약물’ 을 정의함.
- RxNorm에서 파생된 개념 클래스 외에, 국제 코딩 체계는 그 패키지 또는 box 크기에 따라 약물을 관리함. 이는 일반적으로 사전 포장된 상품으로 판매되지만 미국의 대부분의 처방약들은 약국에서 포장된다는 사실 때문임. 이러한 box는 RxNorm팩과 다르다는 점에 유의해야 함. 팩은 약품 조합의 최소 단위를 나타내는 반면, box는 상품으로 판매되는 총량 단위를 정의함. Box 외에 국제 코딩 체계에서도 서로 다른 제조업체 또는 리셀러(공급업체)의 제품을 구별함. 이러한 코딩 체계가 RxNorm 시스템과 원활하게 작동하도록 하기 위해 RxNorm Extension 용어에서 추가 개념 클래스를 정의함:

Vocab	Name	Ingr e dient	Drug strength	Dose form	Brand name	Quantity factor	Box size	Pack content	Supplier
RxNorm	Ingredient	X							
	Clinical Drug Component	X	X						
	Clinical Drug Form	X		X					
	Clinical Drug	X	X	X					
	Branded Drug Component	X	X		X				
	Branded Drug Form	X		X	X				
	Branded Drug	X	X	X	X				
	Clinical Pack	X						X	
	Branded Pack	X			X			X	
Rxnorm (not explicit)	Quantified Clinical Drug	X	X	X		X			
	Quantified Branded Drug	X	X	X	X	X			
RxNorm Extension	Clinical Drug Box	X	X	X			X		
	Branded Drug Box	X	X	X	X		X		
	Quantified Clinical Box	X	X	X		X	X		
	Quantified Branded Box	X	X	X	X	X	X		
	Clinical Pack Box	X	X	X			X		
	Branded Pack Box	X	X	X	X		X		
	Marketed Product	X	X	X	optional	optional	optional	optional	X

- o ① Concept Classes 시스템은 완전히 정상화됨. 즉, 정의된 각각의 브랜드 약물에 대해 동등한 임상 약물(Clinical Drug), 임상 약물 구성 요소(Clinical Drug Component), 브랜드 약물 구성 요소(Branded Drug Component), 임상 약물 형태(Clinical Drug Form) 및 브랜드 약물 형태(Branded Drug Form)를 사용할 수 있음. 여러 성분 조합 약물의 경우 동일한 방법으로 정의하는데, 임상 약물 구성요소에 대해서는 그렇지 않다는 점에 유의해야 함. 이것들은

오직 단일 요소 개념으로만 존재하며, 그 이유는 RxNorm의 기술 설명서에서 논의되지 않았음. 재료는 단일 성분으로만 정의되며, RxNorm의 MIN(다중 성분)에 대해서는 구현되지 않음.

#### 4.3.4.3. 관계(Relationships)

- 다른 것과 마찬가지로 모든 원천 개념은 RxNorm이나 RxNorm 확장 개념에 매핑됨.
- 매핑은 용어 소스로 부터 구성되거나 어떤 경우에는 어휘 팀에 의해 진행됨. 일반적으로, 매핑은 동등한 개념 클래스 간에 있으며, 가능하다면 임상 또는 브랜드 약물에 대한, 또는 계층 내 더 상세한 개념 클래스들에 대한 것들임.
- 그러나, 특정 상황이나(예를 들어, Procedure Drugs을 매핑할 때) 보다 일반적인 개념을 사용할 수 있는 원천 개념에 대해 알려진 정보가 불충분한 경우에는 등가적(equivalent) 매핑이 불가능함. 이에 따른 비등가적 (non-equivalent) 매핑은 다음과 같이 수행됨:
  - 환자 또는 임상 약물 구성요소에 대한 1:n 매핑
  - 다른 모든 개념 클래스에 대한 1:1 매핑
    - . 약물 개념 간의 내적 관계(internal relationships)은 RxNorm 관계에서 도출되고 RxNorm Extension 용어의 부가적 개념 클래스를 cover하기 위해 확장되었음(‘속성(attribute)’이 없는 ‘Drug’은 ‘Clinical’이나 ‘Branded’를 의미함):

Source	Original name	relationship_id	Connecting Concept Classes
RxNorm	consists_of	Consists of	Drug or Quantified Drug and Drug Component
	constitutes	Constitutes	Drug Component and Drug or Quantified Drug
	contained_in	Contained in	Drug or Quantified Drug and Pack
	contains	Contains	Pack and Drug or Quantified Drug
	dose_form_of	RxNorm dose form of	Dose Form and Drug, Drug Form
	form_of	Form of	(Precise) Ingredient and Ingredient
	has_dose_form	RxNorm has dose form	Drug, Drug Form and Dose Form
	has_form	Has form	Ingredient and (Precise) Ingredient
	has_ingredient	RxNorm has ing	Clinical Drug Form/Component and Ingredient, and Branded Drug Form/Component and Brand Name. Not implemented between Branded Drug

Source	Original name	relationship_id	Connecting Concept Classes
RxNorm Extension			and Brand Name
	has_precise_ingredient	Has precise ing	Brand Name/Clinical Drug Component and (Precise) Ingredient
	has_quantified_form	Has quantified form	Drug and Quantified Drug
	has_tradename	Has tradename	Clinical and Branded Concept Class equivalents
	has_tradename	Has brand name	Ingredient and Brand Name
	ingredient_of	RxNorm ing of	Ingredient and Clinical Drug Form/Component, and Brand Name and Branded Drug Form/Component. Not implemented between Branded Drug and Brand Name
	inverse_isa	RxNorm inverse is a	Drug Form and Drug
	isa	RxNorm is a	Drug and Drug Form
	precise_ingredient_of	Precise ing of	(Precise) Ingredient and Brand Name/Clinical Drug Component
	quantified_form_of	Quantified form of	Quantified Drug and Drug
	reformulated_to	Reformulated in	Brand Name and Brand Name(name changes)
	reformulation_of	Reformulation of	Brand Name and Brand Name(name changes)
	tradename_of	Tradename of	Branded and Clinical Concept Class equivalents
	tradename_of	Brand name of	Brand Name and Ingredient
	-	Concept replaced by	Concepts in RXNATOMARCHIVE with merged_to_rxci information
	-	Concept replaces	Concepts in RXNATOMARCHIVE with merged_to_rxci information
RxNorm Extension	-	Available as box	Drug and Drug Box
	-	Box of	Drug Box and Drug
	-	Has marketed form	Drug and Marketed Product
	-	Marketed form of	Marketed Product and Drug
	-	Has supplier	Drug and Supplier
	-	Supplier of	Supplier and Drug

- 약물 및 분류 간의 관계와 분류 내의 관계에 대한 설명은 개별 용어 섹션을 참조할 것.

#### 4.3.4.4. 약물 강도(Drug Strength)

- Drug Domain에는 각 성분의 강도를 정의하는 추가 DRUG\_STRENGTH 테이

블이 있음. 이는 성분(Ingredient), 임상 약물 형태(Clinical Drug Form) 및 브랜드 약품 형태(Branded Drug Form), 투여량 양식(Dose Form), 브랜드 이름 및 공급업체를 제외한 모든 약물 개념 등급에 대해 정의됨.

#### 4.3.5. 민족성 도메인과 용어(Ethnicity Domain and Vocabulary)

- 민족성과 인종 도메인은 한 개인의 민족적, 인종적 배경을 모두 결정함. 그것은 미국 행정부(백악관)의 정보 규제 담당관인 대통령 직속 행정관리예산국(OMB<sup>21)</sup>)의 권고에 따름. 이 분류는 모든 연방보고목적을 위해 인종 및 민족성에 관한 데이터를 유지, 수집 및 제시하기 위한 최소 표준을 제공함.
- 그것이 작용하는 방법은 다음과 같음: 인종 필드는 인종과 민족적 배경을 포함하지만, 민족성의 경우 ‘Hispanic or Latino’ (concept\_id=38003563)와 ‘Not Hispanic or Latino’ (concept\_id=38003564) 두 가지 범주만 있음. 즉, 두 가지 범주는 서로 직교하며 Latinos와 non-Latinos 모두 인종적 또는 민족적 배경을 가질 수 있음.
- 이것은 매우 미국 중심적인 솔루션이므로 미국 이외 국가 사용자에게 혼란을 줄 수 있음.
- 국내에서는 해당 자료를 수집하는 의료기관이 없어 현재까지 국내 CDM 참여 네트워크에서는 수집되는 자료가 없음.
- 그러나 국내 거주 외국인이 증가하면서 향후 각 기관의 수집 자료 보완이 필요하고 이차적으로 CDM 내에 관련 정보가 한국형으로 수집되어야 할 것임.

#### 4.3.6. 성별 도메인과 용어(Gender Domain and Vocabulary)

- 성별 영역은 생물학적, 생리학적 특성을 나타내는 사람의 성별에 대한 모든 개념을 나타냄. 실제로, 성별(PERSON 테이블에 있음)은 ‘gender’ 보다는 ‘sex’라고 불러야 함. 그 이유는 ‘gender’는 사회에서의 행동, 역할, 기대치, 활동을 의미하기 때문임.
- 도메인은 FEMALE(concept\_id=8532)과 MALE(concept\_id=8507)의 두 가지 표준 개념만 포함하고 있음. 많은 데이터 소스에 ‘Unknown’, ‘Refused to

---

21) Office of Management and Budget

tell’ , ‘Hermaphrodite’ , ‘transgender constellations( ‘male to female’ , etc.)’ 이 포함되어 있음. OMOP CDM의 현재 목적을 위해, 성별 개념은 환자의 생물학적 구성을 통해 충화하거나 생물학적 성의 영향을 위한 분석 결과를 조정하기 위해 사용됨. 따라서 다른 모든 성별은 concept\_id=0(알 수 없는 정보)으로 표시됨.

- 원천 개념으로부터의 매핑 외에 성별에 대해 정의된 관계는 없음. 그러나 대부분의 ETL은 매핑 테이블의 도움 없이 성별 개념을 직접 다름.

## V

# CDM 자료를 이용한 추정(Estimation)과 예측(Prediction)

## 5.1. ATLAS에서 Estimation 개요

- Cohort Definitions을 통해 target cohort, comparator cohort, outcome cohort를 정의 후 target cohort와 comparator cohort 사이의 outcome risk를 추정하는데 연구 설계 및 분석 모형을 지정해야 함

## 5.2. ATLAS에서 Estimation 실행 방법

### 5.2.1. Cohort 선정

- ATLAS에서 target cohort, comparator cohort, outcome cohort 설정

The screenshot shows the 'Population Level Effect Estimation' interface. At the top, it displays 'Graham replication: dabigatran vs warfarin for risk of ischemic stroke'. Below this are two tabs: 'Specification' (selected) and 'Utilities'. The main area is titled 'Choose your target cohort:' and contains the entry 'Graham replication: target cohort - dabigatran new users with prior atrial fibrillation' with a file icon and a delete icon. Next is 'Choose your comparator cohort:' with the entry 'Graham replication: comparator cohort - warfarin new users with prior atrial fibrillation' and a file icon and a delete icon. Finally, 'Choose your outcome cohort:' with the entry 'Graham replication: outcome cohort - incident ischemic stroke, observed in inpatient setting' and a file icon and a delete icon.

### 5.2.2. 모형 선정

- ATLAS에서 제공하는 분석모형은 Logistic regression, Poisson regression, Cox proportional hazards 선택 가능

The screenshot shows the 'Specify the statistical model used to estimate the risk of outcome between target and comparator cohorts:' section. A dropdown menu is open, showing 'Select a model...' at the top, followed by three options: 'Logistic regression', 'Poisson regression', and 'Cox proportional hazards'. The 'Logistic regression' option is highlighted with a blue background.

### 5.2.3. 추적 및 워싱아웃 기간 선정

- ATLAS에서 추적 기간 시작 시점과 끝 시점을 선정할 수 있고 워싱아웃 및 최소 추적기간 설정 가능

Define the time-at-risk window start, relative to target/comparator cohort entry:

days from cohort start date

Define the time-at-risk window end:

days from  ▾

Minimum washout period applied to target and comparator cohorts:

Minimum required days at risk, applied to target and comparator cohorts:

▾

### 5.2.4. 환자 선정

- ATLAS에서 target cohort, comparator cohort 모두 선정된 경우 제외 가능하며 대상자 선정 이전에 결과 해당하는 경우 제외할 수 있음.

Remove patients who enter both cohorts?

▾

Remove patients who have observed the outcome prior to cohort entry?

▾

### 5.2.5. Propensity score matching

- ATLAS에서 Propensity score matching 설정할 수 있으며 baseline covariates를 Click으로 설정 가능

Use propensity score adjustment as a confounding adjustment strategy for baseline covariates?

Yes ▾

Which types of baseline covariates do you want to include in the propensity score model?

- Demographics
  - Gender
  - Age group (5-year bands)
  - Index year
  - Index month
  - Race
  - Ethnicity
- Conditions
  - In prior 30d
  - In prior 365d
  - In prior 180d within inpatient setting
  - All time prior
  - Overlapping index date
- Drugs
  - In prior 30d
  - In prior 365d
  - All time prior
  - Overlapping index date
- Procedures
  - In prior 30d
  - In prior 365d
- Measurement
  - Existence in prior 30d
  - Existence in prior 365d
- Risk scores
  - Charlson
  - CHADS2
  - DCSI

- ATLAS에서 baseline covariates 추가 및 삭제 변수 Concept set에서 선택 가능

**Choose a Concept Set**

Repository Concept Sets:				
Show 10 entries	Filter Repository Concept Sets:	New Concept Set		
Id	Title	Created	Modified	Author
10	Warfarin test	12/18/2018 9:14:53 AM	12/18/2018 9:14:53 AM	anonymous
4	valsartan test1	12/17/2018 2:11:37 PM	12/18/2018 10:23:59 AM	anonymous
7	Other ARB's tutorial : rhk	12/17/2018 2:23:08 PM	12/17/2018 2:30:19 PM	anonymous
8	other arb	12/17/2018 2:23:29 PM	12/17/2018 2:29:40 PM	anonymous
9	other ARB	12/17/2018 2:23:43 PM	12/17/2018 2:29:15 PM	anonymous
5	Valsartan test2	12/17/2018 2:19:29 PM	12/17/2018 2:23:22 PM	anonymous
2	valsatan tutorial:rhk	12/17/2018 2:10:30 PM	12/17/2018 2:21:46 PM	anonymous
3	valsartan_s	12/17/2018 2:10:35 PM	12/17/2018 2:10:35 PM	anonymous
296	pmrpmr_visit	04/24/2019 5:41:22 PM	04/24/2019 5:42:10 PM	anonymous
295	pmr준민_visit	04/24/2019 5:40:42 PM	04/24/2019 5:41:55 PM	anonymous

Showing 1 to 10 of 268 entries      Previous **1** 2 3 4 5 ... 27 Next

○  DCSI  
 What concepts do you want to include in baseline covariates in the propensity score model? (Leave blank if you want to include everything)

What concepts do you want to exclude from baseline covariates in the propensity score model? (Leave blank if you want to include everything)

- ATLAS에서 trim
  - ‘by Percentile’ 은 Propensity score의 Bounds 값 percentile 이후 끝 값 제외
  - ‘to Equipoise’ 은 Propensity score의 Bounds 값 이후 끝 값 제외

How do you want to trim your cohorts based on the propensity score distribution?

**to Equipoise**

**None**

**by Percentile**

**to Equipoise**

Bounds (1-100%):  
**5**

- ATLAS에서 ‘Matching’ 및 ‘Stratification’ 선정 가능
  - ‘Matching’ 은 Propensity score를 이용하여 매칭 진행
  - ‘Stratification’ 은 strata 개수를 정한 후 Propensity score를 이용하여 Stratification 분석 진행

Do you want to perform matching or stratification?

Matching	▼
No matching/stratification	
Matching	
Stratification	

- ATLAS에서 Target patient 대비 Comparator patients 배수 선정

How many comparator patients do you want to select for each target patient (within a defined caliper)?

1
---

## 5.2.6. 최종 모형 선정

- ATLAS에서 최종 모형의 baseline covariates 포함 여부 선정 가능

Do you want to adjust for baseline covariates in the outcome model?

No ▼
------

## 5.2.7. Negative control outcomes

- ATLAS에서 Empirical calibration 진행 시 Negative control outcomes 설정 가능

Id	Title	Created	Modified	Author
10	Warfarin test	12/18/2018 9:14:53 AM	12/18/2018 9:14:53 AM	anonymous
4	valsartan test1	12/17/2018 2:11:37 PM	12/18/2018 10:23:59 AM	anonymous
7	Other ARB's tutorial : rhk	12/17/2018 2:23:08 PM	12/17/2018 2:30:19 PM	anonymous
8	other arb	12/17/2018 2:23:29 PM	12/17/2018 2:29:40 PM	anonymous
9	other ARB	12/17/2018 2:23:43 PM	12/17/2018 2:29:15 PM	anonymous
5	Valsartan test2	12/17/2018 2:19:29 PM	12/17/2018 2:23:22 PM	anonymous
2	valsatan tutorial:rhk	12/17/2018 2:10:30 PM	12/17/2018 2:21:46 PM	anonymous
3	valsartan_s	12/17/2018 2:10:35 PM	12/17/2018 2:10:35 PM	anonymous
296	pmrpmr_visit	04/24/2019 5:41:22 PM	04/24/2019 5:42:10 PM	anonymous
295	pmr준민_visit	04/24/2019 5:40:42 PM	04/24/2019 5:41:55 PM	anonymous

Show 10 entries Filter Repository Concept Sets:

New Concept Set

to include everything)  ant to include everything)

Showing 1 to 10 of 268 entries Previous      ...  Next

NO  What outcomes would you like to use as your negative controls? These are concepts known not to be associated with either the target or comparator group, such that we can assume the true relative risk should equal 1. These negative control outcomes will be used for empirical calibration.

## 5.2.8. 저장

- ATLAS에서 모형 저장

Population Level Effect Estimation

Graham replication: dabigatran vs warfarin for risk of ischemic stroke

Specification Utilities

## 5.2.9. R 코드 복사

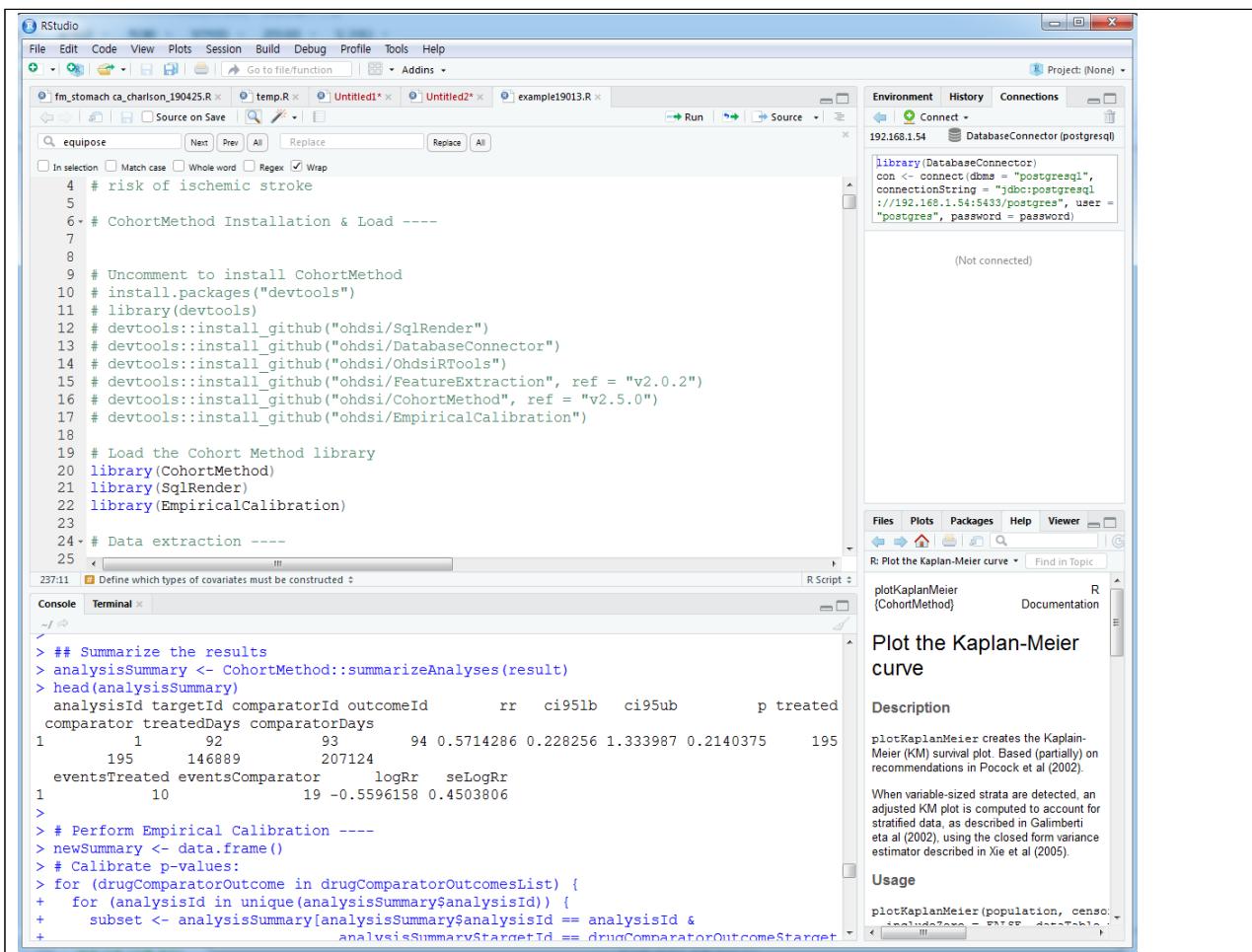
- ATLAS에서 Estimation → Utilities → R Code → Copy to Clipboard → R studio 실행

The screenshot shows the 'Population Level Effect Estimation' interface. At the top, there is a title bar with the text 'Population Level Effect Estimation' and a subtitle 'Graham replication: dabigatran vs warfarin for risk of ischemic stroke'. Below the title bar are two tabs: 'Specification' and 'Utilities', with 'Utilities' being the active tab. Under the 'Utilities' tab, there are four buttons: 'Print Friendly', 'R Code', 'Import', and 'Export'. The 'R Code' button is highlighted with a blue background. Below these buttons is a section titled 'Multi-Analysis Approach'. Within this section, there is a large blue button labeled 'Copy To Clipboard' with a white icon of a clipboard. A red lightning bolt-shaped arrow points from the text 'Copy To Clipboard' in the heading to this button. Below the button is a code block containing the following R script:

```
# Study: ----  
# Graham replication: dabigatran vs warfarin for  
# risk of ischemic stroke  
  
# CohortMethod Installation & Load ----
```

## 5.2.10. R Studio 실행

- R studio 실행하여 R script에 코드를 붙여놓기
- connectionDetails의 서버 IP 주소, user ID 및 password 등 설정
- cdmDatabaseSchema, ..., outputFolder 등 설정
- 선택 코드 드래그 후 ‘Run’ 실행



## 5.3. 모형에 따른 결과 해석

### 5.3.1 Logistic regression

- 로지스틱 회귀분석은 결과변수가 0 또는 1을 가지는 경우 추정 및 비교 분석 할 수 있고 예를 들어 사망/생존, 질병 유/무인 경우로 의학연구에서 많이 사용
- 다음 예제는 과거 심방세동을 앓은 환자 중 Dabigatran 처방받은 환자와 와파린을 처방받은 환자 대상으로 허혈성 뇌졸중의 위험률 추정 결과

> head(analysisSummary) - R studio

analysis Id	target Id	comparator Id	outcome Id	rr	ci95lb	ci95ub	p
1	92	93	94	<b>0.556</b>	0.247	1.181	<b>0.141</b>
treated	comparator	treated Days	comparator Days	events Treated	events Comparator	logRr	seLogRr
195	195	NA	NA	10	18	-0.588	0.399

- summary 결과를 보면 와파린 처방군에 비해 Dabigatran 처방군의 허혈성 뇌출중 위험률이 0.556배로 약 44%로 줄어든다고 추정 가능
- p-value = 0.141로 통계적으로 유의하지 않으므로 의미해석하지 않음

### 5.3.2 Poisson regression

- 포아송 회귀분석은 단위 시간 안에 어떤 사건이 몇 번 발생할 것인지를 추정 및 비교 분석할 수 있고 사건의 횟수가 희박한 경우 빈번하게 사용
- 다음 예제는 과거 심방세동을 앓은 환자 중 Dabigatran 처방받은 환자와 와파린을 처방받은 환자 대상으로 허혈성 뇌출중의 위험률 추정 결과

> head(analysisSummary) - R studio

analysis Id	target Id	comparator Id	outcome Id	rr	ci95lb	ci95ub	p
1	92	93	94	<b>0.450</b>	0.175	1.084	<b>0.086</b>
treated	comparator	treated Days	comparator Days	events Treated	events Comparator	logRr	seLogRr
195	195	151410	219255	10	19	-0.800	0.465

- summary 결과를 보면 와파린 처방군에 비해 Dabigatran 처방군의 허혈성 뇌출중 위험률이 0.450배로 약 55%로 줄어든다고 추정 가능
- p-value = 0.086으로 통계적으로 유의하지 않으므로 의미해석하지 않음

### 5.3.3 Cox regression

- 콕스 회귀분석은 사건 발생 기간 및 유무를 아는 경우 추정 및 비교 분석할 수 있고 의학연구에서 많이 사용
- 다음 예제는 과거 심방세동을 앓은 환자 중 Dabigatran 처방받은 환자와 와파린을 처방받은 환자 대상으로 허혈성 뇌졸중의 위험률 추정 결과

> head(analysisSummary) - R studio

analysis Id	target Id	comparator Id	outcome Id	rr	ci95lb	ci95ub	p
1	92	93	94	<b>0.571</b>	0.228	1.334	<b>0.214</b>
treated	comparator	treated Days	comparator Days	events Treated	events Comparator	logRr	seLogRr
195	195	146889	207124	10	19	-0.560	0.450

- summary 결과를 보면 와파린 처방군에 비해 Dabigatran 처방군의 허혈성 뇌졸중 위험률이 0.571배로 약 43%로 줄어든다고 추정 가능
- p-value = 0.214로 통계적으로 유의하지 않으므로 의미해석하지 않음

## VI 시스템 구성현황 및 Table Layout

### 6.1 시스템 개요

#### 6.1.1. 시스템 구성현황

구분	서버종류	OS	CPU	메모리	디스크	DBMS	ETL
분석 서버	HP DL360G10	CentOS 7.5	2.6Ghz 16core*2	128GB	1.9T * 2ea		
DBMS 서버			2.3Ghz 12core*2	64GB	1.9T * 4ea	PostgreSQL 10.5	nifi

#### 6.1.2. 분석서버

구분	사양	규격	비고
1 Model	HP DL360G10	2U rack server	ETL 수행 및 분석용 서버 1대 구성
2 Prcessors	G-6142 2.6Ghz	CPU * 2EA (16Core)	Intel
3 Memory	128GB, 4 x 32GB	PC4-2666V-R	-
4 SSD	HP 240GB SSD x 2EA HP 192TB SSD x 2EA	SATA 6Gbps	Mirror 1+0 (OS영역) Mirror 1+0 (Data 영역)
5 HDD	-	-	-
6 Raid Controller	HPE Smart Array P408i-a 2GB	12G RAID CONTROLLER	-
7 Management	RESTful API iLO 5 관리엔진	Web base	웹 기반 서버 관리 툴
8 Networking option	10Gb 2P Network card	* 1	-
9 PowerSupply	Dual 800W	100 ~ 240 V 입력 전원 대비 출력 전력 전환 효율 94%	서버 전력 사용에 따라
10 OS	CentOS	7.5	

#### 6.1.3. DB 서버

구분	사양	규격	비고
1 Model	HP DL360G10	2U rack server	CDM DB 서버 1대 구성

구분		사양	규격	비고
2	Prcessors	G-5118 2.3Ghz	CPU * 2EA (12Core)	Intel
3	Memory	64GB, 2 x 32GB	PC4-2666V-R	-
4	SSD	HP 240GB SSD x 2ea HP 1.92TB SSD x 2ea HP 1.92TB NVMe * 2ea	SATA 6Gbpsx4	Mirror 1+0 (OS영역) Mirror 1+0 (Data 영역) Mirror 1+0 (Data 영역)
5	HDD	-	-	-
6	Raid Controller	HPE Smart Array P408i-a 2GB	12G RAID CONTROLLER	-
7	Management	RESTful API iLO 5 관리엔진	Web base	웹 기반 서버 관리툴
8	Networking option	10Gb 2P Network card	* 1	-
9	PowerSupply	Dual 800W	100 ~ 240 V 입력 전원 대비 출력전력 전환 효율 94%	서버 전력 사용에 따라
10	OS	CentOS	7.5	
11	DB	PostgreSQL 10.5		

## 6.2. 테이블 목록

CDM Table명	테이블 설명
Person	환자의 인종, 생년월일, 성별 등의 기본정보를 알 수 있는 테이블
Visit_occurrence	환자가 병원에 방문한 기록으로 방문시작일과 방문종료일을 알 수 있는 테이블
Observation_period	환자가 병원에 내원하여 치료가 완료되기까지의 기간을 알 수 있는 테이블
Procedure_occurrence	진단 및 치료 목적을 가지고 의사에 의해 처방 및 지시된 활동 또는 프로세스를 알 수 있는 테이블
Drug_exposure	환자에게 처방된 약 정보로 처방약, 약에 노출된 일시, 중지사유, 처방일수, 처방용량, 투약방법, 복용법, 단위 등을 알 수 있는 테이블
Condition_occurrence	환자의 진단명, 건강상태, 진단일시, 진단의사 등을 알 수 있는 테이블
Device_exposure	기기 및 각종 치료재료 사용 정보를 알 수 있는 테이블
Measurement	체계적인 검사로부터 얻어진 측정값으로 측정일, 측정값 최대/최소 범위, 측정 값, 측정단위 등을 알 수 있는 테이블
Specimen	검체 채취일시, 양, 위치, 단위 등을 알 수 있는 테이블
Observation	문진, 과거력, 알러지 등의 정보를 알 수 있는 테이블
Death	환자의 사망일, 원인을 알 수 있는 테이블
Note	자유 기술문으로 구성된 각종 기록지의 정보를 알 수 있는 테이블
<b>Health System data</b>	
Location	물리적 위치 또는 주소정보를 캡처하는 일반적인 방법을 나타냄
Care_site	의료서비스가 시행되는 조직 단위의 고유한 목록 포함(사무실 병동, 병원, 진료소 등)
Provider	의료 서비스제공자 식별 목록 포함
<b>Meta-data</b>	
CDM_source	원천 데이터베이스를 OMOP CDM으로 변환하는데 사용되는 소스에 대한 세부 정보 포함 (sourcename, holder, source description)
<b>Health economics</b>	
Cost	환자가 지불하는 비용, 보험 지급자가 지불하는 비용 등을 알 수 있는 테이블
Payer_plan_period	환자의 보험 종류와 적용기간의 정보를 알 수 있는 테이블
Cohort	지속 시간 동안 선정 기준으로 포함된 특정 세트를 충족하는 대상의 집합으로 파생된 레코드가 있음
Cohort_attribute	시간의 기간 동안 선정 기준의 주어진 집합에 의해 정의 된 바와 같이, 집단 내의 각 주제와 관련된 속성 포함
Drug_era	환자들이 특정 약성분에 노출된 것으로 예상되는 시간의 길이를 정의한 레코드

CDM Table명	테이블 설명
Dose_era	환자가 특정 약성분의 일정 투여량에 노출된 것으로 가정되는 시간의 길이를 정의한 레코드
Condition_era	환자가 처방된 진단상태를 갖는 것으로 가정되는 시간의 길이를 정의한 레코드
Concept	OMOP에서 제공하는 표준화된 용어에 대한 고유코드(concept_id) 및 정보
Concept_class	제공된 vocabulary 내에 포함된 Concept을 구별하기 위해 사용되는 분류의 리스트
Concept_synonym	Concept에 대한 대체 이름과 설명을 저장하는 데 사용되는 테이블
Concept_ancestor	모든 표준 concept들 사이의 추론된 계층(hierarchical) 관계들을 정의한 레코드
Concept_relationship	두 개의 Concept들 간의 직접적인 관계와 관계의 본질(nature)을 정의한 레코드
Relationship	Concept relationship 테이블에서 임의의 두 concept을 연결하는데 사용될 수 있는 모든 허용 가능한 관계의 종류의 참조(reference) 리스트를 제공
Vocabulary	다양한 소스들로부터 모아진 vocabulary 리스트나 OMOP community에 의해 새로 생성된 리스트
Domain table	OMOP 공통데이터모델 내에 포함된 데이터 요소의 도메인들의 리스트
Drug_strength	특정 의약품의 양 또는 농도 및 특정 성분 관련 단위에 대한 구조화된 콘텐츠 포함
Source_to_concept_map	소스와 컨셉 사이의 관계를 정의하는 레코드
Cohort_definition	연관된 설명 및 구문을 통해 각각의 파생된 코호트를 정의하는 레코드
Attribute_definition	연관된 설명 및 구문을 통해 각각의 속성을 정의하는 레코드

### 6.3. 테이블 상세 목록

#### 1) Person - 환자의 기본정보를 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
PERSON_ID	환자 ID	integer	환자별로 부여되는 고유 식별자(unique identifier)
GENDER_CONCEPT_ID	성별	integer	외래키-CONCEPT 테이블 참조 환자의 고유한 성별 * Concept table에서 vocabulary_id= 'gender' 인 코드검색 예) MALE:8507, FEMALE:8532
YEAR_OF_BIRTH	출생년도	integer	원천 데이터의 생일에서 추출된 년도 정확한 출생년에 대한 데이터가 없고 연령데이터가 있는 경우, 이를 이용해서 출생년도를 추정할 수 있음.
MONTH_OF_BIRTH	출생월	integer	정확한 출생일이 기록되어 있는 원천 데이터가 있을 경우, 월(month)이 추출되어 이 필드에 저장됨.
DAY_OF_BIRTH	출생일	integer	정확한 출생일이 기록되어 있는 원천 데이터가 있을 경우, 일(day)이 추출되어 이 필드에 저장됨.
BIRTH_DATETIME	생년월일시	datetime	태어난 일시
RACE_CONCEPT_ID	인종	datetime	외래키-CONCEPT 테이블 참조 사람의 고유한 인종(유전적 인종, race) * Concept table에서 vocabulary_id= 'Race' 인 코드 검색 예) KOREAN:38003585, ASIAN:8515
ETHNICITY_CONCEPT_ID	민족성	integer	외래키-Standardized Vocabularies 의 민족성 개념 식별자 참조. 사람의 민족성 * Concept table에서 vocabulary_id= 'ethnicity' 인 코드 검색 → Hispanic or Latino:38003563, Not Hispanic or Latino:38003564
LOCATION_ID	거주지	integer	외래키-location table의 거주지(place of residency) 참조 상세 주소 정보가 저장
PROVIDER_ID	의료서비스 제공자	integer	외래키-provider 테이블 참조 주치의(primary care provider)
CARE_SITE_ID	요양기관	integer	외래키-care_site 테이블 참조 주요 치료 장소(care site)

CDM 컬럼		Type	Description
Eng	Kor		
PERSON_SOURCE_VALUE	환자 ID 원본값	varchar(50)	원천 데이터에서 사용되는 (암호화된) 환자 식별 키 원천 데이터와 다시 연계할 필요가 있을 경우 꼭 필요함.
GENDER_SOURCE_VALUE	성별 원본값	varchar(50)	원천 데이터 상에서 보이는 환자의 성별에 대한 코드 값. 원천 데이터의 환자의 성별은 표준 용어 내에 정의된 성별 컨셉과 매핑되며, 원천 데이터의 값이 참고용으로 이 필드에 저장됨.
GENDER_SOURCE_CONCEPT_ID	성별 소스코드 체계	Integer	외래 키-원천데이터에서 사용하는 성별 개념 코드.
RACE_SOURCE_VALUE	인종 원본값	varchar(50)	원천 데이터 상에서 보이는 환자의 인종(유전적 인종, race)에 대한 코드 값. 원천 데이터의 환자의 인종은 표준 용어 내에 정의된 인종 컨셉과 매핑되며, 원천 데이터 값이 참고용으로 이 필드에 저장됨.
RACE_SOURCE_CONCEPT_ID	인종 소스코드 체계	Integer	외래 키-원천 데이터에서 사용되는 인종 구분 컨셉 참조
ETHNICITY_SOURCE_VALUE	민족 원본값	varchar(50)	원천 데이터 상에서 보이는 환자의 민족성 (ethnicity)에 대한 코드 값. 원천 데이터의 환자의 민족성은 표준 용어 내 컨셉과 매핑되며, 원천 데이터의 값이 참고용으로 이 필드에 저장됨.
ETHNICITY_SOURCE_CONCEPT_ID	민족성 원본 개념 ID	Integer	외래 키-원천 데이터의 민족성 개념 참조

\* 변환데이터에 포함된 모든 환자가 포함되어야 함 (예: 건진, 외래, 응급, 입원 환자 등)

## 2) Visit\_occurrence - 환자가 병원에 내원한 기록을 알 수 있는 테이블

- EMR에서 해당 방문기록이 없는 검사 등이 정상적으로 발생 가능하므로(선행 제 후 검사 등), ETL 후 방문 건이 없으나 정상적인 데이터인 경우, 외래 방문한 것처럼 visit occurrence를 추가할 수 있음
- visit occurrence는 날짜 중복하여 발생할 수 있음

CDM 컬럼		Type	Description
Eng	Kor		
VISIT_OCCURRENCE_ID	방문 ID	integer	환자가 공급자를 방문한 건에 대한 고유 식별자
PERSON_ID	환자 ID	integer	외래 키-PERSON 테이블의 PERSON_ID를 참조 방문 레코드가 있는 환자에 대한 식별자 환자의 인구학적 상세 정보는 PERSON테이블에 저장됨.

CDM 컬럼		Type	Description
Eng	Kor		
VISIT_CONCEPT_ID	방문 유형	integer	외래키-Standardized Vocabularies의 visit Concept identifier 참조 *Concept table에서 domain_ID='Visit' 인코드 검색 예) InpatientVisit:9201, OutpatientVisit:9202
VISIT_START_DATE	방문 개시일자	date	내원시작일(연/월/일)
VISIT_START_TIME	방문 개시시간	datetime	내원시작일시(연/월/일/시)
VISIT_END_DATE	방문 마감일자	date	내원종료일(연/월/일) 하루 방문일 경우 내원시작일과 동일
VISIT_END_TIME	방문 마감시간	datetime	내원종료일시(연/월/일/시)
VISIT_TYPE_CONCEPT_ID	방문데이터소스의 유형(예: 청구,EMR )	integer	외래키-Standardized Vocabularies에 정의되어있는 개념 식별자. 내원기록을 가져온 원천데이터의 유형 *Concept table에서 vocabulary_ID='VisitType'인 코드검색
PROVIDER_ID	의료서비스 제공자	integer	외래키-provider 테이블의 방문과 관련된 provider 참조
CARE_SITE_ID	요양기관	integer	외래키-CARE_SITE 테이블의 (방문과 관련된) CARE_SITE 참조
VISIT_SOURCE_VALUE	방문 유형 원본값	integer	원천 데이터의 내원 구분 코드 값
VISIT_SOURCE_CONCEPT_ID	방문 유형 소스코드 체계	integer	외래키-원천 데이터에서 사용된 내원구분 개념 코드 참조
ADMITTING_SOURCE_CONCEPT_ID	내원 전 서비스 장소 ID	integer	외래키-Vocabulary의 기 정의된 개념 ; 병원에 방문하기 전 서비스를 받은 장소
ADMITTING_SOURCE_VALUE	내원 전 서비스 장소 원본값	varchar (50)	외래키-Vocabulary의 기 정의된 개념 ; 방문에 대한 admitting source를 반영한 Service 장소
DISCHARGE_TO_CONCEPT_ID	퇴원 후 서비스 장소 ID	integer	외래키-Vocabulary의 기 정의된 개념 ; 환자가 병원에서 퇴원 후 서비스를 받을 장소
DISCHARGE_TO_SOURCE_VALUE	퇴원 후 서비스 장소 원본값	varchar (50)	외래키-Vocabulary의 기 정의된 개념 ; 방문에 대한 discharge to source를 반영한 Service 장소
PRECEDING_VISIT_OCCURRENCE_ID	방문 ID	integer	원천 데이터의 병원에서 퇴원한 후 서비스를 받을 장소에 대한 코드

### 3) Observation\_period

- 환자가 병원에 내원하여 치료가 완료되기까지의 기간을 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
OBSERVATION_PERIOD_ID	관찰기간 ID	integer	각 관찰기간에 대한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-관찰 기간이 정의된 환자의 고유 식별자 환자의 인구학적 상세 정보는 PERSON테이블에 저장됨.
OBSERVATION_PERIOD_START_DATE	관찰 개시일자	date	원천 데이터에서 사용 가능한 관찰기간의 시작일
OBSERVATION_PERIOD_END_DATE	관찰 마감일자	date	원천 데이터에서 사용 가능한 관찰기간의 종료일
PERIOD_TYPE_CONCEPT_ID	관찰기간 데이터 소스 유형	integer	외래키-관찰 기간 정보의 소스를 반영하여 표준 용어(StandardizedVocabularies)에 정의되어 있는 개념 식별자 *Concept table에서 vocabulary_id= 'ObsPeriodType' 인 코드 검색 예)Period covering healthcare encounters:44814724

### 4) Procedure\_occurrence

- 환자가 병원에 내원하여 진단 및 치료 목적으로 의사에게 처방 및 지시되어 받은 모든 행위와 프로세스에 대해 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
PROCEDURE_OCCURRENCE_ID	시술/처치 ID	integer	각 시술 발생에 대하여 시스템이 생성한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-시술을 받은 환자에 대한 고유 식별자 환자의 인구학적 상세 정보는 PERSON테이블에 저장됨.
PROCEDURE_CONCEPT_ID	시술/처치 유형	integer	외래키-Standardized Vocabularies의 표준 시술 개념 식별자 참조
PROCEDURE_DATE	시술/처치일자	date	시술이 시행된 날짜 (연/월/일)
PROCEDURE_DATETIME	시술/처치일시	datetime	시술이 시행된 일시 (연/월/일/시)

CDM 컬럼		Type	Description
Eng	Kor		
PROCEDURE_TYPE_CONCEPT_ID	시술/처치 데이터소스 유형	integer	외래키-시술 레코드를 가져온 원천 데이터의 유형을 반영하는 StandardizedVocabularies에 정의되어있는 개념 식별자 *Concept table에서 vocabulary_ID='procedureType'인코드검색 예)EHR order list entry:38000275
MODIFIER_CONCEPT_ID	시술/처치 수식어	integer	외래키-시술 수식어(Modifier)에 대한 표준 개념 식별자(e.g. bilateral)
QUANTITY	시술/처치 량	integer	시술이 처방되거나 수행된 양
PROVIDER_ID	의료서비스 제공자	integer	외래키-PROVIDER 테이블의 의료제공자(provider) 식별자 :시술 수행 책임자 정보
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRENCE 테이블의 방문에 대한 고유 식별자 :시술이 수행된 방문 정보
VISIT_DETAIL_ID	방문상세 ID	integer	외래키-시술이 수행된 방문에 관한 VISIT_DETAIL테이블의 고유 식별자
PROCEDURE_SOURCE_VALUE	시술/처치 유형 원본값	varchar(50)	원천 데이터의 시술 코드 이 시술코드는 Standardized Vocabularies의 표준 시술 컨셉에 매핑되어 있음 원천 시술 코드는 일반적으로 ICD-9-Proc, CPT-4, HCPCS 또는 OPCS-4를 사용함
PROCEDURE_SOURCE_CONCEPT_ID	시술/처치 유형 소스코드 체계	integer	외래키-원천 데이터에서 사용된 코드를 참조하는 시술 개념
MODIFIER_SOURCE_VALUE	시술/처치 수식어구에 대한 원본값	varchar(50)	원천 데이터에서 보이는 시술에 대한 수식어구(Qualifier)에 대한 소스코드값 - 원천 데이터에 기록되어 있는 그대로 표현

## 5) Drug\_exposure - 처방되어 환자에게 노출된 약에 대해 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
DRUG_EXPOSURE_ID	의약품 처방 ID	integer	각 약물 사용 이벤트에 대한 시스템이 생성한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-약물 처방을 받은 환자에 대한 고유 식별자 환자의 인구학적 상세 정보는 PERSON테이블에 저장됨.
DRUG_CONCEPT_ID	처방 의약품	integer	외래키-약물 개념에 대한 Standardized Vocabularies의 표준 개념 식별자 참조
DRUG_EXPOSURE_START_DATE	의약품 개시일자	date	약물을 사용하게 된 처방 사례의 시작일(연/월/일) 처방 시작일, 처방 입력일 또는 약물이 주입되는 시술이 시행된 일자 등이 유효한 입력값임.

CDM 컬럼		Type	Description
Eng	Kor		
DRUG_EXPOSURE_START_DATETIME	의약품 개시일자	datetime	약물을 사용하게 된 처방 사례의 시작일시(연/월/일/시) 처방 시작일, 처방 입력일 또는 약물이 주입되는 시술이 시행된 일자 등이 유효한 입력값임.
DRUG_EXPOSURE_END_DATE	의약품 마감일자	date	약물을 사용하게 된 처방 사례의 종료일(연/월/일) 모든 원천 데이터들에서 사용 가능하지는 않음.
DRUG_EXPOSURE_END_DATETIME	의약품 마감일자	datetime	현재 약물 사용사례에 대한 종료일시(연/월/일/시) 모든 원본데이터들에서 사용 가능하지는 않음.
VERBATIM_END_DATE	종료일자	date	원천 데이터로부터 알려진 약물 노출의 종료일
DRUG_TYPE_CONCEPT_ID	의약품 데이터소스 유형	integer	외래키-약물 노출의 유형을 나타내는 Standardized Vocabularies의 기 정의된 개념 식별자 참조 원천 자료에서 약물 노출이 표현된 방법을 표시함 *Concept table에서 Vocabulary_ID= 'drug Type' 인 코드검색 예)Prescription written:38000177, Prescription dispensed in pharmacy: 38000175
STOP_REASON	중단 사유	varchar(20)	regimen 완료, 변경, 제거 등 약물 사용을 중단하게 된 사유
REFILLS	리필	integer	첫 처방 이후로 리필된 횟수 첫 처방은 0으로 표시하여 건수로 집계하지 않음.
QUANTITY	처방 용량	float	원본 처방이나 지급 기록(spensing record)에 기록되어 있는 약물의 양 * '1회처방량 x 1일처방횟수 x 투여일수'로 계산
DAYS_SUPPLY	총 투여일수	integer	원본 처방이나 지급기록에 기록되어 있는 약물의 처방일수
SIG	의약품 투여방법	clob	원본 처방전( 약 통에 기록되어 있는) 또는 지급 기록에 기록되어 있는 약물 복용 방법
ROUTE_CONCEPT_ID	의약품 투여경로	integer	외래키- 약물 투여경로를 나타내는 Standardized Vocabularies에서 기 정의된 개념. *Concept table에서 domain_ID= 'route' 인코드검색 예) Oral:4128794, Inhaling:4120036
LOT_NUMBER	Lot 숫자	varchar(50)	특정량(particular quantity) 또는 의약품 생산 로트번호와 같이 제조자가 제품에 부여한 식별자
PROVIDER_ID	의료서비스 제공자	integer	외래키-약물노출이 시작된(처방, 투약) 사람에 대한 Provider를 PROVIDER table에서 참조
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRECE 테이블의 VISIT_OCCURRENCE_ID 참조 :약물노출이 시작된 방문에 대한 정보
VISIT_DETAIL_ID	방문 상세 ID	integer	외래키-VISIT_DETAIL 테이블의 VISIT_DETAIL_ID 참조 :약물노출이 시작된 방문에 대한 정보

CDM 컬럼		Type	Description
Eng	Kor		
DRUG_SOURCE_VALUE	처방 의약품 원본값	varchar (50)	원본데이터의 약물 코드 원본데이터에서 사용된 약물코드는 Standardized Vocabularies의 표준약물 컨셉과 매핑되어 있음. 참조용으로 원본데이터의 값이 이 필드에 저장되어 있음.
DRUG_SOURCE_CONCEPT_ID	처방 의약품 소스코드 체계	integer	외래키 원본데이터에서 사용되는 약물 코드체계에 관한 CONEPT_ID
ROUTE_SOURCE_VALUE	투여경로 원본값	varchar (50)	원천 데이터에 기록되어 있는 약물 투여 경로에 대한 상세 내역 정보
DOSE_UNIT_SOURCE_VALUE	용량 단위 원본값	varchar (50)	원천 데이터에 기록되어 있는 용량 단위에 대한 상세 정보

## 6) Condition\_occurrence - 환자의 건강상태 및 진단명을 알 수 있는 테이블

- 일반 진단테이블과 퇴원진단테이블을 모두 포함 (퇴원진단테이블은 환자 퇴원 시 퇴원요약지에 기재된 진단명을 의무기록팀에서 KCD로 분류하는 과정을 통해 검토/수정되는 진단명 기록 테이블을 말하며, 이 경우는 provider ID를 통해 구분하게 됨)

CDM 컬럼		Type	Description
Eng	Kor		
CONDITION_OCCURRENCE_ID	진단 ID	integer	각 진단 발생 이벤트에 관한 고유 식별자
PERSON_ID	환자 ID	integer	외래키 진단 정보를 가지고 있는 환자에 대한 고유 식별자 환자의 자세한 인구학적 정보는 PERSON 테이블에 저장됨
CONDITION_CONCEPT_ID	상병	integer	외래키 표준사전(StandardizedVocabularies)에 의해 사전 정의된 진단 컨셉에 관한 CONCEPT_ID CONCEPT 테이블 참조
CONDITION_START_DATE	진단 개시일	date	진단 사례가 기록된 날짜 (연/월/일)
CONDITION_START_DATETIME	진단 개시일	datetime	진단 사례가 기록된 일시 (연/월/일/시)
CONDITION_END_DATE	진단 종료일	date	진단 사례가 종료되는 것으로 판단되는 날짜 (연/월/일) *입원/응급환자 - 퇴원약 있으면 처방일+처방기간 - 퇴원약 없으면 퇴원일 *외래/통원환자 - 약처방 있으면 처방일+처방기간 - 약처방 없으면 내원일 동일

CDM 컬럼		Type	Description
Eng	Kor		
CONDITION_END_DATETIME	진단 종료 일시	date	진단 사례가 종료되는 것으로 판단되는 일시(연/월/일/시)
CONDITION_TYPE_CONCEPT_ID	주진단/부진단 여부	integer	외래키-표준 용어의 개념 식별자(Concept ID) 참조. 기록된 Condition의 발병 유형과 표준화 수준을 나타냄. *Concept table에서 vocabulary_ID='conditionType'인 코드 검색 예)EHRChiefComplaint:42894222,PrimaryCondition: 44786627
STOP_REASON	중단 사유	varchar (20)	진단이 더 이상 존재하지 않는 이유 - 원천 데이터에 명시된 내용
PROVIDER_ID	의료 서비스 제공자	integer	외래키-PROVIDER table의 Provider ;진단 결정에 책임이 있는 공급자(주치의)
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRENCE 테이블의 VISIT_OCCURRENCE_ID 참조 ;진단 결정 당시
VISIT_DETAIL_ID	방문 상세 ID	integer	외래키-VISIT_DETAIL 테이블의 VISIT_DETAIL_ID 참조 ; 진단 결정 당시
CONDITION_SOURCE_VALUE	상병 원본값	integer	원천 데이터에 있는 condition code. StandardizedVocabularies와 매핑되고, 참조값으로 이 필드에 저장됨.
CONDITION_SOURCE_CONCEPT_ID	상병 소스코드 체계	integer	외래키-원천 데이터의 Condition Concept
CONDITION_STATUS_SOURCE_VALUE	병태 상태 원본값	varchar (50)	원천 데이터의 condition status 값(진단 당시) *인정된 진단:concept_id 4203942를 사용 *사전에 진단:concept_id 4033240을 사용 *최종진단:concept_id 4230359를 - 입원시 진단 ‘입원진단’, - 퇴원시 진단 ‘퇴원진단’
CONDITION_STATUS_SOURCE_CONCEPT_ID	병태 상태 소스코드 체계	varchar (50)	외래키-Standard Vocabulary의 CONCEPT_ID ; 기 정의된 condition status *인정된 진단:concept_id 4203942를 사용 *사전에 진단:concept_id 4033240을 사용 *최종진단:concept_id 4230359를 사용 - 퇴원 진단을 위해서도 사용됨

## 7) Device\_exposure - 기기 및 각종 치료재료 사용 정보를 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
DEVICE_EXPOSURE_ID	기기 노출 ID	integer	각 의료기기 노출에 대한 시스템이 생성하는 고유 식별자
PERSON_ID	환자 ID	integer	외래키-의료기기에 노출된 환자에 대한 고유 식별자 환자의 자세한 인구학적 정보는 PERSON 테이블에 저장됨
DEVICE_CONCEPT_ID	기기 Concept ID	integer	외래키-StandardizedVocabularies의 Standard Concept id 참조 ; 의료기기 개념
DEVICE_EXPOSURE_START_DATE	기기적용 시작일	date	의료기기나 물품을 적용하거나 사용을 시작한 날(연/월/일)
DEVICE_EXPOSURE_START_DATETIME	기기적용 시작일시	datetime	의료기기나 물품을 적용하거나 사용을 시작한 일시(연/월/일/시) *v5.2 added
DEVICE_EXPOSURE_END_DATE	기기적용 종료일	date	의료기기나 물품을 적용하거나 사용을 종료한 날(연/월/일)
DEVICE_EXPOSURE_END_DATETIME	기기적용 종료일시	datetime	의료기기나 물품을 적용하거나 사용을 종료한 일시(연/월/일/시) *v5.2 added
DEVICE_TYPE_CONCEPT_ID	기기 노출 ID	integer	외래키-Standardized Vocabularies의 기 정의된 Concept id 참조 ; 기록된 기기 노출의 유형
UNIQUE_DEVICE_ID	기기 객체 식별자	varchar(50)	환자에게 사용된 기기의 instance를 식별하는 UDI 또는 이와 동등한 식별자
QUANTITY	수량	integer	노출에 사용된 개별 기기의 수
PROVIDER_ID	의료 서비스 제공자	integer	외래키-PROVIDER table의 provider 참조 ; 기기 노출을 시작한 의료 제공자
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRENCE 테이블의 VISIT ; 기기가 사용된 방문
VISIT_DETAIL_ID	방문 상세 ID	integer	외래키-VISIT_DETAIL 테이블의 VISIT ; 기기 노출이 시작된 방문
DEVICE_SOURCE_VALUE	기기 코드 원본값	varchar(50)	원천 데이터에서의 기기의 소스코드값. StandardizedVocabularies와 매핑되고, 참조값으로 이 필드에 저장됨.
DEVICE_SOURCE_CONCEPT_ID	기기 코드 ID	integer	외래키-원천 데이터에서 사용되는 기기 개념 코드

## 8) Measurement - 환자의 검사 결과 값을 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
MEASUREMENT_ID	검진 ID	integer	각 MEASUREMENT에 관한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-MEASUREMENT가 기록된 환자의 고유 식별자 환자의 자세한 인구학적 정보는 PERSON 테이블에 저장됨.
MEASUREMENT_CONCEPT_ID	검진항목	integer	외래키-StandardizedVocabularies의 standard measurement 개념 식별자
MEASUREMENT_DATE	검진일자	date	MEASUREMENT 수행일 (연/월/일)
MEASUREMENT_DATETIME	검진시간	datetime	MEASUREMENT 수행 일시 (연/월/일/시) 일부 데이터베이스는 시간에 대한 datatype이 없고, 시간분석을 위해, MEASUREMENT의 date와 time을 결합하여 사용할 수 있음
MEASUREMENT_TYPE_CONCEPT_ID	검진 데이터소스 유형	integer	외래키- MEASUREMENT레코드가 기록된 원천 데이터를 반영한 표준용어에서 기 정의된 개념 식별자 *Concept table에서 vocabulary_ID= 'MeasType' 인코드검색 예)Pathology finding:44818703, Lab result:44818702
OPERATOR_CONCEPT_ID	측정결과 적용 수학기호	integer	외래키-표준용어에 기정의된 개념 식별자 value_as_number에 적용되는 수학 연산자: <, <=, =, >=, > *관측결과(value)는 operator + number, number, 또는 concept id로 표현될 수 있음
VALUE_AS_NUMBER	숫자 측정값	float	수치형 값(numeric value)으로 표현된 측정값 *비율로 표기되는 검사결과 (예: 항체검사결과 '1:6.1')의 경우, 나누어 계산한 수(0.163934426)를 입력하며 이 때 unit_concept_id는 'titer(8525)'로 입력
VALUE_AS_CONCEPT_ID	측정값의 코드	integer	외래키-Standardized Vocabularies에 개념으로 표현된 측정 결과 (e.g., positive/negative, present/absent, low/high, etc.).
UNIT_CONCEPT_ID	단위	integer	외래키-Standardized Vocabularies에 있는 측정단위에 대한 표준 Concept_ID *Concept table에서 domain_id= 'unit' 인코드검색 예)kg:4122383, U/day:4117658
RANGE_LOW	정상범위의 하한치	float	정상 하한값. Measurement value와 동일한 단위 사용
RANGE_HIGH	정상범위의 상한치	float	정상 상한값. Measurement value와 동일한 단위 사용
PROVIDER_ID	의료서비스 제공자	integer	외래키-ROVIDER table의 provider. 검사를 시작하거나 측정에 대한 책임자

CDM 컬럼		Type	Description
Eng	Kor		
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-MEASUREMENT가 기록된 VISIT_OCCURRENCE 테이블의 VISIT_OCCURRENCE_ID 참조
VISIT_DETAIL_ID	방문 상세 ID	integer	외래키-VISIT_DETAIL 테이블의 VISIT_DETAIL_ID 참조 ;MEASUREMENT가 기록된 방문 v5.3 added
MEASUREMENT_SOURCE_VALUE	검진 항목 원본값	varchar(50)	원본데이터에서 보이는 Measurement name Standardized Vocabularies의 표준컨셉과 매핑되고 참조용으로 소스코드가 이 필드에 저장됨
MEASUREMENT_SOURCE_CONCEPT_ID	검진 항목 소스코드 체계	integer	외래키-원천 데이터에서 사용되는 코드를 참조하는 Standardized Vocabularies의 CONCEPT
VALUE_SOURCE_VALUE	측정값 원본값	varchar(50)	원천 데이터에 저장되어 있는 value_as_number 또는 value_as_concept_id의 콘텐트와 관련있는 source value

## 9) Specimen - 검체 채취일시, 양, 위치, 단위등을 알 수 있는 테이블

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
SPECIMEN_ID	검체 ID	integer	각 검체에 대한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-Specimen이 기록이 있는 사람에 대한 식별자
SPECIMEN_CONCEPT_ID	검체	integer	외래키-StandardVocabularies의 CONCEPT_ID ; 검체
SPECIMEN_TYPE_CONCEPT_ID	검체 타입	integer	외래키-StandardVocabularies의 CONCEPT_ID ; 원천 데이터에 표현된 검체의 기록 체계 반영
SPECIMEN_DATE	검체 채취일	date	환자로부터 검체를 채취한 날짜
SPECIMEN_DATETIME	검체 채취일시	datetime	환자로부터 검체 시료를 채취한 일시 *v5.2 added
QUANTITY	수량	float	샘플링 과정 동안 환자로부터 채취한 검체의 양
UNIT_CONCEPT_ID	단위	integer	외래키-Standard Concept id ; 검체 채취의 수치적 양과 관련된 단위
ANATOMIC_SITE_CONCEPT_ID	해부학적 위치 ID	integer	외래키-Standard Concept id ; 검체 채취의 해부학적 위치
DISEASE_STATUS_CONCEPT_ID	질병 상태 ID	integer	외래키-Standard Concept id ; 검체 채취된 질병 상태
SPECIMEN_SOURCE_ID	검체 시료 ID	varchar(50)	원천 데이터 상 검체 식별자

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
SPECIMEN_SOURCE_VALUE	검체 시료 값	varchar (50)	원천 데이터의 검체 값 표준용어(StandardizedVocabularies)에서 표준 컨셉과 매핑 되어있음 참조용으로 원천 데이터의 코드가 이 필드에 저장되어 있음
UNIT_SOURCE_VALUE	시료 단위	varchar (50)	원천 데이터에 명시된 단위에 대한 정보
ANATOMIC_SITE_SOURCE_VALUE	해부학적 위치	varchar (50)	원본 데이터에 명시된 해부학적 부위 정보
DISEASE_STATUS_SOURCE_VALUE	질병 상태	varchar (50)	원본 데이터에 명시된 질병의 상태 정보

- 10) Observation - 환자의 가족력여부, 흡연여부, 알러지 여부 등을 문진 또는 설문조사를 통해 얻어지는 정보를 알 수 있는 테이블
- 알러지, 과거력, 가족력 및 흡연/음주 등의 정보를 포함하여 구조적으로 기록된 환자에 대한 관찰내용을 최대한 포함할 것을 권장

CDM 컬럼		Type	Description
Eng	Kor		
OBSERVATION_ID	문진 ID	integer	각 observation에 관한 고유 식별자 * unique해야 함
PERSON_ID	환자 ID	integer	외래키-PERSON table의 PERSON_ID ;observation이 기록된 환자 환자에 관한 자세한 인구학적 정보는 PERSON 테이블에 저장됨
OBSERVATION_CONCEPT_ID	문진 항목	integer	외래키-Standardized Vocabularies의 concept_id ;observation concept
OBSERVATION_DATE	문진일자	date	OBSERVATION 기록일(연/월/일)
OBSERVATION_DATETIME	문진일시	datetime	OBSERVATION 기록 일시(연/월/일/시) *v 5.2 added
OBSERVATION_TYPE_CONCEPT_ID	문진 데이터소 스 유형	integer	외래키-Standardized Vocabularies의 concept_id ;OBSERVATION type *Concept table에서 vocabulary_ID= ‘observationType’ 인코드검색 예)Patientreported:44814721,Problemlist fromEHR:38000276
VALUE_AS_NUMBER	숫자 측정값	float	숫자로 저장된 OBSERVATION 결과값. 결과가 숫자값으로 표현된 OBSERVATION에 적용할 수 있음.
VALUE_AS_STRING	문자 측정값	varchar (60)	문자로 저장된 OBSERVATION 결과값. 축약된 문자(verbatim text)로 표현된 observation에 적용할 수 있음

CDM 컬럼		Type	Description
Eng	Kor		
VALUE_AS_CONCEPT_ID	측정값의 코드 체계	Integer	외래키-CONCEPT_ID로 저장된 OBSERVATION 결과 결과가 표준용어(Standardized Vocabularies)의 표준 컨셉으로 표현될 수 있는 observations에 적용할 수 있음. (예: positive/negative, present/absent, low/high, etc.).
QUALIFIER_CONCEPT_ID	기준	integer	외래키-Standard Concept ID 수식어(qualifier)(e.g., severity of drug-drug interaction alert)
UNIT_CONCEPT_ID	단위	integer	외래키-Standard Concept ID 표준용어(Standardized Vocabularies)의 OBSERVATION 단위 *Concepttable에서 domain_id= 'unit' 인 코드검색 예)kg:4122383,U/day:4117658
PROVIDER_ID	의료서비 스 제공자	integer	외래키-PROVIDER table의 provider OBSERVATION에 책임이 있는 PROVIDER
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRENCE 테이블의 visit OBSERVATION이 기록된 방문 정보
VISIT_DETAIL_ID	방문상세 ID	integer	외래키-VISIT_DETAIL 테이블의 visit OBSERVATION이 기록된 방문 정보 *v5.3 added
OBSERVATION_SOURCE_VALUE	문진 항목 원본값	integer	원천 데이터에서 보이는 OBSERVATION 코드값 이 코드는 표준용어(Standardized Vocabularies)의 표준 컨셉에 매핑되고 소스코드는 참조용으로 여기에 저장됨.
OBSERVATION_SOURCE_CONCEPT_ID	문진 항목 소스코드 체계	varchar (50)	외래키-원본데이터에서 사용된 코드를 참조하는 CONCEPT
UNITS_SOURCE_VALUE	단위 원본값	integer	원천 데이터에서 보이는 단위에 대한 소스코드. 표준용어(Standardized Vocabularies)의 unit 컨셉에 매핑되고 소스코드는 참조용으로 여기에 저장됨.
QUALIFIER_SOURCE_VALUE	기준 원본값	varchar (50)	observation을 특징화하기 위한 수식어와 관련된 소스 값

## 11) Death - 환자의 사망기록에 대해 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
PERSON_ID	환자 ID	integer	외래키-person table의 person_ID ; 사망한 환자 환자의 자세한 인구학적 정보는 PERSON 테이블에 저장됨 *unique해야 함
DEATH_DATE	사망일자	date	환자 사망일(년/월/일) 정확한 사망일 또는 사망 월이 없는 경우 12월 또는 그 달의 마지막 날을 default로 함
DEATH_DATE_TIME	사망일시	datetime	환자 사망일시(년/월/일/시) 정확한 사망일 또는 사망 월이 없는 경우 12월 또는 그 달의 마지막 날을 default로 함. *v5.2 added
DEATH_TYPE_CONCEPT_ID	사망 진단 유형	integer	외래키-Standardized Vocabularies의 concept id 원천 데이터에 표현된 사망 유형에 대한 표준용어에서 사전 정의된 concept id *Concept table에서 vocabulary_ID= 'DeathType' 인코드검색 예)Death Certificate underlying cause:38003618
CAUSE_CONCEPT_ID	사망원인	integer	외래키-Standardized Vocabularies의 concept id conditions에 대한 표준 용어에서의 표준 concept id *사망원인이 자유 기술문으로 작성된 경우, cause_source_value만 작성
CAUSE_SOURCE_VALUE	사망원인 원본값	varchar (50)	원천 데이터의 사망원인에 관한 코드 값 이 코드는 표준 용어(Standardized Vocabularies)의 표준 컨셉과 매핑 되고, 원천 데이터의 값이 참조용으로 이 필드에 저장됨.
CAUSE_SOURCE_CONCEPT_ID	사망원인 소스코드 체계	integer	외래키-원본데이터에서 사용된 코드를 탐조하는 컨셉 주의: 이 변수명은 여러 데이터베이스 플랫폼에서 작동 되도록 축약되었음

## 12) Note 테이블 - 자유 기술문으로 구성된 각종 기록지의 정보를 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
NOTE_ID	노트 ID	integer	각 노트에 관한 고유 식별자
PERSON_ID	환자 ID	integer	외래키-PERSON table의 PERSON_ID ; 노트 기록이 있는 환자 환자의 자세한 인구학적 정보는 PERSON 테이블 에 저장됨
NOTE_DATE	기록일자	date	노트 작성일 (연/월/일)
NOTE_DATETIME	기록일시	datetime	노트 작성 일시 (연/월/일/시) *v 5.2 added
NOTE_TYPE_CONCEPT_ID	기록타입 or 출처	integer	외래키-Standardized Vocabularies의 CONCEPT_ID 노트의 출처를 반영한 표준 용어에서 사전 정의 된 개념
NOTE_CLASS_CONCEPT_ID	어휘분류	integer	외래키-Standardized Vocabularies의 Concept_ID 노트의 HL7 LOINC Document Type 용어 분류를 반영한 표준용어(Standardized Vocabularies)에서 사전에 정의된 컨셉 *v 5.2 added
NOTE_TITLE	노트 제목	string (250)	원천 데이터에 보이는 노트의 제목 *v 5.2 added
NOTE_TEXT	노트 내용	varchar (MAX)	노트 내용 *v 5.3 changed RDBMSdependenttext → varchar(MAX)
ENCODING_CONCEPT_ID	인코딩 유형	integer	외래키-Standardized Vocabularies의 Concept_ID 인코딩 유형의 노트 문자를 반영하는 표준용어(Standardized Vocabularies)에 사전 정의된 컨셉 *v 5.2 added
LANGUAGE_CONCEPT_ID	언어 ID	integer	외래키-Standardized Vocabularies의 Concept_ID 노트에 사용된 언어를 반영하는 표준용어 (Standardized Vocabularies)에 사전에 정의된 컨 셉 *v 5.2 added
PROVIDER_ID	의료서비 스제공자	integer	외래키-PROVIDER table의 provider ; 노트 작성자
VISIT_OCCURRENCE_ID	방문 ID	integer	외래키-VISIT_OCCURRENCE table의 visit NOTE가 기록된 방문
VISIT_DETAIL_ID	방문상세 ID	integer	외래키-VISIT_DETAIL table의 visit NOTE가 기록된 방문 *v5.3 added
NOTE_SOURCE_VALUE	원본값	varchar (50)	원천 데이터에서 노트의 출처와 관련된 소스코드 값

13) Location - 환자나 병원의 위치 정보를 알 수 있는 테이블

- 지원기관에서 제공하는 주소 테이블 이용가능

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
LOCATION_ID	위치 ID	integer	각 지리정보에 관한 고유 식별자 *unique해야 함
ADDRESS_1	주소 1	varchar(50)	원천 데이터의 주소필드1, 일반적으로 거리 주소(street address)가 사용됨
ADDRESS_2	주소 2	varchar(50)	원천 데이터의 주소필드2 일반적으로 빌딩,층수등과 같은 추가적인 상세주소가 사용됨
CITY	도시	varchar(50)	원천 데이터의 도시 필드
STATE	주	varchar(2)	원천 데이터의 주 필드
ZIP	우편번호	varchar(9)	우편번호
COUNTRY	국가	varchar(20)	나라
LOCATION_SOURCE_VALUE	위치 원본값	varchar(100)	원천 데이터의 지역 식별에 고유하게 사용되는 축약 정보

14) Care\_site - 기관(병원)의 정보를 알 수 있는 테이블

- 참여기관이 실제 물리적으로 분리된 여러 의료기관의 데이터를 포함하지 않는 경우 실제 1행의 데이터만 포함됨

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
CARE_SITE_ID	요양기관 ID	integer	각 의료 시설에 관한 고유 식별자 *unique해야 함
CARE_SITE_NAME	요양기관 명	varchar (255)	원천 데이터에서 의료시설 이름 또는 축약 설명(verbatim description)
PLACE_OF_ SERVICE_ CONCEPT_ID	요양기관 유형 (예: 대학병원)	integer	외래키-Standardized vocabularies의 CONCEPT_ID ; 서비스를 받은 장소
LOCATION_ID	위치	integer	외래키-LOCATION table(상세 주소 정보)의 LOCATION_ID ; 지리적 위치 정보
CARE_SITE_SOURCE_VALUE	요양기관 원본값	varchar(50)	원천 데이터에 있는 Care Site 식별자로 참조용으로 저장됨
PLACE_OF_ SERVICE_SOURCE _VALUE	요양기관 유형 원본값	varchar(50)	원천 데이터에 있는 place of service에 대한 소스코드. 참조용으로 여기 저장됨.

## 15) Provider - 의료 서비스 제공자 식별 목록 포함

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
PROVIDER_ID	의료서비스 제공자	Integer	각 의료서비스 제공자에 대한 고유 식별자
PROVIDER_NAME	의료서비스 제공자명	Varcahr (255)	의료서비스 제공자명 *v5.3changed Varcahr(50)→Varchar(255)
NPI	NPI 정보	Varchar (20)	의료서비스 제공자 NPI(National Provider Identifier)
DEA	DEA 정보	Varchar (20)	의료서비스 제공자 DEA(Drug Enforcement Administration) 번호
SPECIALTY_CONCEPT_ID	전공 ID	Integer	외래키-StandardizedVocabularies의 Concept ID ; 전공과목(Specialty)
CARE_SITE_ID	요양기관 ID	Integer	외래키-CARE_SITE 테이블의 CARE_SITE_ID ; 의료서비스 제공자가 근무하는 주 의료시설
YEAR_OF_BIRTH	출생연도	Integer	의료서비스 제공자의 출생연도
GENDER_CONCEPT_ID	성별 ID	Integer	의료서비스 제공자의 성별
PROVIDER_SOURCE_VALUE	제공자 식별자	Varchar (50)	원천 데이터의 의료서비스 제공자 식별자. 참조를 위해 이 필드에 저장됨.
SPECIALTY_SOURCE_VALUE	전공	Varchar (50)	원천 데이터에서의 의료서비스 제공자 전공에 대한 소스코드. 참조를 위해 이 필드에 저장됨.
SPECIALTY_SOURCE_CONCEPT_ID	전공 ID	Integer	외래키-CONCEPT테이블의 CONCEPT_ID ; 원천 데이터 코드 참조
GENDER_SOURCE_VALUE	성별	Varcahr (50)	원천 데이터에서의 의료서비스 제공자의 성별 코드값. 참조를 위해 이 필드에 저장됨.
GENDER_SOURCE_CONCEPT_ID	성별 ID	Integer	외래키-CONCEPT테이블의 CONCEPT_ID ; 원천 데이터의 성별 코드

## 16) Cost

- 환자가 병원에 내원하여 진료를 받고 지불한 계산 내역을 알 수 있는 테이블

CDM 컬럼		Type	Description
Eng	Kor		
COST_ID	식별번호	integer	각 COST 레코드에 대한 고유 식별자
COST_EVENT_ID	이벤트 ID	integer	외래키-COST 데이터가 기록되는 이벤트에 대한 고유 식별자(예시:MEASUREMENT,PROCEDURE,VISIT,D RUGEXPOSURE)
COST_DOMAIN_ID	도메인 ID	varchar(20)	COST 이벤트의 도메인을 나타내는 개념. COST 정보가 기록된 Entity의 도메인을 알 수 있는 Entity별 도메인 정보 테이블에서 가져옴. *v5.3changed string(20) → varchar(20)
COST_TYPE_CONCEPT_ID	구분 ID	integer	외래키-CONCEPT table의 CONCEPT_ID ; COST 데이터의 원본 또는 출처. 보험 청구 정보로부터 계산된 provider revenue(회계 DB에서 보고된 비용대비 요금 비[cost-to-charge ratio]로 산출)
CURRENCY_CONCEPT_ID	국제통화 ID	integer	외래키-CONCEPT 테이블의 CONCEPT_ID ; 미국 달러를 USD로 나타내는 것과 같이 국제 통화단위를 기술하는 3글자 CONCEPT_ID
TOTAL_CHARGE	총금액	float	의료서비스 제공자(PROVIDER)가 상품이나 서비스에 대해 Payer에게 부과(청구)한 총액 -Payer: 보험회사, 환자 -Provider: 병원, 의사, 약국, DME 제공자
TOTAL_COST	비용	float	상품이나 서비스에 대해 의료서비스 제공자가 발생시킨 비용
TOTAL_PAID	총지불금	float	의료서비스 제공자의 상품이나 서비스에 대해 모든 Payer로부터 실제로 지불된 총 금액
PAID_BY_PAYER	지불금	float	상품이나 서비스에 대해 해당 Payer가 지불한 금액
PAID_BY_PATIENT	환자부담금	float	진료비 중 환자가 부담한 총액
PAID_PATIENT_COOPAY	정액환자 부담금	float	진료비 중 정액 환자 부담금
PAID_PATIENT_COINSURANCE	정액환자 공동보험	float	위험에 대하여 공동 감수(joint assumption)하는 부분으로 환자가 지불하는 금액. 일반적으로 환자 본인부담액이 일정수준을 초과한 양에 대하여 PAYER PLAN에 명시된 비용의 비율
PAID_PATIENT_DEDUCTIBLE	정액환자 본인 부담금	float	PAYER PLAN에 의해 규정된 환자 본인부담액(deductible). PAID_PATIENT_DEDUCTIBLE은 PAID_BY_PATIENT 변수에 영향을 미침.
PAID_BY_PRIMARY	주지불자 지불액	float	조정(coordination of benefits)을 통해 주지불자(primary Payer)에 의해 지불된 금액

CDM 컬럼		Type	Description
Eng	Kor		
PAID_INGREDIENT_COST	약값 지불액	float	PAYER가 약국에 지불하는 약값(조제료 제외). 이 필드가 ‘0’ 이 아니면 paid_ingredient_cost는 paid_by_payer 필드에 영향을 미침.
PAID_DISPENSING_FEE	조제료 지불액	float	PAYER가 약국에 지불하는 조제료(약값 제외). 이 필드가 ‘0’ 이 아니면 paid_dispensing_fee는 paid_by_payer에 영향을 미침.
PAYER_PLAN_PERIOD_ID	보험자 약정정보 ID	integer	외래키-PAYER_PLAN_PERIOD table. PAYER_PLAN_PERIOD: 보험회사, 약정(Plan) 및 가족에 대한 상세 내역이 저장된 테이블. paid_by_payer 필드에 기여한 보험회사와 관련된 payer_plan_id를 기록함.
AMOUNT_ALLOWED	계약금액	float	PAYER과 의료 서비스 제공자 간의 합의된 계약금액
REVENUE_CODE_CONCEPT_ID	수익코드ID	integer	외래키-Standardized Vocabularies의 CONCEPT_ID; 수익코드(Revenue codes)
REVENUE_CODE_SOURCE_VALUE	수익코드 원본값	varchar(50)	원천 데이터의 수익코드 -참조를 위해 이 필드에 저장됨
DRG_CONCEPT_ID	DRG ID	integer	외래키-DRG Vocabulary의 CONCEPT_ID ; visit에 대한 DRG
DRG_SOURCE_VALUE	DRG코드	varchar(3)	원천 데이터의 3자리 DRG 코드

### 17) Payer\_plan\_period - 환자의 보험 정보를 알 수 있는 테이블

- 자동차보험, 산재 등의 일시적인 보험자격을 포함
- 건강보험은 출생부터 쭉 자격이 이어진 것으로 볼 수 있음

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
PAYER_PLAN_PERIOD_ID	보험가입 기간 ID	integer	PAYER, PLAN, 가족 코드와 시간 범위( time span)에 대한 고유 조합 식별자 *unique해야 함
PERSON_ID	환자 ID	integer	외래키-PERSON 테이블의 PERSON_ID ; PAYER 의해 보장되는 환자 식별자. 환자의 자세한 인구학적 정보는 PERSON 테이블에 저장됨
PAYER_PLAN_PERIOD_START_DATE	보험 개시일자	date	PAYER_PLAN 기간의 시작일
PAYER_PLAN_PERIOD_END_DATE	보험 마감일자	date	PAYER_PLAN 기간의 종료일
PAYER_CONCEPT_ID	보험가입 ID	integer	외래키-StandardizedVocabularies의 CONCEPT_ID ; standard Payer concept identifier *v5.3 added

CDM 컬럼		Type	Description (Kor)
Eng	Kor		
PAYER_SOURCE_VALUE	보험가입 원본값	varchar (50)	원천 데이터의 PAYER 코드
PAYER_SOURCE_CONCEPT_ID	보험가입 원본ID	integer	외래키-원천 데이터의 PAYER CONCEPT_ID *v5.3 added
PLAN_CONCEPT_ID	계획 개념 ID	integer	외래키-StandardizedVocabularies의 CONCEPT_ID ;건강보험 보장 약관(health benefit plan) *v5.3 added
PLAN_SOURCE_VALUE	가입 보험 원본값	varchar (50)	원천 데이터에서의 환자의 건강보험 보장 약관 (health benefit plan) 코드
PLAN_SOURCE_CONCEPT_ID	가입 보험 원본 ID	integer	외래키-원천 데이터에서의 PLAN 개념 코드 *v5.3 added
SPONSOR_CONCEPT_ID	스폰서 ID	integer	외래키-StandardizedVocabularies의 CONCEPT_ID ;SPONSER *v5.3 added
SPONSOR_SOURCE_VALUE	스폰서 원본값	varchar (50)	원천 데이터의 건강보험 약관(Health Plan)에 대하여 환자의 SPONSER 코드 *v5.3 added
SPONSOR_SOURCE_CONCEPT_ID	스폰서 원본 ID	integer	외래키-원천 데이터의 SPONSOR CONCEPT_ID *v5.3 added
FAMILY_SOURCE_VALUE	가족 원본값	varchar (50)	원천 데이터의 환자의 가족에 관한 코드값
STOP_REASON_CONCEPT_ID	중단사유 ID	integer	외래키-StandardizedVocabularies의 CONCEPT_ID ;보험 보장이 중단된 이유 *v5.3 added
STOP_REASON_SOURCE_VALUE	중단사유	varchar (50)	원천 데이터의 보험 보장 중단 사유 *v5.3 added
STOP_REASON_SOURCE_CONCEPT_ID	중단사유 ID	integer	외래키- 원천 데이터의 보험보장 중단 개념 코드 *v5.3 added

## 6.4. 표준화된 2차 자료(Standardized Derived Elements)

- 다음 테이블은 원천 데이터에서 직접 획득하지 않고 변환된 CDM 테이블로 부터 2차적으로 가져온 환자의 임상 이벤트에 대한 정보가 포함되어 있음.
  - DRUG\_ERA
  - DOSE\_ERA
  - CONDITION\_ERA
- OMOP CDM의 표준화된 2차 자료 내의 테이블 설명 ERD(entity-relationship diagram)



## 1) DRUG\_ERA - Drug\_exposure 테이블로부터 2차 생성

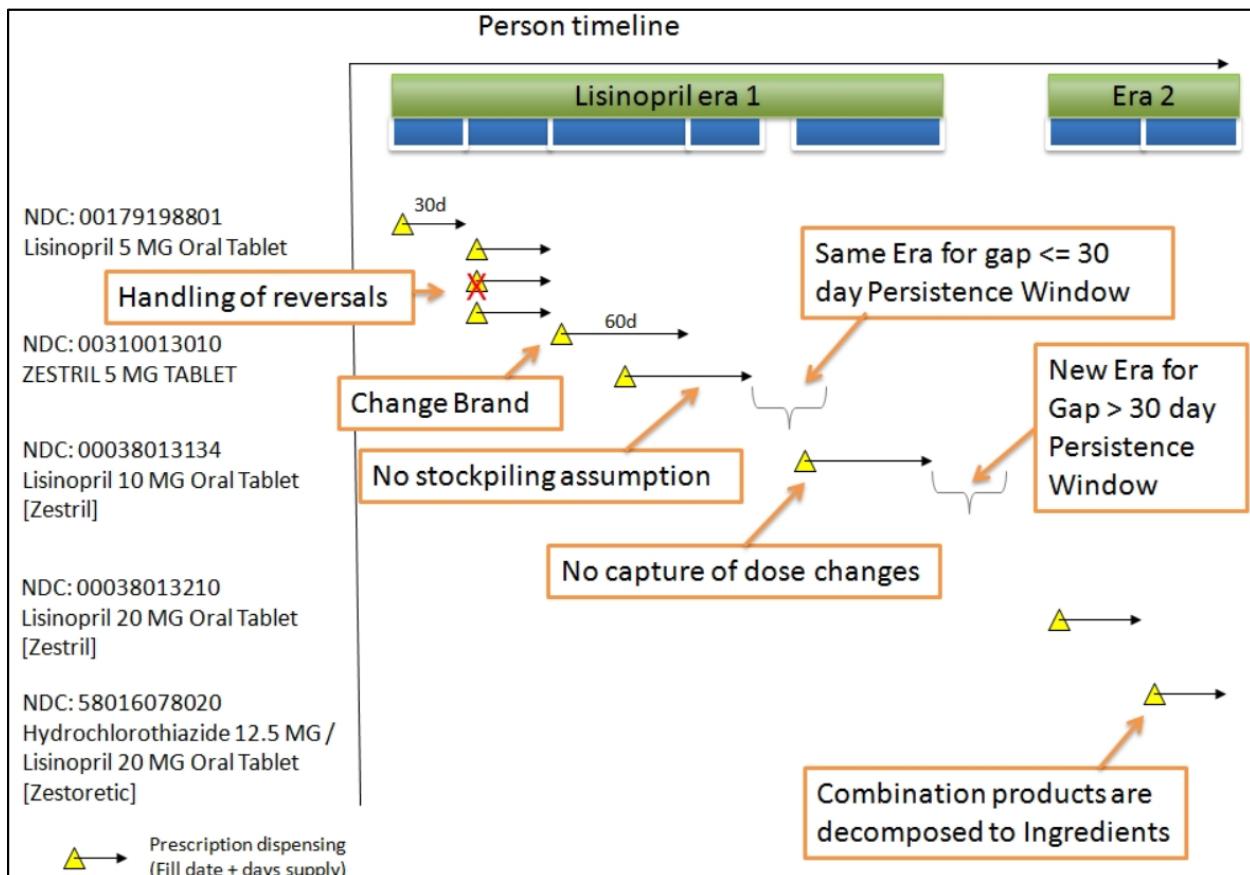
- Drug Era는 사람이 특정 활성 성분(active ingredient)에 노출된다고 가정되는 시간의 범위로 정의됨. Drug Era는 Drug Exposure와 같지 않음: Exposures는 약물이 사람에게 전달된 시점의 소스에 해당하는 개별 기록이며, Drug Exposure의 연속된 기간을 특정 규칙에 의해 결합하여 continuous Drug Eras를 생성함.

CDM 컬럼		Type	Description
Eng	Kor		
DRUG_ERA_ID	의약품 노출기간 ID	integer	각 DRUG ERA(약물 노출기간)에 관한 고유 식별자 *unique해야 함
PERSON_ID	환자 ID	integer	외래키 DRUG ERA 동안 DRUG에 노출된 사람에 대한 고유 식별자(foreign key identifier). PERSON_ID의 인구학적 상세내용은 PERSON테이블에 저장됨.
DRUG_CONCEPT_ID	의약품	integer	외래키-Standard Vocabulary의 Concept_ID ; 성분 개념(Ingredient Concept) *제품명 단위가 아닌 성분명 단위로 연속처방기간을 계산
DRUG_ERA_START_DATE	노출 개시일자	date	개별 약물 노출 사례로부터 생성된 Drug Era의 시작일. 시간 순으로 기록된 약물 사용이 발생한 가장 첫 날임.
DRUG_ERA_END_DATE	노출 마감일자	date	개별 약물 노출로부터 구성된 DRUG ERA의 종료일. 시간 순으로 기록된 약물 사용 사례의 마지막 날짜
DRUG_EXPOSURE_COUNT	기간 내 노출 횟수	integer	DRUG ERA를 구성하는데 사용된 각각의 약물 노출 발생 횟수
GAP_DAYS	경과일수	integer	ERA 기록을 작성하는데 사용되었으나 DRUG_EXPOSURE 레코드가 적용되지 않은 날수

## 규약(Conventions)

No.	Convention Description
1	표준화된 알고리즘을 사용하여 DRUG_EXPOSURE의 레코드로부터 2차적으로 생성
2	각 Drug Era는 연속된 간격을 형성하고 동일 약물 성분(활성 화합물)을 가지고 있는 하나 이상이 약물 노출에 해당함.
3	drug_concept_id 필드는 ‘Ingredient’라는 concept_class를 가진 개념만 포함하고 있음. Ingredient는 Drug Era 레코드로 집계된 DRUG_EXPOSURE 테이블의 약물 개념에서 도출됨.
4	Drug Era 시작일은 Drug Exposure 첫날임.

No.	Convention Description
	Drug Era End Date는 마지막 Drug Exposure의 종료일임. 각 약물 노출의 종료 날짜는 drug_exposure_end_date 필드에서 가져오거나, 일반적으로 사용할 수 없으므로 다음 규칙을 사용하여 추론함.
5	약제 처방 데이터의 경우, 약물이 처방된 날에 투약일수를 더하여 약물 노출 종료일을 추정함. 국가별 의료 시스템에 따라 이 투약일수 정보는 day_production 필드에 명시적으로 제공되거나 패키지 크기 등의 정보에서 추론됨. 의료행위에 사용하는 약물의 경우, 일반적으로 하루(즉, 투여일)에 투여됨. 두 개의 DRUG_EXPOSURE 레코드 사이에 30일 이내의 standard Persistent Window(공백)는 단일 Drug Era로 통합됨.
6	투약 공백일은 DUG_ERA 기록에 기여하는 모든 Drug Exposure 이벤트 간에 관찰되는 총 비투약 일수를 말함. 새로운 약물 처방이나 리필이 관찰되면(새로운 DRUG_EXPOSURE 기록) 이전 이벤트에서 남은 양은 무시됨.
7	Persistence Window와 Gap Days의 차이는 전자는 두 개의 연속된 DRUG_EXPOSURE 레코드 사이에 허용되는 최대 약물 비사용 시간이고 후자는 위의 비재고 가정(assumption of non-stockpiling) 하에 주어진 Drug Era의 실제 무투약일의 합임.
8	30의 standard Persistence Window와 비재고(non-stockpiling) 가정 선택은 모순이 있지만 약물 결과 추정에서는 좋은 결과를 도출하는 것으로 나타남. 복약 순응 추정과 같은 문제에는 다른 또는 약물 의존적 Persistence Window/stockpiling 가정이 필요함. 연구자들은 코호트처럼 다른 매개변수로 그들 자신의 Drug Eras를 만들어 COHORT 테이블에 보관할 것을 고려하라고 권고함.



2) Dose\_era - 환자가 특정 약성분의 일정 투여량에 노출된 것으로 가정되는 시간의 범위를 정의한 레코드

CDM 컬럼		Type	Description
Eng	Kor		
DOSE_ERA_ID	투여량 기간 ID	integer	각 DOSE ERA(복용량 기간)에 대한 고유 식별자
PERSON_ID	환자 ID	integer	외래키 DRUG ERA 동안 DRUG에 노출된 사람에 대한 고유 식별자(foreign key identifier). PERSON_ID의 인구학적 상세내용은 PERSON테이블에 저장됨.
DRUG_CONCEPT_ID	약물 개념 ID	integer	외래키 활성 성분 개념(Active Ingredient Concept)에 대한 표준 용어의 표준 개념 식별자를 참조하는 외래 키
UNIT_CONCEPT_ID	단위 개념 ID	integer	외래키 단위 개념의 표준 용어에 대한 표준개념 식별자에 대한 외래키
DOSE_VALUE	용량값	float	투여량에 대한 수치 값
DOSE_ERA_START_DATE	투여 시작일	date	개별 약물 노출 사례로부터 생성된 Drug Era의 시작일. 시간 순으로 기록된 약물 사용이 발생한 가장 첫 날임.
DOSE_ERA_END_DATE	투여 종료일	date	개별 약물 노출로부터 구성된 DRUG ERA의 종료일. 시간순으로 기록된 약물 사용 사례의 마지막 날짜

## Conventions

No.	Convention Description
1	Dose Eras는 DRUG_EXPOSURE 테이블의 기록과 DRUG_STRENGTH 테이블의 용량 정보에서 표준 알고리즘을 사용하여 2차적으로 생성됨.
2	각 Dose Era는 연속된 간격을 형성하고 동일한 유효 일일 용량에서 동일한 약물 성분(활성 화합물)을 포함하는 하나 이상의 약물 노출에 해당함.
3	Dose Form 정보는 고려되지 않음. 따라서 환자가 다른 제제(formulations) 또는 동일한 제제가 다른 제조사의 제품으로 변경되더라도 Dose Era의 성분 노출 기간은 지속됨.
4	일일 용량은 각 DRUG_EXPOSURE 기록에 대해 기록된 총용량을 기간으로 나누어 계산함.

- DRUG\_EXPOSURE 기록의 총용량은 다음과 같이 각 약물에 대한 용량 정보가 포함된 DRUG\_STRENGTH 테이블을 참조하여 계산됨.

5	<b>Tablets and other fixed amount formulations</b>
	Example: Acetaminophen (Paracetamol) 500 mg, 20 tablets.
DRUG_STRENGTH	The denominator_unit is empty
DRUG_EXPOSURE	The quantity refers to the number of pieces, e.g. tablets
	In the example: 20
Ingredient dose=	quantity x amount_value [amount_unit_concept_id] Acetaminophen dose = 20 x 500mg = 10,000mg

6	<b>Puffs of an inhaler</b>
	참고: DRUG_STRENGTH 테이블은 분모 단위{actuat}에 배치할 수 있다는 점 이외에 사례 1을 사용하는 차이는 없음. 이 경우 강도는 문자에 제공됨.
DRUG_STRENGTH	The denominator_unit is {actuat}
DRUG_EXPOSURE	The quantity refers to the number of pieces, e.g. puffs
Ingredient dose=	quantity x numerator_value [numerator_unit_concept_id]

7	<b>Quantified Drugs which are formulated as a concentration</b>
	예: 임상 약물은 5 mL 경구용 혼탁액에 들어있는 아세트아미노펜 250 mg/mL임. 정량화된 임상 약물은 DRUG_STRENGTH 표에 1250mg/5ml가 있을 것임. 혼탁액 두 개가 지급됨.
DRUG_STRENGTH	denominator_unit은 mg 또는 mL임. denominator_value가 1이 아닐 수 있음.
DRUG_EXPOSURE	quantity는 팩의 일부(fraction) 또는 배수를 의미함.
	Example: 2
Ingredient dose=	quantity x numerator_value [numerator_unit_concept_id] Acetaminophen dose = 2 x 1250mg = 2500mg

8	<b>제공된 총량을 수량으로 가진 의약품, e.g. chemotherapeutics</b>
	예: 42799258 'Benzyl Alcohol 0.1 ML/ML/Pramoxine hydrochloride 0.01 MG/MG Topical Gel '을 1.25oz 팩에 나누어 제공.
DRUG_STRENGTH	denominator_unit은 mg 또는 mL임.
DRUG_EXPOSURE	Example: Benzyl Alcohol in mL and Pramoxine hydrochloride in mg 수량(quantity)은 mL 또는 g.를 말함.
Ingredient dose=	Example: 1.25 x 30 (conversion factor oz → mL) = 37 quantity x numerator_value [numerator_unit_concept_id] Benzyl Alcohol dose = 37 x 0.1mL = 3.7mL Pramoxine hydrochloride dose = 37 x 0.01mg x 1000 = 370mg
	Note: analytical side는 DRUG_STRENGTH 테이블의 분모를 체크해봐야 함. 두 번째 성분에 mg가 사용되면 계수 1000이 g와 mg 사이 변환에 적용됨.

9	<b>복합약물(Compounded drugs)</b>
	Example : Ibuprofen 20 % / Piroxicam 1 % 크림, 5ml 튜브에 30ml.

9	복합약물(Compounded drugs)
DRUG_STRENGTH	우리는 Ibuprofen과 Piroxicam의 성분에 대한 항목(entries)이 필요하며, 아마도 amount_value가 1이고 mg 단위 일 것임.
DRUG_EXPOSURE	quantity는 복합약물의 총량을 의미함. 각 화합물에 대해 DRUG_EXPOSURE 테이블에서 하나의 레코드를 사용할 것.
	Example: 20% Ibuprofen of 30ml = 6mL, 1% Piroxicam of 30ml = 0.3mL
Ingredient dose=	관련 약물에 따라 다름 : 위의 use cases 중 하나 Ibuprofen dose = $6 \times 1\text{mg} \times 1000 = 6000\text{mg}$ Piroxicam dose = $0.3 \times 1\text{mg} \times 1000 = 300\text{mg}$
	참고: 분석 측면에서는 DRUG_STRENGTH 테이블에 있는 두 성분의 분모가 mg이고 mL / g와 mg 사이의 변환을 인자 factor 1,000을 적용하여 결정함.

10	시간 경과에 따라 유효 성분이 방출되는 약물, e.g. patches
	Example: Ethinyl Estradiol 0.000833 MG/HR / norelgestromin 0.00625 MG/HR Weekly Transdermal Patch
DRUG_STRENGTH	분모 단위는 시간을 나타냄.
	Example: Ethinyl Estradiol 0.000833 mg/h / norelgestromin 0.00625 mg/h
DRUG_EXPOSURE	수량은 개수(no. of pieces) 수를 나타냄.
	Example: 1 patch
Ingredient rate=	numerator_value [numerator_unit_concept_id]
	Ethinyl Estradiol rate = 0.000833 mg/h
	norelgestromin rate 0.00625 mg/h
	Note: 이것은 24를 곱하여 일일 투여량으로 변환 할 수 있음. (최소 24 시간 동안 한 번에 1 패치 사용 가정)

### 3) Condition\_era - Condition\_occurrence 테이블로부터 2차 생성

- Condition Era는 사람이 주어진 condition을 가지고 있다고 가정되는 시간의 범위로 정의됨. Drug Eras와 마찬가지로 Condition Eras는 Condition 발생의 연대기(chronological periods)임. 개별 Condition Occurrences를 단일 Condition Era로 결합하는 것은 두 가지 목적이 있음.
  - 각 Condition을 독립적인 이벤트로 취급하는 대신 빈번하게 지속적인 관리가 필요한 chronic condition을 통합할 수 있음.
  - Condition Occurrence를 이중으로 계산하지 않도록, 동일한 Condition으로 가까운 기간 내에 다회의 의사 방문에 대하여 통합할 수 있음.
- 예를 들어, PCP(Primary Care Physician)를 방문하고 전문의에게 의뢰되는 사

람을 생각해 보면, 나중에 PCP의 원래 진단을 확인하고 적절한 치료법을 제공할 전문의를 방문함. 이 두 개의 독립적인 의사 방문은 하나의 Condition Era.로 통합되어야 함.

CDM 컬럼		Type	Description
Eng	Kor		
CONDITION_ERA_ID	진단기간 ID	integer	각 CONDITION ERA(질병 혹은 증상의 기간)에 대한 고유 식별자
PERSON_ID	환자 ID	integer	외래키 CONDITION ERA 동안 CONDITION(질환/증상)을 경험한 환자에 대한 고유 식별자. PERSON_ID의 인구학적 상세내용은 PERSON테이블에 저장됨.
CONDITION_CONCEPT_ID	진단유형	integer	외래키 표준 용어의 표준 Condition 개념 ID를 참조하는 외래 키.
CONDITION_ERA_START_DATE	진단 개시일자	date	Condition Occurrences의 개별적인 사례로부터 구성된 CONDITION ERA의 시작일. 시간순으로 기록된 condition 중 처음 시작일임.
CONDITION_ERA_END_DATE	진단 마감일자	date	Condition Occurrences의 개별적인 사례로부터 구성된 CONDITION ERA의 종료일. 연속 기록된 Condition 중 마지막 레코드의 종료일임. *condition_occurrence의 end_date가 NULL인 경우에는 연속기간에 포함되는 마지막 condition occurrence 데이터의 condition_start_date로 입력
CONDITION_OCCURRENCE_COUNT	기간 내 진단횟수	integer	CONDITION ERA를 구성하는데 사용된 개별 Condition Occurrences의 횟수

## 규약(Conventions)

No.	규약 내용(Convention Description)
1	Condition Era 레코드는 표준 알고리즘을 사용하여 CONDITION_OCCURRENCE 테이블의 레코드에서 2차적으로 생성됨.
2	각 Condition Era는 연속 간격을 형성하는 하나 또는 여러 개의 Condition Occurrence 레코드에 상응하는 자료임. condition_concept_id 필드에는 Condition Era를 구성하는 CONDITION_OCCURRENCE 테이블 레코드의 개념과 동일한 개념이 들어 있음. Drug Eras와 달리 Condition Eras는 서로 다른 계층 layer의 Condition을 포함하여 집계되지 않음.
	Condition Era 시작일은 첫 번째 Condition Occurrence의 시작일임.
	Condition Era 종료일은 마지막 Condition Occurrence의 종료일임.
3	Condition Eras는 30 일의 Persistence Window로 작성됨. 즉, 한 번 발생 한 후 30 일 이내에 동일한 condition_concept_id가 발생하지 않으면 condition_era_end_date로 간주됨.

## 7.1. 단일 기관 연구

- 단일 기관 연구의 경우 해당기관 IRB 승인 후 해당 기관 데이터를 활용하여 연구를 진행할 수 있음.



## 7.2. 다기관 연구

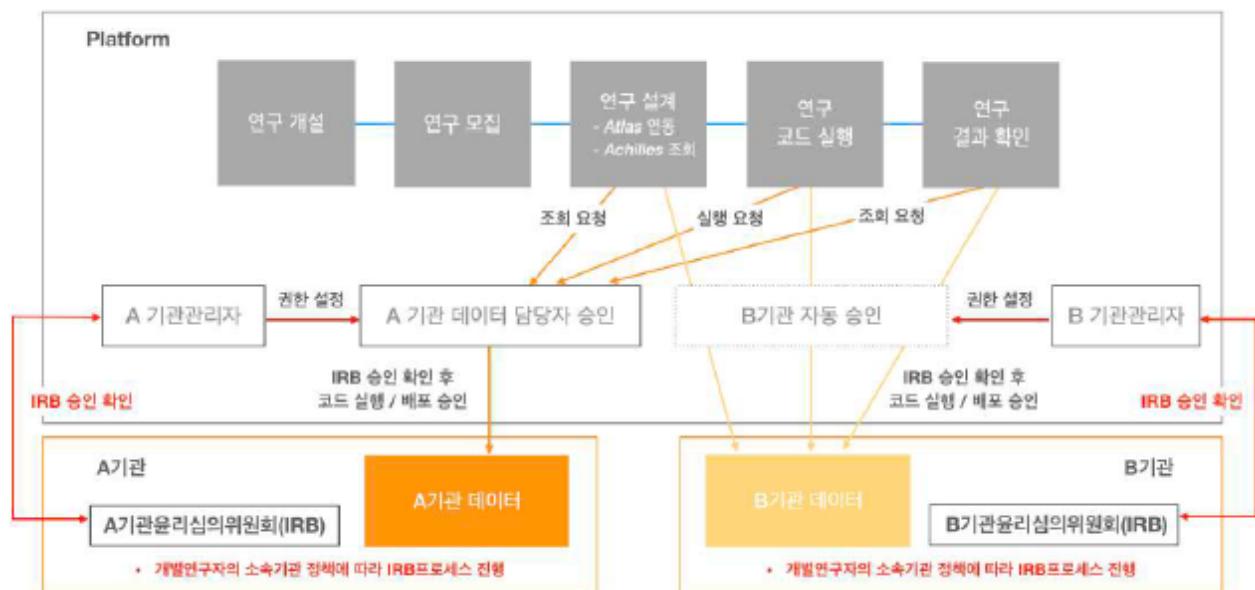
- 단일 기관 연구 결과 타기관(OHDSI 네트워크) 자료를 추가 분석할 필요가 있는 경우, 요청 기관 IRB 승인 내역과 분석 코드를 요청 대상 기관 송부함.
  - 요청을 받은 기관은 요청을 한 기관의 정책에 따라 IRB 프로세스를 진행하여 연구가 원활하게 진행될 수 있도록 함.
- 다기관 데이터 통합분석은 기본적으로 하나의 분석코드로 가능함.
  - 코드는 ‘cohort definition’의 ‘Export - JSON’ 템<sup>22)</sup>에서 복사하여 텍스트 파일로 송부함.
- 다기관 연구의 구체적인 진행은 FeederNet<sup>23)</sup>을 통해 수행하며, 구체적 수행 절차는 다음과 같음.

22) 코호트의 정의와 concept set 지원의 Java Script Object Notation (JSON) 형식의 표현

23) 부록2.FeederNet 매뉴얼 참조

- STEP 1. 연구 개설 : 프로젝트 생성
- STEP 2. 연구 모집 : 연구원 초대 -> 연구 계획 Discussion에서 공유
- STEP 3. 프로젝트 승인요청 (IRB 승인여부 확인)
  - IRB 승인상태 확인 요청 (기관관리자의 IRB승인 확인 및 권한 설정 시 다음 단계 수행 가능)
- STEP 4. 연구 설계
  - 일산병원 데이터 조회/통계 확인 → Atlas에서 연구설계 → 연구 코드 가져오기
  - 타기관 데이터 통계 조회
- STEP 5. 연구 실행
  - 일산병원 연구 코드 실행 -> 데이터담당자 승인
  - 타기관에서 연구 코드 실행
- STEP 6. 결과 조회
  - 일산병원 연구 결과 조회 -> 데이터담당자 승인
  - 타기관에서 연구 결과 조회

#### 연구 권한을 다르게 설정한 다기관 연구 흐름도



## 참 고 자 료

- 1) OHDSI. Retrieved December 1-31, 2018 from <http://www.ohdsi.org/web/wiki/doku.php?id=documentation:software:atlas>
- 2) 일산병원. Retrieved December 20-31, 2018 from <http://192.168.1.53:8080/Atlas/#/home>
- 3) 에비드넷(2018). CDM 기반 임상데이터 구축 및 운영관리 계약서
- 4) 에비드넷 CDM 분석팀(2018). CDM Interface Table 정의
- 5) NLM Retrieved December 1-31, <https://www.nlm.nih.gov/research/umls/rxnorm/docs/2015/appendix5.html>
- 6) 국민건강보험(2016). 연구용역보고서-2016-2-0002, 건강보험 빅데이터를 활용한 의약품 안전사용 모니터링 구축 방안
- 7) 에비드넷 [분산형 바이오 헬스 통합 데이터망 구축] 결과보고서
- 8) OHDSI/CommonDataModel. Retrieved January 1-31, 2019 <https://github.com/OHDSI/CommonDataModel/wiki>
- 9) Korean OHDSI. Retrieved January 1-31, 2019 from <https://www.youtube.com/watch?v=pMtJ3aBQ6sk>
- 10) 에비드넷(2019.1.31.). 참여병원 연구지원을 위한 Executor(SQL/R) 기술지원 포럼 플랫폼 MVP(제11차 리더십 미팅 발표 자료)
- 11) 에비드넷(2019). Feedernet 사용자 매뉴얼(일반연구자) v.1.1



# **부 록 1**

## **R환경설정 매뉴얼**



# R 환경 설정 매뉴얼

---

해당 버전: 1.1.0

## 주의사항:

이 문서 내에 있는 글에서는 R 3.5.0을 설치하고 있지만, 스크린샷은 R 3.2.2 설치 당시에 편집한 것을 사용하고 있습니다. 꼭 설치하실 때 글을 잘 읽고 R 3.5.0에 맞게 설치해주세요.

## 목차

---

1. 사전 점검 - Java 설치
2. R 설치
3. 패키지 설치
4. 이슈 및 트러블 슈팅

## 사전 점검 - Java 설치

---

OHDSI에서 사용하는 R 패키지들은 rJava를 사용하고 있습니다. rJava를 사용하기 위해서는 JDK 8 버전의 설치가 필요합니다. 본 문단에서는 Java를 설치하고 rJava와 연결하기 위해 `JAVA_HOME`이라는 환경 변수를 설정하는 것 까지 설명하고자 합니다.

## 주석:

분석 에이전트 인스톨러에서는 R과 Rtools를 자동으로 설치해주고 있습니다만, 라이센스 상의 이유로 Java는 별도로 설치가 필요합니다.

1. [Oracle JDK 8 다운로드 홈페이지](#)에 들어갑니다.

## Java SE Development Kit 8 Downloads

Thank you for downloading this release of the Java™ Platform, Standard Edition Development Kit (JDK™). The JDK is a development environment for building applications, applets, and components using the Java programming language.

The JDK includes tools useful for developing and testing programs written in the Java programming language and running on the Java platform.

See also:

- [Java Developer Newsletter](#): From your Oracle account, select **Subscriptions**, expand **Technology**, and subscribe to **Java**.
- [Java Developer Day hands-on workshops \(free\) and other events](#)
- [Java Magazine](#)

JDK 8u171 [checksum](#)

JDK 8u172 [checksum](#)

### Java SE Development Kit 8u171

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

Accept License Agreement

Decline License Agreement

Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	77.97 MB	<a href="#">jdk-8u171-linux-arm32-vfp-hflt.tar.gz</a>
Linux ARM 64 Hard Float ABI	74.89 MB	<a href="#">jdk-8u171-linux-arm64-vfp-hflt.tar.gz</a>
Linux x86	170.05 MB	<a href="#">jdk-8u171-linux-i586.rpm</a>
Linux x86	184.88 MB	<a href="#">jdk-8u171-linux-i586.tar.gz</a>
Linux x64	167.14 MB	<a href="#">jdk-8u171-linux-x64.rpm</a>
Linux x64	182.05 MB	<a href="#">jdk-8u171-linux-x64.tar.gz</a>
Mac OS X x64	247.84 MB	<a href="#">jdk-8u171-macosx-x64.dmg</a>
Solaris SPARC 64-bit (SVR4 package)	139.83 MB	<a href="#">jdk-8u171-solaris-sparcv9.tar.Z</a>
Solaris SPARC 64-bit	99.19 MB	<a href="#">jdk-8u171-solaris-sparcv9.tar.gz</a>
Solaris x64 (SVR4 package)	140.6 MB	<a href="#">jdk-8u171-solaris-x64.tar.Z</a>
Solaris x64	97.05 MB	<a href="#">jdk-8u171-solaris-x64.tar.gz</a>
Windows x86	199.1 MB	<a href="#">jdk-8u171-windows-i586.exe</a>
Windows x64	207.27 MB	<a href="#">jdk-8u171-windows-x64.exe</a>

### Java SE Development Kit 8u172

You must accept the [Oracle Binary Code License Agreement for Java SE](#) to download this software.

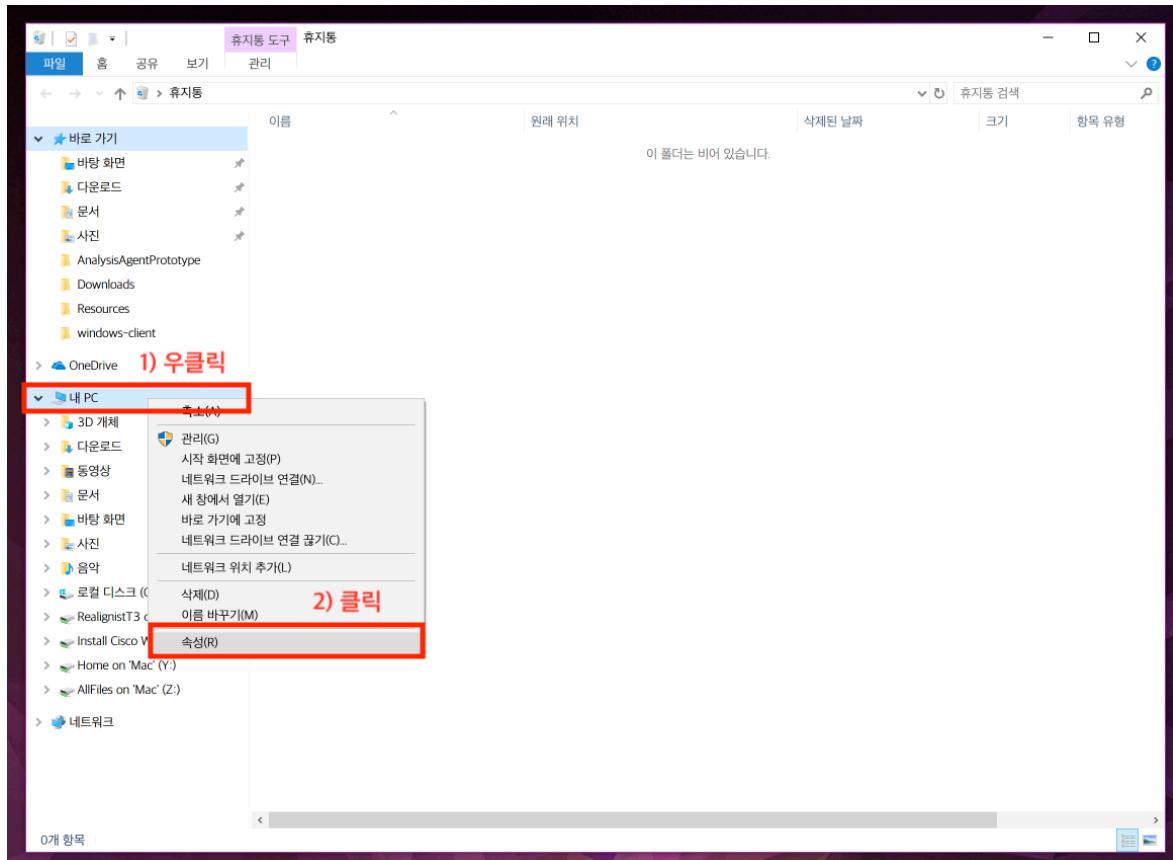
Accept License Agreement

Decline License Agreement

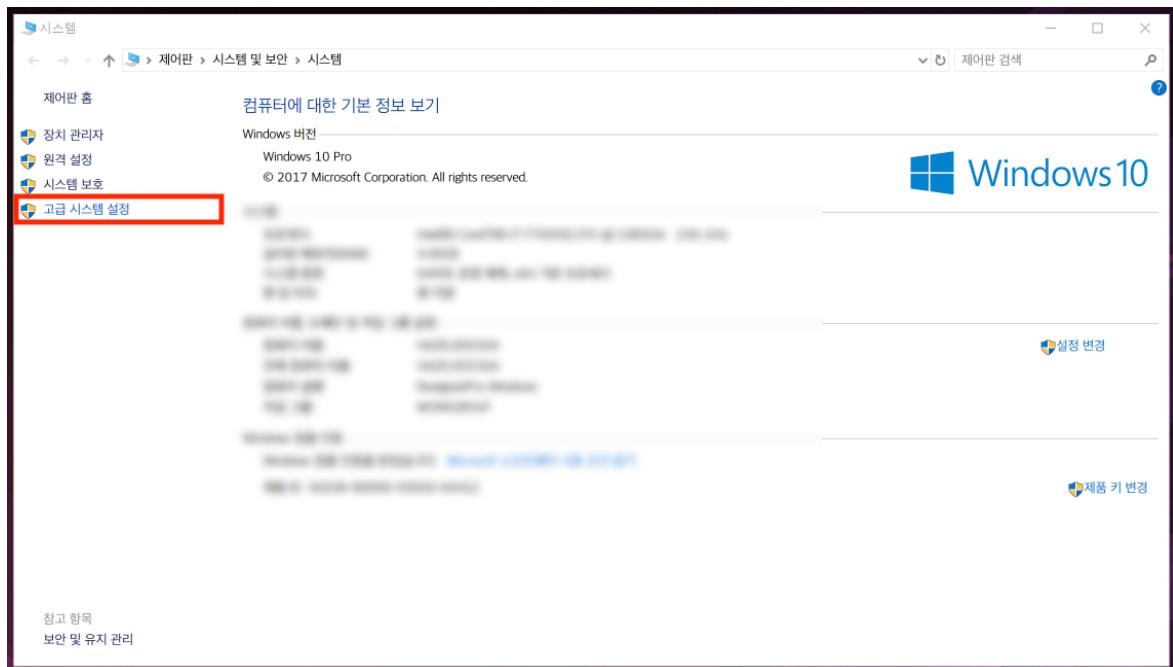
Product / File Description	File Size	Download
Linux ARM 32 Hard Float ABI	77.99 MB	<a href="#">jdk-8u172-linux-arm32-vfp-hflt.tar.gz</a>
Linux ARM 64 Hard Float ABI	74.9 MB	<a href="#">jdk-8u172-linux-arm64-vfp-hflt.tar.gz</a>
Linux x86	170.07 MB	<a href="#">jdk-8u172-linux-i586.rpm</a>
Linux x86	184.91 MB	<a href="#">jdk-8u172-linux-i586.tar.gz</a>
Linux x64	167.15 MB	<a href="#">jdk-8u172-linux-x64.rpm</a>
Linux x64	182.08 MB	<a href="#">jdk-8u172-linux-x64.tar.gz</a>
Mac OS X x64	247.87 MB	<a href="#">jdk-8u172-macosx-x64.dmg</a>
Solaris SPARC 64-bit (SVR4 package)	140.05 MB	<a href="#">jdk-8u172-solaris-sparcv9.tar.Z</a>
Solaris SPARC 64-bit	99.35 MB	<a href="#">jdk-8u172-solaris-sparcv9.tar.gz</a>
Solaris x64 (SVR4 package)	140.63 MB	<a href="#">jdk-8u172-solaris-x64.tar.Z</a>
Solaris x64	97.06 MB	<a href="#">jdk-8u172-solaris-x64.tar.gz</a>
Windows x86	199.1 MB	<a href="#">jdk-8u172-windows-i586.exe</a>
Windows x64	207.3 MB	<a href="#">jdk-8u172-windows-x64.exe</a>

- 자바 8까지의 JDK 버전 규칙은 {메이저 버전}u{マイ너 버전} 입니다. 마이너 버전 업데이트에는 보안 패치 등이 들어있으니 가급적 최신 버전의 사용을 권장합니다.
- 2. 오라클의 라이센스에 동의한 후, 현재 최신 버전(글을 쓴 시점에선 8u172 버전이 최신입니다.) 다운로드 받습니다. 본 프로그램은 64bit 윈도우에서의 사용을 전제로 만들어졌기 때문에, Windows x64 항목을 다운로드해야합니다.

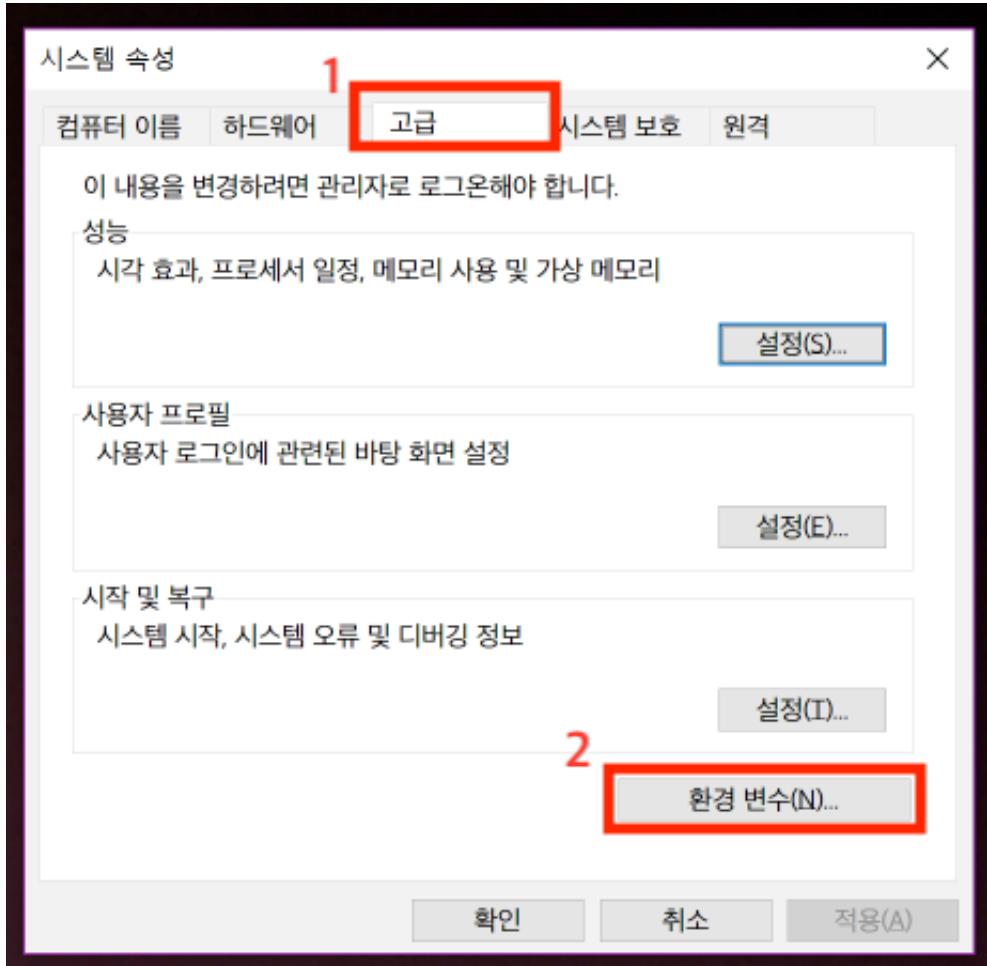
3. 인스톨러 파일을 다운로드 한 후, 실행하여 Java를 설치합니다.
4. 윈도우 탐색기에서 좌측 사이드바에 있는 "내 PC" 항목을 우클릭한 뒤, "속성" 항목을 클릭합니다.



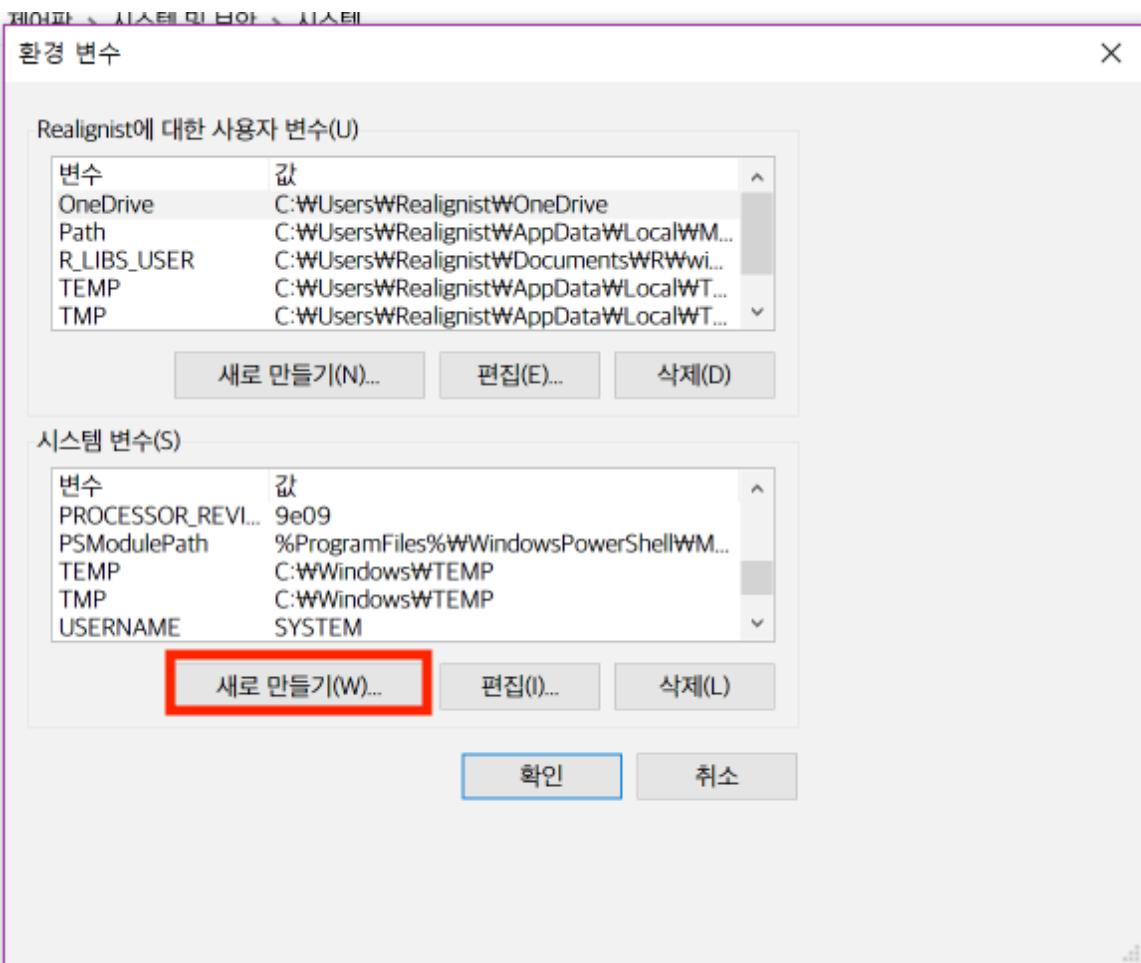
5. 좌측 메뉴의 "고급 시스템 설정"을 클릭합니다.



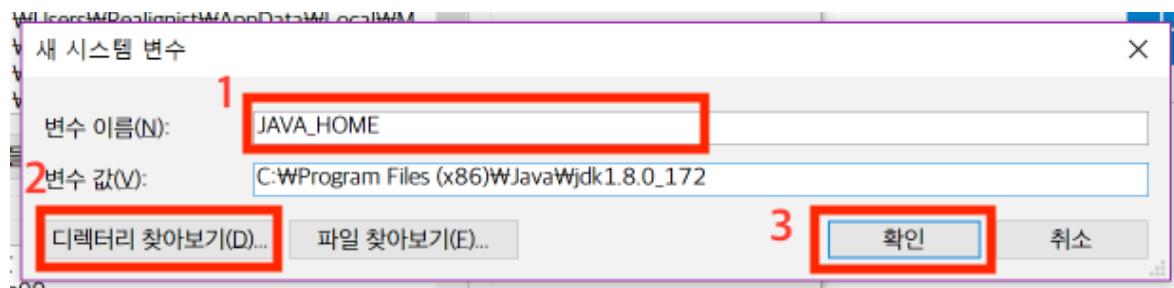
6. 시스템 설정 창의 "고급" 탭에서 하단에 있는 "환경 변수" 버튼을 눌러줍니다.



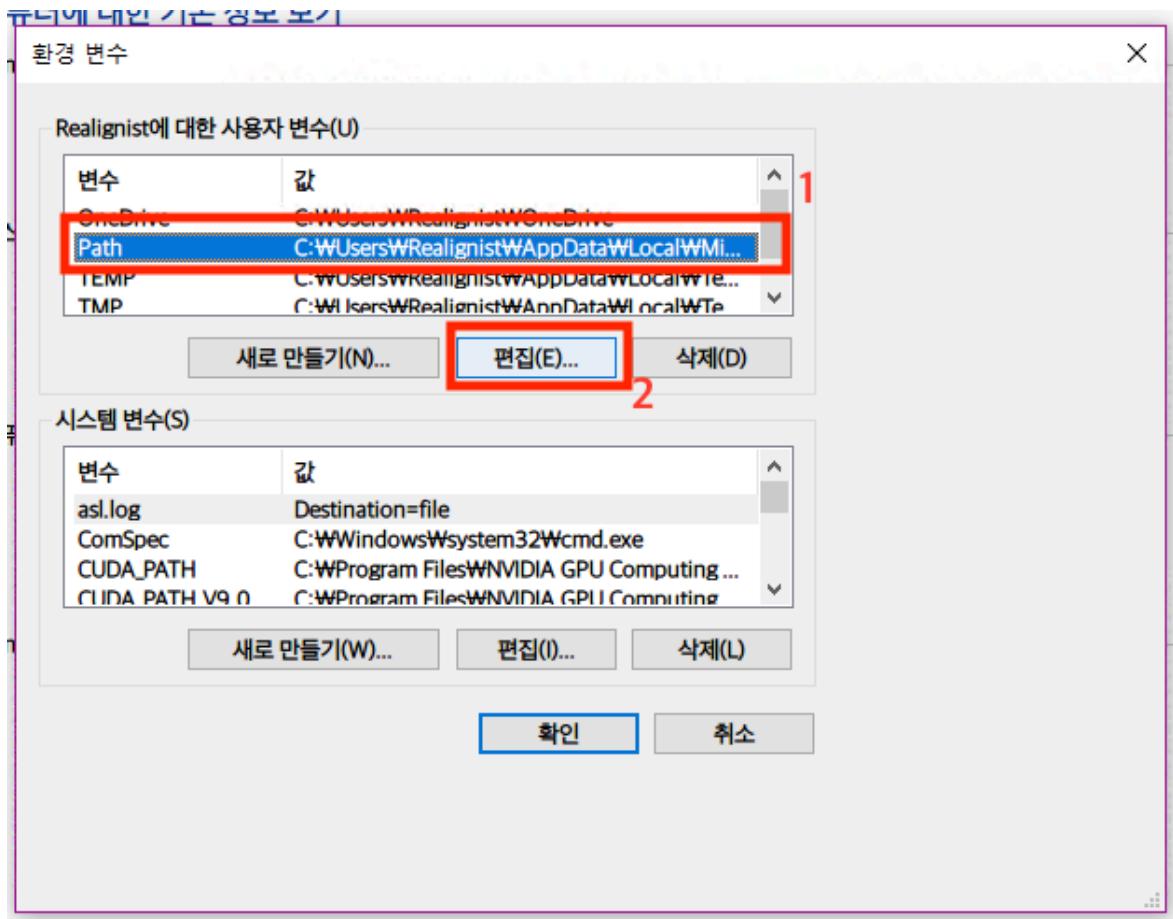
7. "시스템 변수" 항목의 "새로 만들기" 버튼을 클릭해줍니다.



8. 환경 변수 추가 창에서 변수 이름에 `JAVA_HOME` 이라고 입력해줍니다. 그 뒤 "디렉터리 찾아보기" 버튼을 눌러 위에서 설치한 자바의 경로를 선택합니다. (보통은 `C:\Program Files (x86)\Java` 폴더 안에 `jdk1.8`로 시작합니다.) 선택한 후, 변수 값에 제대로 선택이 된 걸 확인하고 "확인" 버튼을 누릅니다.



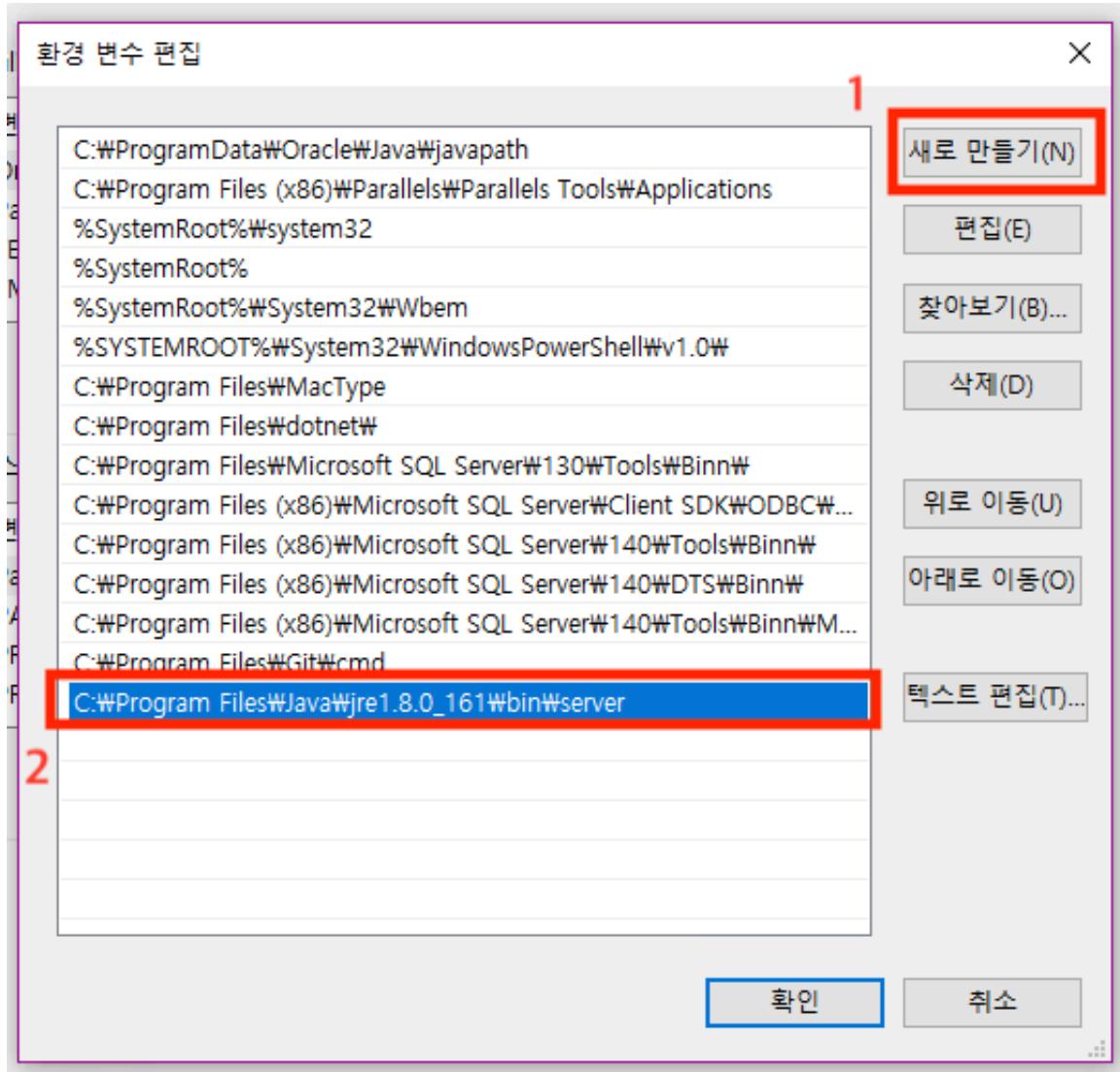
9. "사용자 변수" 항목의 `PATH` 변수를 누르고 "편집" 버튼을 클릭해줍니다.



10. "새로 만들기" 버튼을 클릭하고, `C:\Program Files\Java\jre1.8.0_{마이너 버전}\bin\server`를 추가해줍니다.

**경고:**

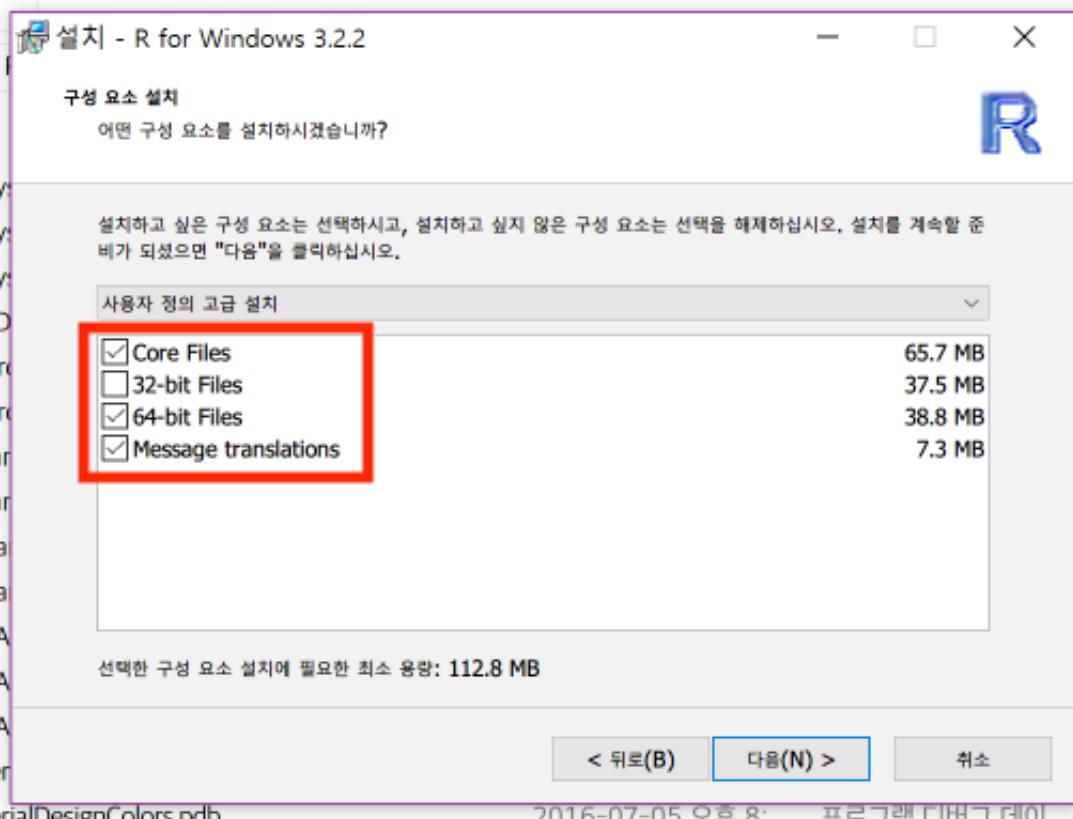
절대로 `C:\Program Files\Java\jre1.8.0_{마이너 버전}\bin\server`를 그대로 복사 / 붙여넣기해선 안됩니다. {마이너 버전} 을 지우고 자신이 받은 JDK 버전(8u172라면 172가 마이너 버전입니다.)을 넣어야 합니다.



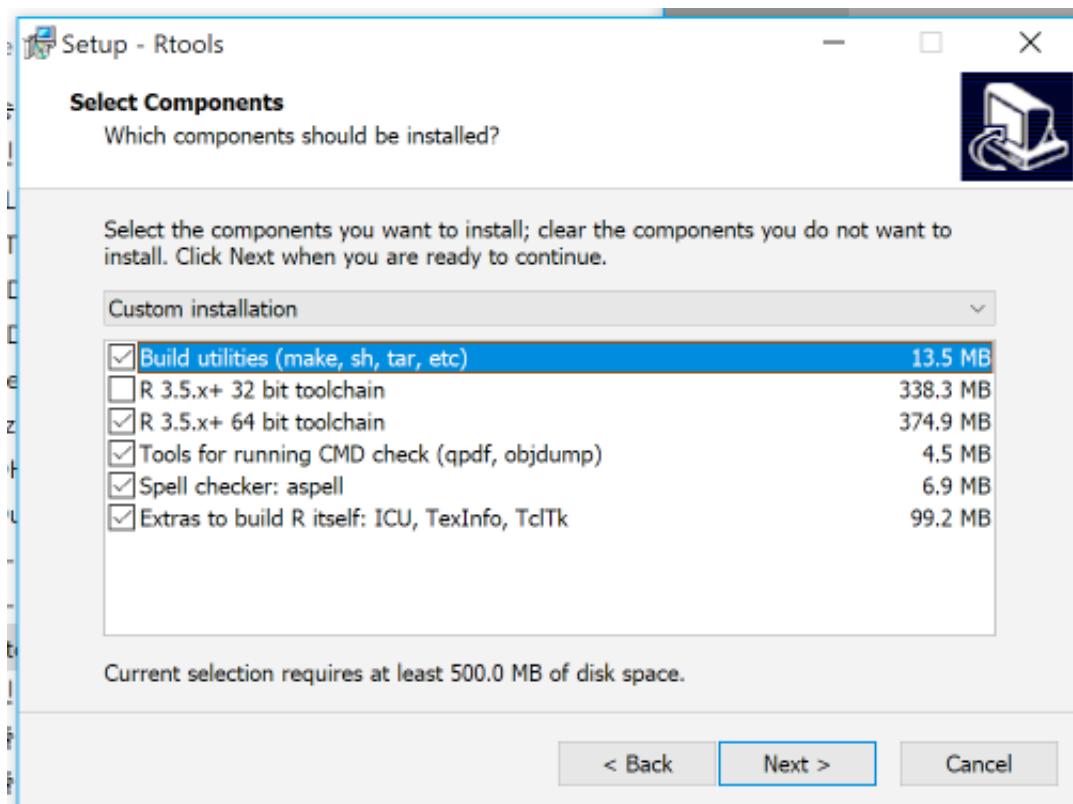
## R 설치

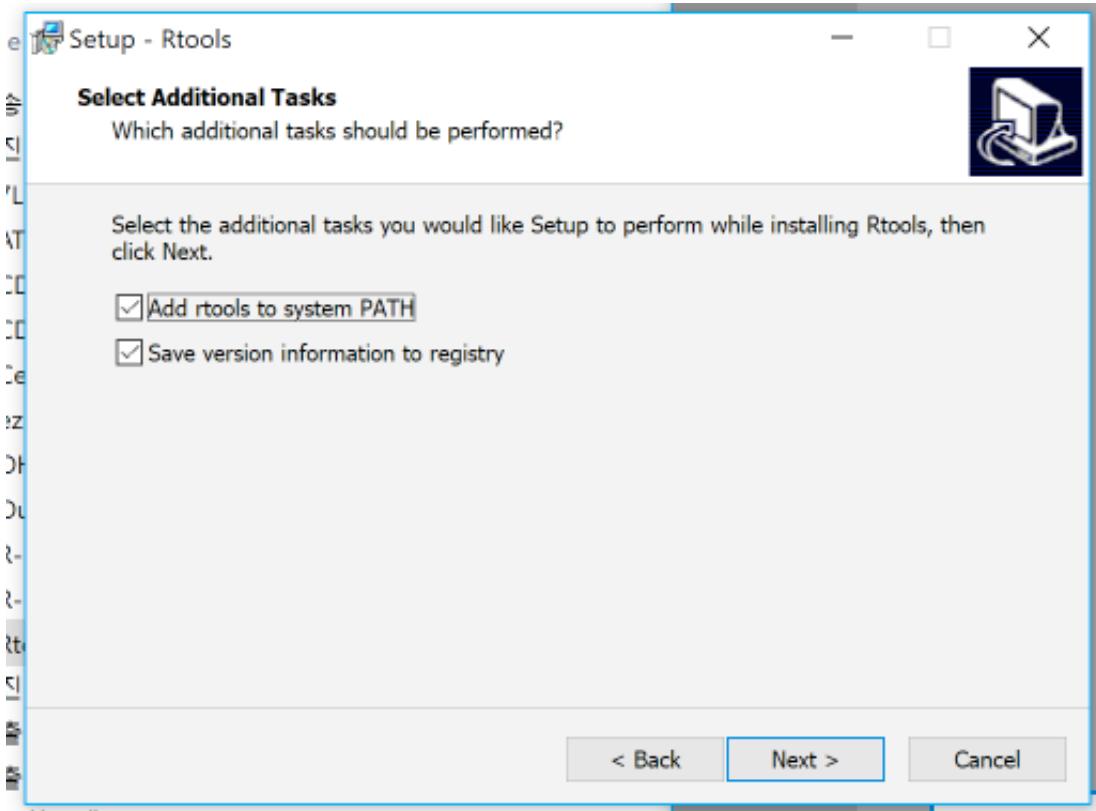
본 문단에서는 프로그램이 정상적으로 설치되지 않았을 때나 트러블 슈팅을 위해서 자동 설치 후, 수동 설치 과정을 안내하고자 합니다.

1. [링크](#)에서 R 3.5.0 설치 파일을 받습니다.
2. `R-3.5.0-win.exe` 파일을 실행해서 R 3.5.0 버전을 설치합니다.
3. R 설치 파일에서 설치할 컴포넌트를 고를 때, **32bit를 제외한 전체**를 설치합니다.

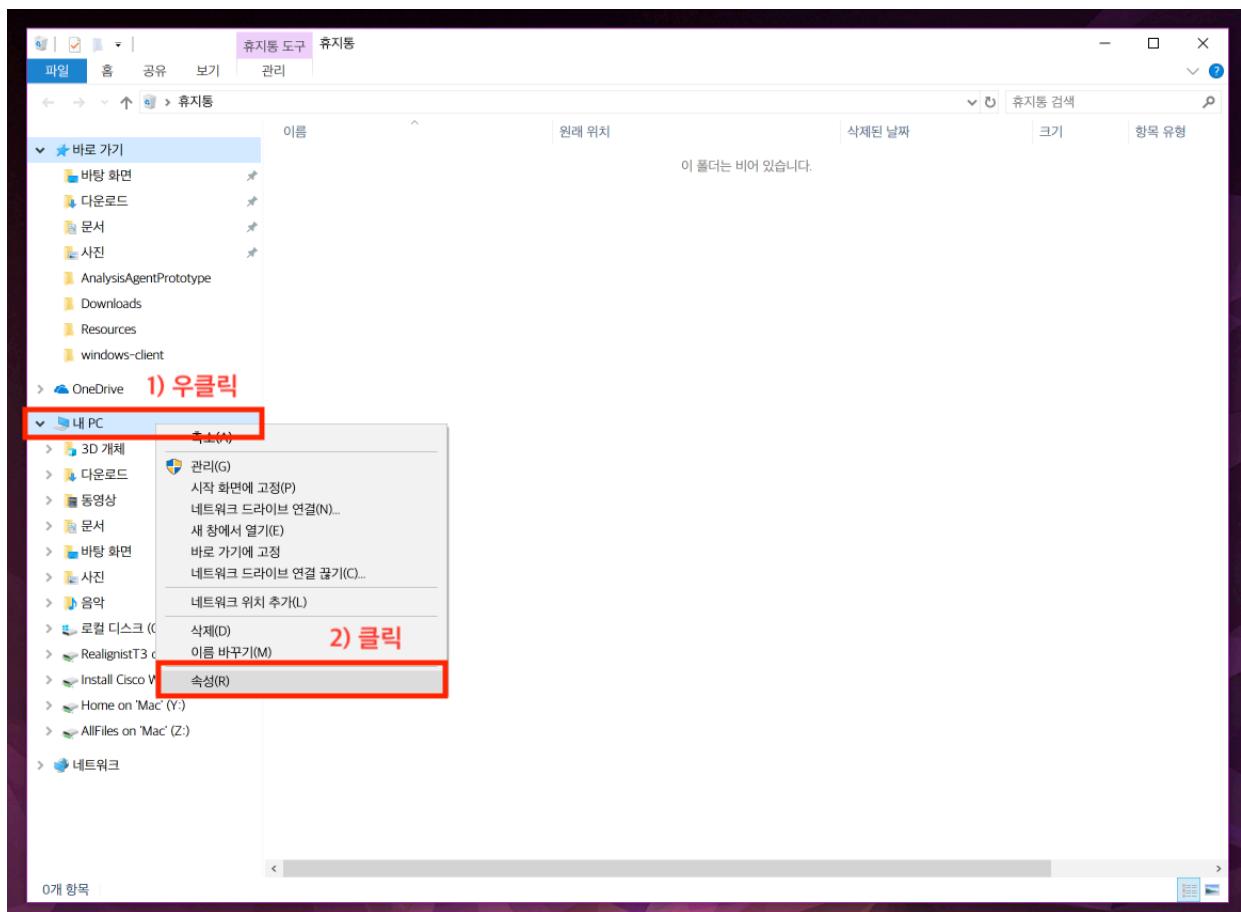


4. 링크에서 Rtools 3.5 설치 파일을 받습니다.
5. `rtools.exe` 파일을 실행해서 Rtools를 설치합니다.
6. Rtools 설치 파일에서 설치할 컴포넌트를 고를 때, 다음 항목들만을 설치합니다.

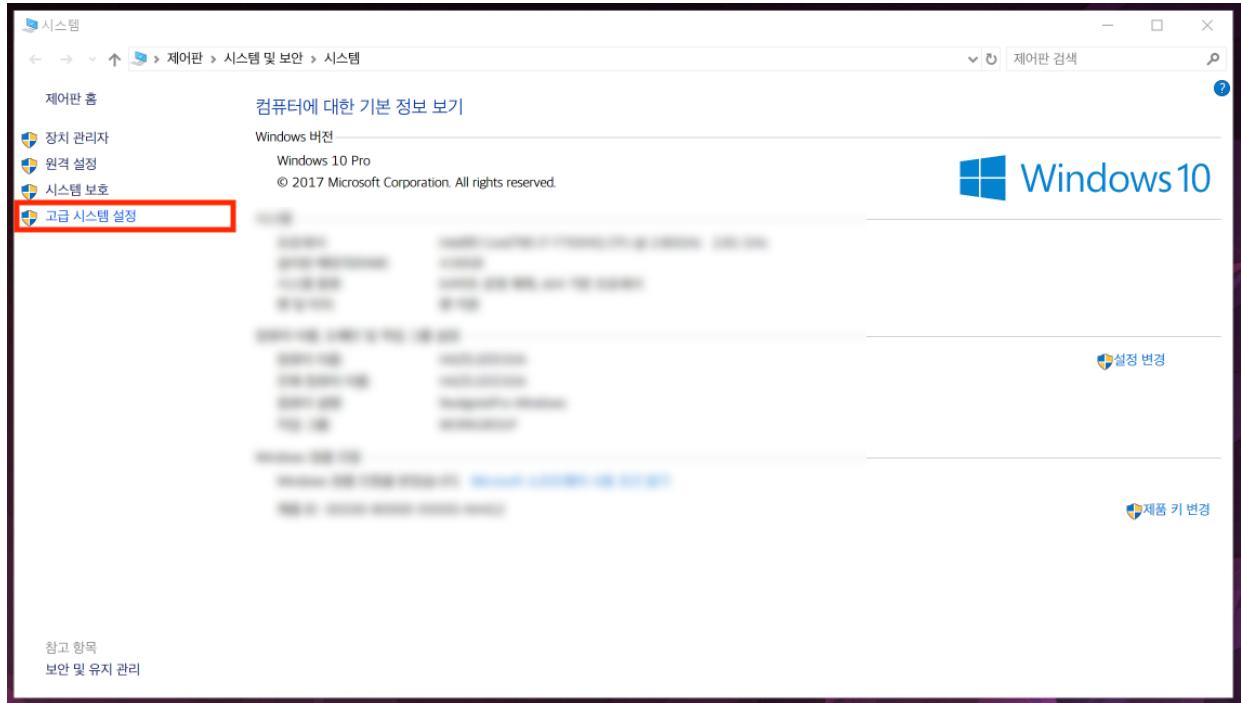




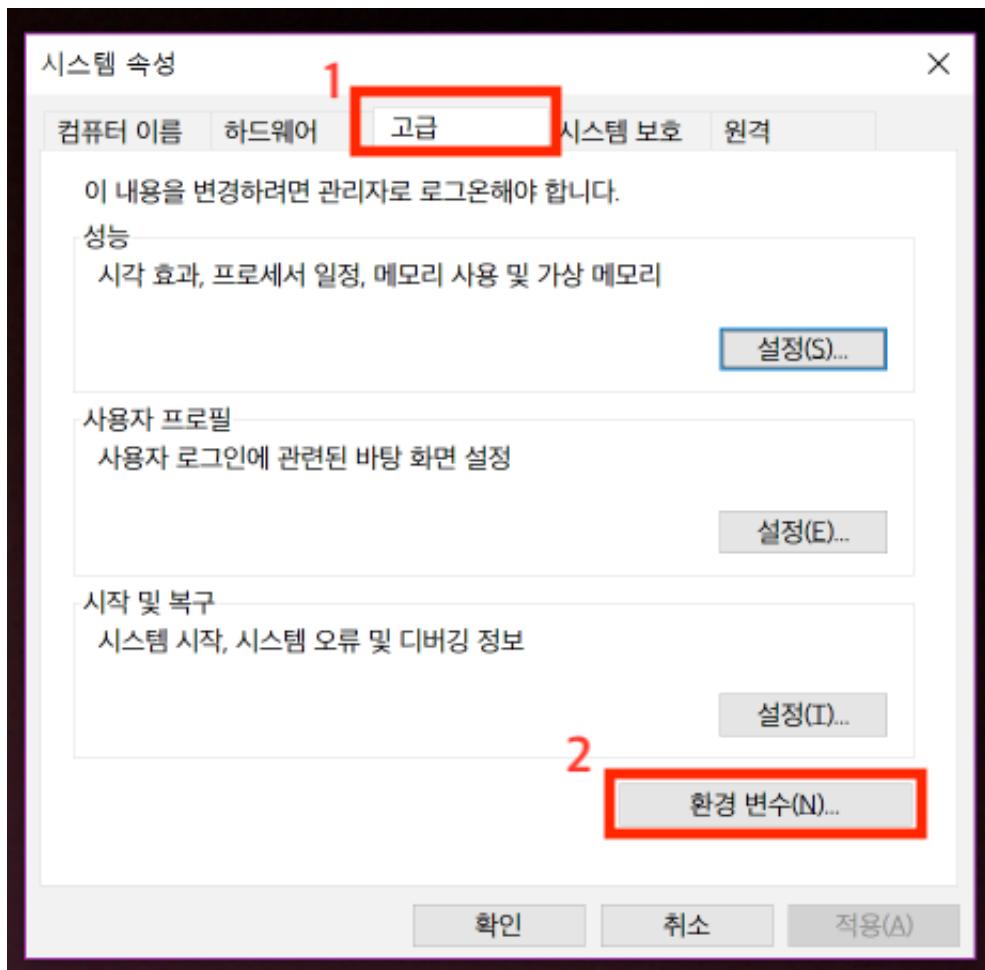
7. 원도우 탐색기에서 좌측 사이드바에 있는 "내 PC" 항목을 우클릭한 뒤, "속성" 항목을 클릭합니다.



8. 좌측 메뉴의 "고급 시스템 설정"을 클릭합니다.



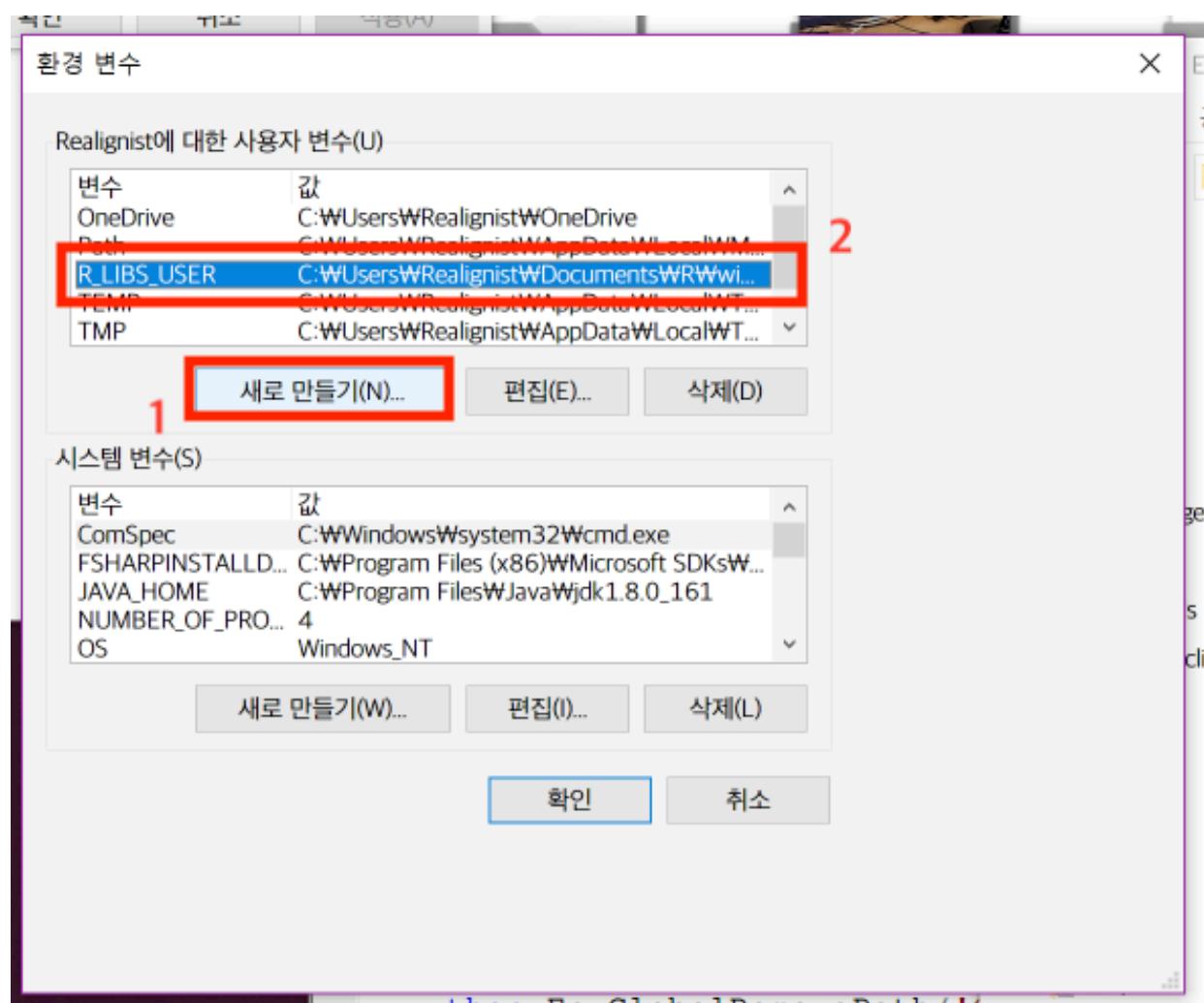
9. 시스템 설정 창의 "고급" 탭에서 하단에 있는 "환경 변수" 버튼을 눌러줍니다.



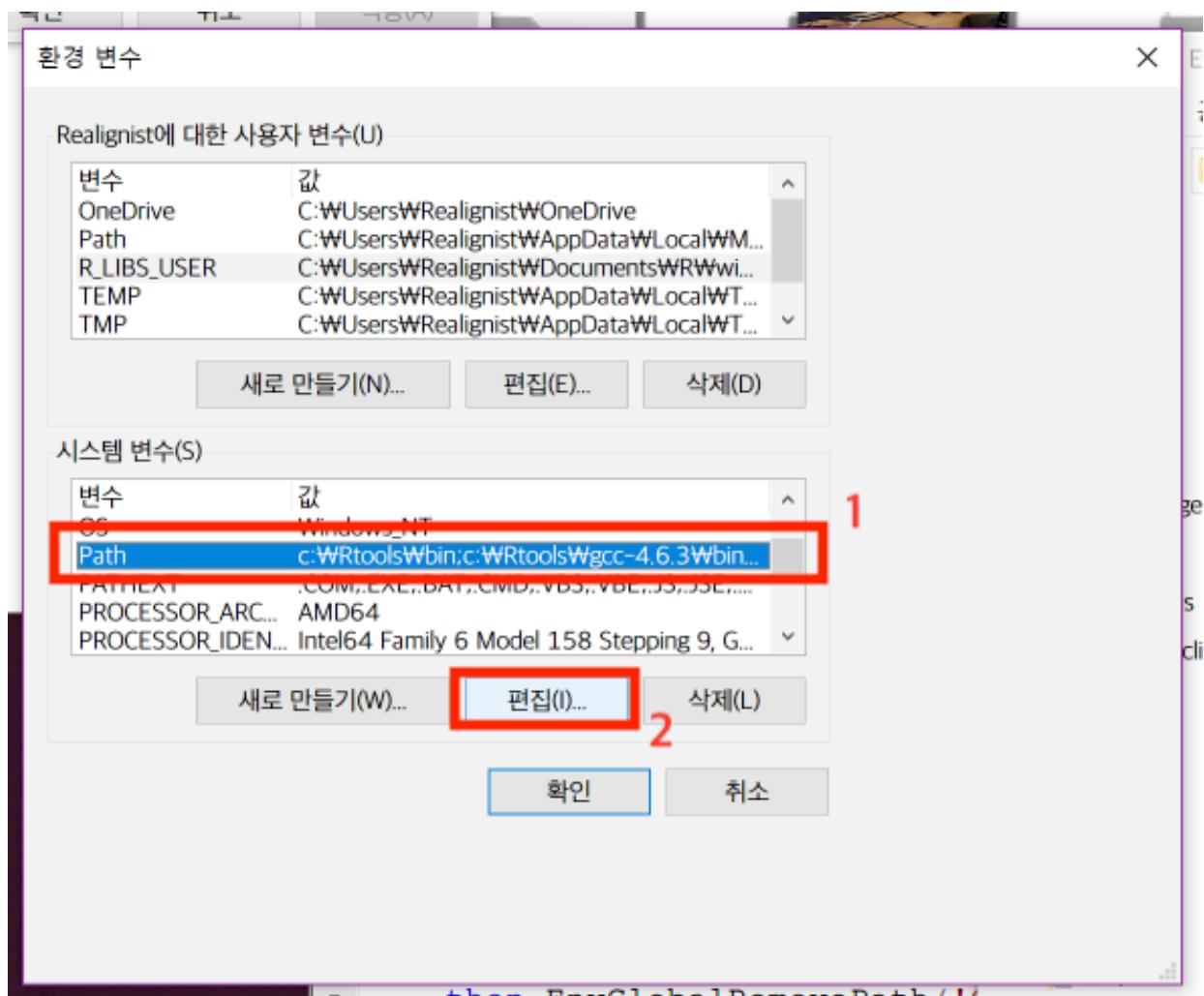
10. "사용자 변수" 항목에서 "새로 만들기" 버튼을 누르고, 변수 이름은 `R_LIBS_USER`를 변수 값에는 `C:\Users\{사용자 이름}\Documents\R\win-library\3.5` 또는 자신이 R 패키지 설치에 사용할 폴더를 입력해줍니다.

경고:

절대로 C:\Users\{사용자 이름}\Documents\R\win-library\3.5 를 그대로 복사해선 안됩니다.  
{사용자 이름} 을 지우고 자신의 사용자 이름을 넣어줘야합니다. 값을 입력하기 전에 C:\Users\{사용자 이름}\Documents 폴더가 존재하는지 확인하세요.

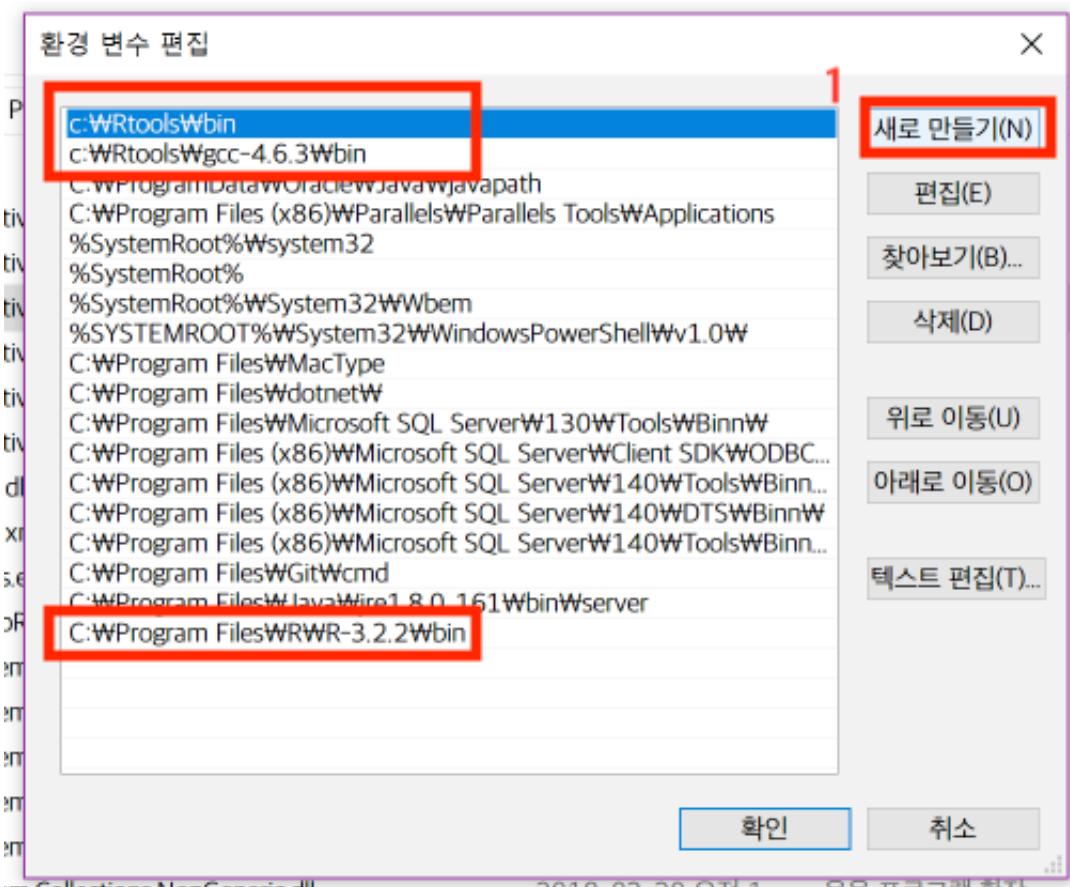


11. "시스템 변수" 항목에서 Path 항목을 찾아 선택한 후, "편집" 버튼을 누릅니다.



12. 리스트에 다음 항목들이 있는지 확인한 뒤, 없다면 "새로 만들기" 버튼을 눌러 다음 항목들을 추가해줍니다.

- C:\Rtools\bin
- C:\Rtools\mingw\_64\bin
- C:\Program Files\R\R-3.5.0\bin



13. 이후 마지막으로 컴퓨터를 재부팅합니다.

## OHDSI 패키지 설치

1. 아래와 같은 R 코드를 한줄 한줄 실행합니다.
2. 주의: 다음 스크립트는 ATLAS 2.6 이상을 위한 환경 설정용 스크립트입니다.

```

install.packages("devtools")
install.packages("drat")
drat::addRepo("OHDSI")
library(devtools)

install.packages("SqlRender")
install_github("ohdsi/DatabaseConnector", ref = "v2.2.0")
install_github("ohdsi/OhsdiSharing", ref = "v0.1.3")
install_github("ohdsi/FeatureExtraction", ref = "v2.1.5")
install_github("ohdsi/CohortMethod", ref = "v3.0.1")
install_github("ohdsi/EmpiricalCalibration", ref = "v1.3.6")
install.packages("MethodEvaluation")

# 아킬레스 및 Prediction 연구 시에만 설치
devtools::install_github("ohdsi/Achilles")
devtools::install_github("ohdsi/PatientLevelPrediction", ref="v3.0.0")

```

3. ATLAS 2.5까지는 다음 스크립트를 사용해주세요.

```
install.packages("drat")
drat::addRepo("OHDSI")
install.packages("BigKnn")

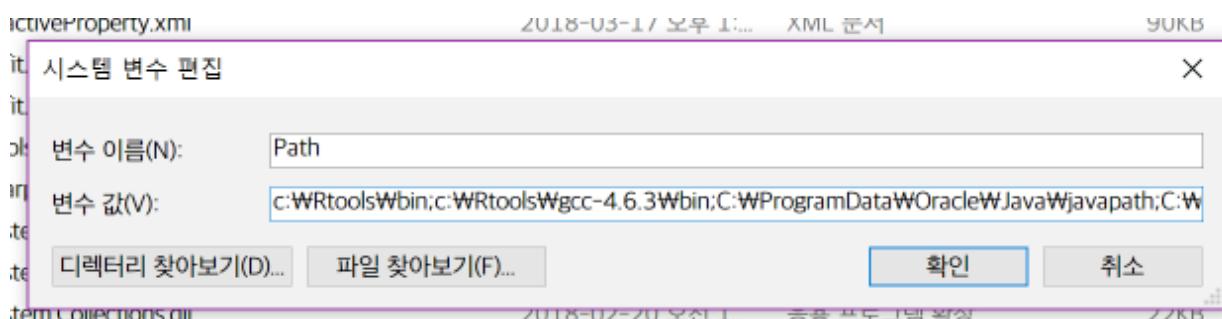
devtools::install_github("ohdsi/SqlRender", ref="v1.5.0")
devtools::install_github("ohdsi/DatabaseConnector", ref="v2.0.5")
devtools::install_github("ohdsi/OhsdsiRTools", ref="v1.5.5")
devtools::install_github("ohdsi/FeatureExtraction", ref="v2.0.2")
devtools::install_github("ohdsi/CohortMethod", ref="v2.6.2")
devtools::install_github("ohdsi/EmpiricalCalibration", ref="v1.3.6")

# 아킬레스 및 Prediction 연구 시에만 설치
devtools::install_github("ohdsi/Achilles")
devtools::install_github("ohdsi/PatientLevelPrediction", ref="v2.0.5")
```

## 이슈 및 트러블 슈팅

## 원도우 10 이전 버전에서의 PATH 환경 변수 편집

원도우 10 이전 버전, 혹은 Windows Server 2016 이전 버전에서는 PATH 환경 변수를 수정할 때도 다음과 같이 텍스트로만 수정해야합니다.



여러 개의 항목을 나열할 때 우리가 **,** 를 사용해서 표현하듯, 환경변수에서는 여러 개의 폴더를 나열할 때 매 폴더 경로의 마지막에 **:** 를 사용합니다.

자연어 예시: 사과, 배, 감, 복숭아

환경변수 예시: C:\Rtools\bin;c:\Rtools\gcc-

4.6.3\bin;C:\ProgramData\Oracle\Java\javapath;

## **부 록 2**

# **FeederNet 사용자 매뉴얼**

## **(일반연구자, v.1.1)**



# FeederNet®

사용자 매뉴얼(일반연구자) v.1.1

# Index

## FeederNet® 이란?

## FeederNet®에서 연구하기

다기관 연구의 순서를 확인 할 수 있어요

## 회원가입

## 프로젝트

- 프로젝트 생성 및 초대
- Dashboard
- Analyses
- IRB
- Data
- Members
- Permissions
- Discussion

## Appendix

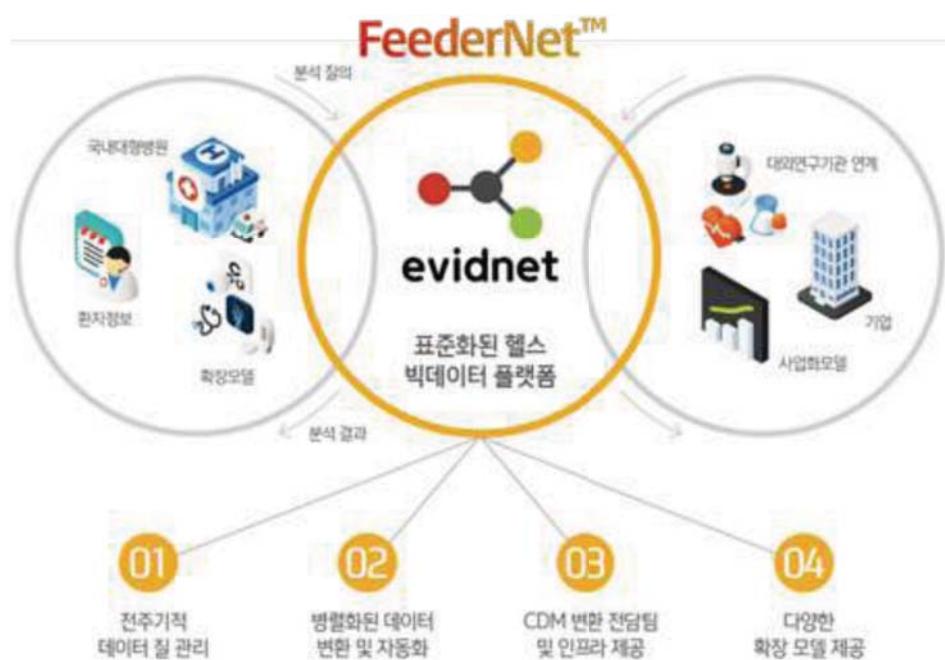
# FeederNet® 이란?

FeederNet®은 환자들의 표준화된 데이터를 개인정보 제공이나 공유없이 분산 형태로 관리하면서 다기관 데이터를 분석해 의학 연구에 활용가능한 통계 결과데이터를 제공하고, 공동 연구 네트워크를 구축하는 CDM 기반 분산연구망 플랫폼입니다.

현존하는 병원 간 이형적인 데이터에 따른 데이터 공유의 어려움 문제를 데이터 표준화 이후 FeederNet®을 통해 기관 간 데이터를 완벽하게 상호 운용하여 해결할 수 있으며, 이 때 데이터는 한 곳에 모아지는 것이 아니라 각 병원에 그대로 두고 통계적 분석 결과만을 연구자에게 제공합니다. 따라서 개별 환자의 정보에 접근하지 않고, 기관 내 의료정보 데이터를 반출하지 않으면서 개인정보 제공 및 공유 없이 많은 연구자가 공동 활용할 수 있습니다.

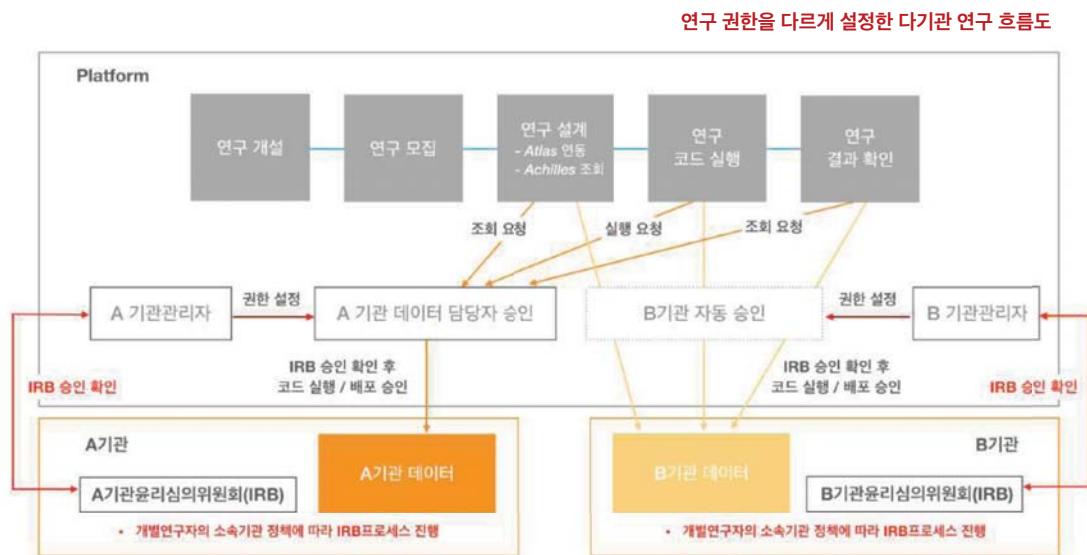
또한 하나의 분석코드로 다기관 데이터 통합분석이 가능하며, 연구에 소요되는 시간을 단축시킬 수 있습니다. 뿐만 아니라, FeederNet®을 통해 전주기적 데이터의 질 관리는 물론 병렬화 된 데이터의 변환 및 자동화, CDM 변환 인프라, 나아가 다양한 확장모델을 제공받을 수 있어 보다 효율적이고 효과적인 연구를 가능하게 합니다.

이제 Data가 아닌 Evidence를 공유하는 CDM 기반 분산연구망 플랫폼 FeederNet®을 통해 혁신적인 임상 의학연구를 실현하실 수 있습니다.



# FeederNet®에서 연구하기

다기관 연구



## A기관 연구원 기준으로 한 다기관 연구 순서

### STEP 1. 연구 개설

- 프로젝트 생성

### STEP 2. 연구 모집

- 연구원 초대 -> 연구 계획 Discussion에서 공유

### STEP 3. 프로젝트 승인요청 (IRB 승인여부 확인)

- IRB 승인상태 확인 요청 (기관관리자의 IRB승인 확인 및 권한 설정 시 다음 단계 수행 가능)

### STEP 4. 연구 설계

- A 기관 데이터 통계 조회 -> 데이터 담당자 승인 -> Atlas에서 연구 설계 -> 연구 코드 가져오기
- B 기관 데이터 통계 조회

### STEP 5. 연구 실행

- A 기관에서 연구 코드 실행 -> 데이터 담당자 승인
- B 기관에서 연구 코드 실행

### STEP 6. 결과 조회

- A 기관에서 연구 결과 조회 -> 데이터 담당자 승인
- B 기관에서 연구 결과 조회

# 회원가입

- feedernet.co.kr 접속

로그인 화면

The screenshot shows the Feedernet login page. It has fields for '이메일 주소' (Email Address) and '비밀번호' (Password). Below these fields is a link '로그인 정보를 잊으셨나요?' (Forgot your login information?). At the bottom left is a red-bordered button labeled '가입하기' (Sign Up). On the right side, there is a dark button labeled '로그인' (Login).

- 가입하기 클릭

약관 동의서 화면

This screenshot displays the '약관 동의서' (Agreement Consent) page. It contains several checkboxes for accepting terms and conditions:

- 이용약관, 개인정보 수집 및 이용, 보안서약서에 모두 동의합니다.
- FeederNet(피더넷) 플랫폼 서비스 이용약관 (필수)
- 개인정보 처리방침 (필수)
- 개인정보 수집 및 이용에 대한 동의 (필수)
  - [필수사항] 개인정보 수집 및 이용에 대한 동의  
주식회사 에비드넷(이하 '회사')은 정보통신망 이용촉진 및 정보보호 등에 관한 법률, 개인정보보호법, 통신비밀보호법, 전기통신사업법 등 정보통신서비스제공자가 준수하여야 할 관련 법령상의 개인정보보호 규정을 준수합니다. 취급하는 모든 개인정보는 관련 법령에に基づ거나 정보주체의 동의에 이어져 수집, 보유 및 처리되고 있으며 개인 정보 처리
  - ① 회의 가이 및 과리 • 회의가이 의사화이 회의제 서비스 이용 유지 및 과리 보이 시변 및 이즈 각주 고지 및 토지
- 보안서약서 동의 (필수)
  - 본인은 본인 소속기관의 데이터를 이용함에 있어 다음 사항을 엄숙히 서약한다.
    - 본인은 벨몬트 리포트, ICH-GCP E6 등 관련 지침의 윤리적 원칙에 따라 연구를 수행할 것이며 이를 위하여 관련 법규와 소속 기관의 연구관련 규정 및 지침을 준수한다.
    - 본인은 소속기관의 기관윤리심의위원회(Institutional Review Board)의 지침을 준수한다.

③ 보이오 개인정보 보호법 및 생명윤리 및 안전에 관한 법률을 포함한 구내부과 헤시키 서어 두 규제 지침을 준수하

At the bottom center is a large blue button with the text '동의합니다.' (I Agree.) which is also highlighted with a red border.

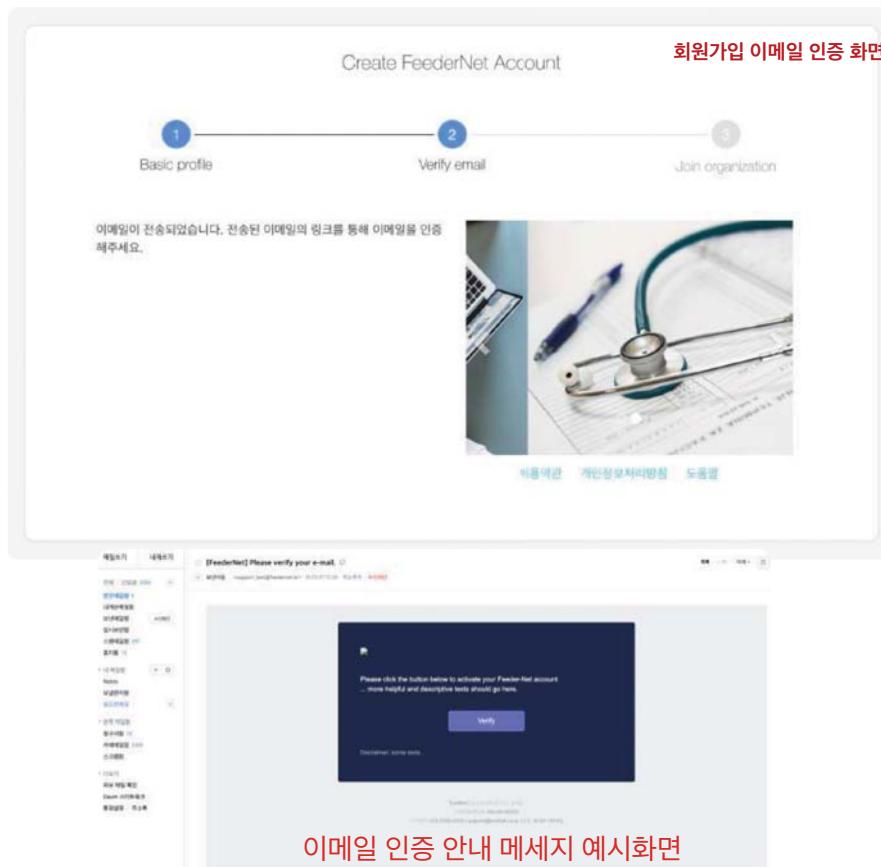
- 이용약관, 개인정보처리방침, 개인정보 수집 및 이용 약관, 보안서약서 약관을 모두 확인하고 동의합니다 클릭

# 회원가입

- 사용자 기본정보 입력

회원가입 화면

- 이메일주소는 사용자 소속기관의 이메일 주소 입력을 권장함
- 비밀번호는 영문,숫자,특수기호 조합 8자 이상으로 입력함



- 입력한 이메일 계정으로 발송된 이메일 인증 안내 메세지에 따라 인증 진행

# 회원가입

회원가입 완료 화면

The screenshot shows the final step of account creation, 'Join organization'. A success message at the top says '가입이 완료되었습니다.' (Account creation completed). Below it, a note states: '기관 소속 연구자의 경우 해당 기관으로부터 인증을 받아 기관 데이터에 대한 접근을 요청할 수 있습니다.' (If you are a researcher affiliated with an institution, you can request access to the institution's data by accepting authentication from the institution). Two bullet points provide instructions: '소속 기관 이메일로 인증을 완료하면, 자동으로 기관 연구자 등록이 신청됩니다.' (After completing authentication via the institution's email, automatic registration as an institutional researcher will be requested) and '이후에 언제든지 [설정] > [이메일] 메뉴를 통해서 이메일을 관리하실 수 있습니다.' (You can manage your email through the [Setting] > [Email] menu at any time). The main form area shows the primary email 'youmi.lee@evidnet.com' and a secondary email 'another@evidnet.com' which is currently 'Unverified'. A large '가입완료' (Completion) button is at the bottom right.

- Primary 메일은 '로그인' 또는 '비밀번호 재설정'에 사용 되는 메일이며, 추가 인증된 다른 이메일로 변경이 가능함

추가 인증 이메일 화면

This screenshot shows the 'Verify email' step of the account creation process. It displays a modal window asking if the user wants to receive a verification email to 'qa.feedernet.kr'. The modal has two buttons: '취소' (Cancel) and '확인' (Confirm). The main form below shows the primary email 'youmi.lee@evidnet.com' and a secondary email 'another@evidnet.com' which is currently labeled 'Unverified'. A red box highlights the 'Unverified' status of the secondary email. A large '가입완료' (Completion) button is at the bottom right.

- 추가한 이메일의 인증 하려면 이메일주소 우측에 **Unverified** 버튼을 클릭하여 인증 이메일 전송을 함
- 이메일을 추가함 (선택사항)

# 프로젝트

## 프로젝트 생성 및 초대



- '신규 프로젝트' 버튼 클릭 하여 새로운 프로젝트 생성

A screenshot of the 'Create Project' dialog box. It has fields for 'Project Name\*' (Project name), 'Study Type' (Select Study Type), and 'Project Description' (A brief description of the project for other users to understand). Below these are 'Visibility' buttons for 'Private' and 'Public'. At the bottom right are 'Cancel' and 'Create Project' buttons, with the latter being a blue button and having a red box highlighting it.

- 프로젝트 정보 입력 후 프로젝트 생성 클릭
- **프로젝트명** : 프로젝트를 구분할 수 있는 제목 입력
- **Study Type**: Comparative Effectiveness, Safety Surveillance, Others
- **프로젝트 소개** : 프로젝트를 소개하는 설명 입력
- **공개여부:**
  - **Private**(참여연구원들만 조회 가능)
  - **Public** (현재 MVP 버전 이후 지원예정)

A screenshot of the 'Research Project' list. It shows a table with columns for ID, NAME, OPEN SCOPE, STATUS, and STARTED AT. A new project entry is visible with the ID '87', NAME 'Test', OPEN SCOPE 'Private', STATUS 'Open', and STARTED AT 'May 9, 2019 1:16 PM'. This row is highlighted with a red box.

- 프로젝트 생성 후 메인 화면에서 생성 된 프로젝트 확인 가능

# 프로젝트

Dashboard

Dashboard 화면

The screenshot shows a project dashboard interface. At the top, there's a header with a lock icon and the text 'Test'. Below the header is a navigation bar with tabs: 'Dashboard' (which is highlighted with a red box), 'Analyses', 'IRB', 'Data', 'Members', 'Discussion', and 'Settings'. The main content area is titled '프로젝트 대시보드'. Underneath this, there's a section titled '프로젝트 활동내역' (Project Activity Log). It contains two entries:

- [pl idc1]님이 [Test] 프로젝트에 관리자로 참여하였습니다.  
May 9, 2019 1:40 PM
- [user idc1]님이 [pl idc1]님을 프로젝트에 관리자로 초대하였습니다.  
May 9, 2019 1:39 PM

- 생성한 프로젝트의 활동내역에 대하여 확인 가능

# 프로젝트

Analyses > Analysis 생성

## 생성한 프로젝트의 분석(Estimation) 실행방법 예시

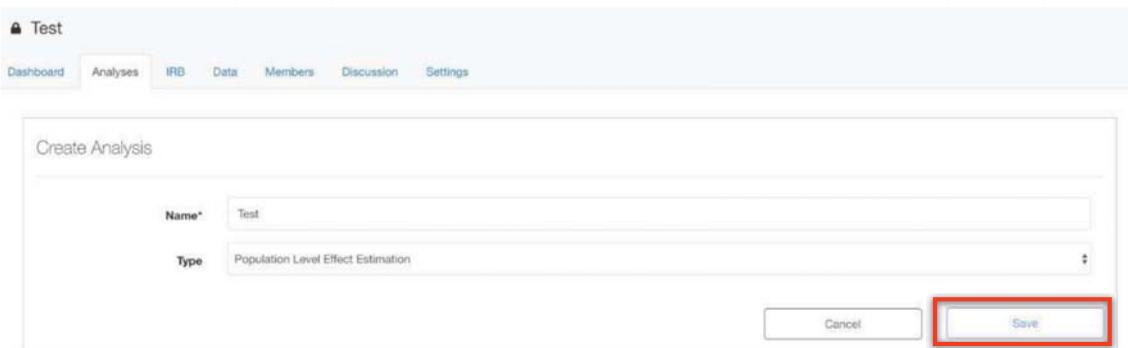
Analyses main 화면



Analyses main 화면 캡처입니다. 헤더에는 'Dashboard', 'Analyses' (선택됨), 'IRB', 'Data', 'Members', 'Discussion', 'Settings'가 있습니다. 검색 필드에는 'Search'가 있고, 오른쪽에는 'New Analysis' 버튼이 보입니다. 테이블 헤더에는 'ANALYSIS ID', 'NAME', 'TYPE', '연구자', '수정'があります. 테이블 아래에는 'No result.'라고 표시됩니다.

- 연구 분석을 진행하려는 프로젝트에 접속하여 Analyses 화면에서 'New Analysis' 클릭

New Analyses 화면



New Analyses 화면 캡처입니다. 헤더에는 'Dashboard', 'Analyses' (선택됨), 'IRB', 'Data', 'Members', 'Discussion', 'Settings'가 있습니다. 중앙에는 'Create Analysis' 폼이 있습니다. 'Name\*' 필드에 'Test'가 입력되어 있고, 'Type' 필드에는 'Population Level Effect Estimation'이 선택되어 있습니다. 오른쪽에는 'Cancel'과 'Save' 버튼이 있습니다. 'Save' 버튼은 빨간색 박스로 표시되어 있습니다.

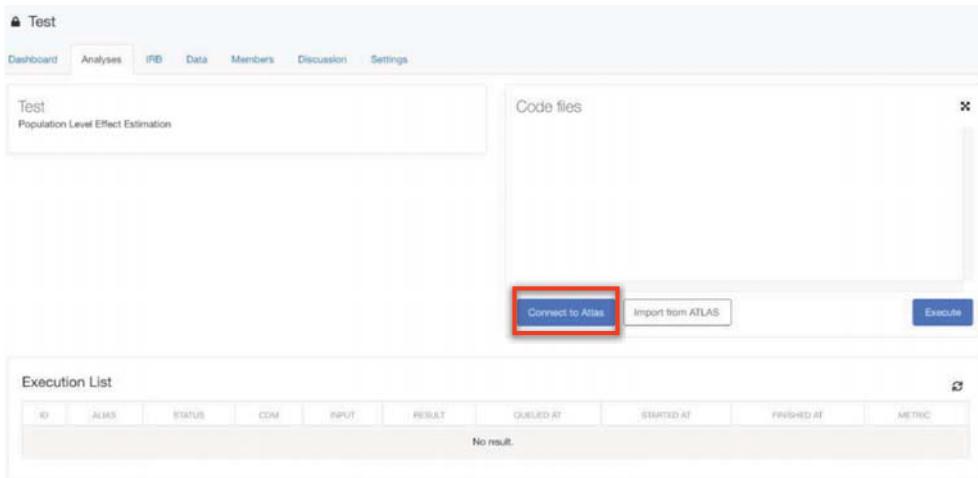
- 연구 분석 이름 및 연구 종류 (현재는 Estimation만 가능) 선택 후 저장 버튼 클릭

# 프로젝트

Analyses > Atlas 연결

- 소속 기관 ATLAS 연결 방법

Analyses 상세 화면



- 'Connect to ATLAS' 클릭

Analyses 속 ATLAS 연결 화면



- 'Connect' 클릭 시 ATLAS web page로 자동 이동

# 프로젝트

Analyses > 연구 코드 가져오기

Analyses 상세 화면

The screenshot shows the 'Analyses' section of a project. At the top, there are tabs for Dashboard, Analyses, IRB, Data, Members, Discussion, and Settings. The 'Analyses' tab is selected. Below the tabs, there's a 'Test' section with a sub-section 'Population Level Effect Estimation'. To the right, there's a 'Code files' panel with a 'Connect to ATLAS' button highlighted by a red box. Below the code files panel is an 'Execution List' table with columns: ID, ALIAS, STATUS, COM, INPUT, RESULT, QUEUED AT, STARTED AT, FINISHED AT, and METRIC. The table shows 'No result.'

- 'Import from ATLAS' 클릭시 연결된 ATLAS에서 연구 코드 불러오기 가능

Analyses 속 데이터 불러오는 화면

This is a modal dialog titled 'Pick research from Atlas'. It has a progress bar at the top with five steps: 1. Select Atlas, 2. Select Research, 3. Confirm, 4. Working, and 5. Complete. Below the progress bar is a table with four columns: ORGANIZATION NAME, API ALIAS, API VERSION, and CONNECT. The table contains one row with values: idc1, IDC 1, v2.5.0, and a 'Select' button highlighted by a red box. At the bottom right of the dialog is a 'Cancel' button.

- STEP 1
- 코드를 불러올 ATLAS 선택 (소속 기관 ATLAS만 접속 가능)

Analyses 속 데이터 불러오는 화면

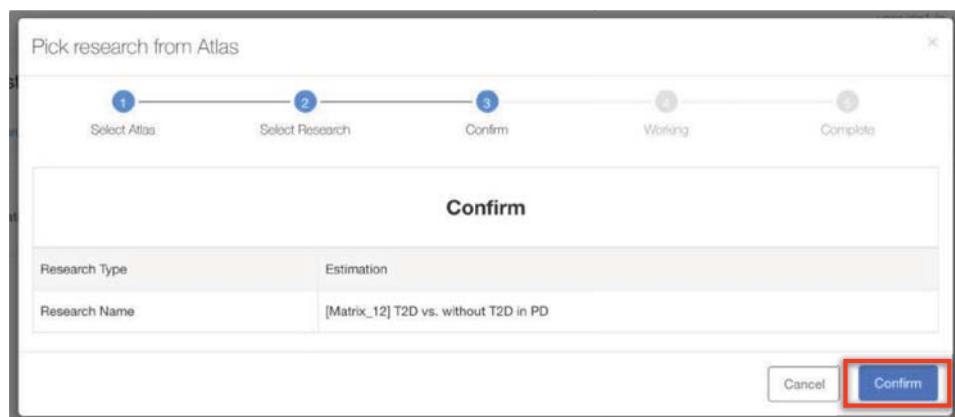
This is a modal dialog titled 'Pick research from Atlas'. It has a progress bar at the top with five steps: 1. Select Atlas, 2. Select Research, 3. Confirm, 4. Working, and 5. Complete. Below the progress bar is a 'Filter by Research Type:' dropdown set to 'Estimation'. A table below shows research entries. The second entry, 'Estimation' with ID 2 and name '[Matrix\_12] T2D vs. without T2D in PD', is highlighted by a red box. The table has columns: ID, ATLAS ESTIMATION NAME, and UPDATED AT. The 'UPDATED AT' column shows 'Apr 26, 2019 3:24 PM' for both rows. At the bottom right is a 'Cancel' button.

- STEP2
- 코드를 가져 올 연구 선택

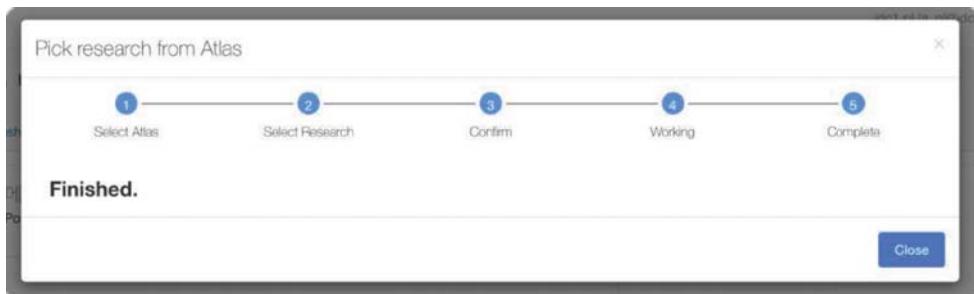
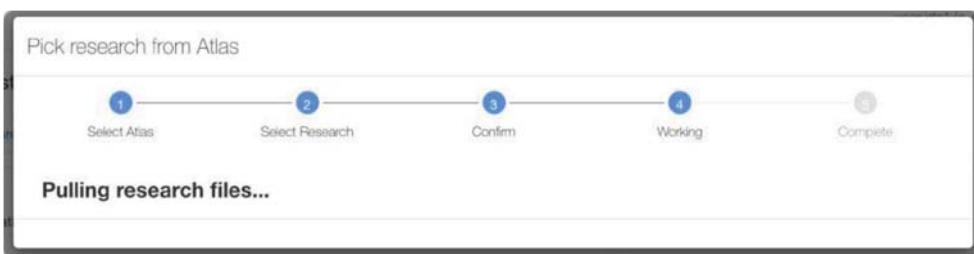
# 프로젝트

Analyses > 연구 코드 가져오기

Analyses 속 데이터 불러오는 화면



- STEP3
- **Confirm**을 눌러 원하는 연구 코드 가져오기



- STEP 4&5
- 코드 가져오기 및 완료 화면

# 프로젝트

Analyses > 연구 코드 가져오기

- Code files란에 불러온 code들 확인 가능

Analyses 상세 화면

The screenshot shows the 'Analyses' tab selected in the top navigation bar. Below it, there's a 'Test' project with a sub-project 'Population Level Effect Estimation'. On the right, the 'Code files' section is displayed, listing four files: 'comparator.sql', 'main.R', 'outcome.sql', and 'target.sql', each with its creation date and time. A red box highlights the close button for this section. At the bottom, there are three buttons: 'Connect to Atlas', 'Import from ATLAS', and 'Execute'.

Execution List

ID	ALIAS	STATUS	CDM	INPUT	RESULT	QUEUED AT	STARTED AT	FINISHED AT	METRIC
No result.									

- 확장 버튼 클릭시 상세 코드 확인 가능

연구 코드 상세 화면

The screenshot shows a modal window titled 'Files' with the 'comparator.sql' file selected. The main area displays the SQL code for creating temporary tables and inserting data from the 'CONCEPT' table into the 'Codesets' table. A red box highlights the expand/collapse icon next to the file name. At the bottom right of the modal is a 'Close' button.

```
1 CREATE TEMP TABLE Codesets :codeset_id int NOT NULL,
2   concept_id bigint NOT NULL
3 )
4 ;
5
6 INSERT INTO Codesets :codeset_id, concept_id)
7 SELECT @ as codeset_id, c.concept_id FROM (select distinct I.concept_id FROM
8 (
9   select concept_id from #vocabulary_database_schema.CONCEPT where concept_id in (201826)and invalid_reason is null
10
11 ) I
12 ) C;
13 INSERT INTO Codesets :codeset_id, concept_id)
14 SELECT I as codeset_id, c.concept_id FROM (select distinct I.concept_id FROM
15 (
16   select concept_id from #vocabulary_database_schema.CONCEPT where concept_id in (1503297,2160019,4333507,401919
17 UNION
18   select c.concept_id
19   from #vocabulary_database_schema.CONCEPT c
20   join #vocabulary_database_schema.CONCEPT_ANCESTOR ca on c.concept_id = ca.descendant_concept_id
21   and ca.ancestor_concept_id in (1503297,2160019,4333507,40191957,4338416,4333628,42704966)
22   and c.invalid_reason is null
23 )
24 ) I
25 INSERT INTO Codesets :codeset_id, concept_id)
26 SELECT 2 as codeset_id, c.concept_id FROM (select distinct I.concept_id FROM
27 )
```

- 파일을 클릭하여 원하는 코드를 확인

# 프로젝트

## Analyses > 연구 코드 실행

**Analyses 실행 화면**

The screenshot shows the 'Analyses' section of a project dashboard. At the top, there are tabs for 'Dashboard', 'Analyses' (which is selected), 'IRB', 'Data', 'Members', 'Discussion', and 'Settings'. Below the tabs, there's a 'Test' section titled 'Population Level Effect Estimation'. To the right, a 'Code files' panel lists several SQL and R scripts: 'comparator.sql', 'main.R', 'outcome.sql', and 'target.sql', each with a timestamp. At the bottom of the code panel are three buttons: 'Connect to Atlas', 'Import from ATLAS', and a red-bordered 'Execute' button.

- Import 된 Code를 실행하기 위하여 Execute 클릭

**Analyses code 실행 화면**

A modal dialog box titled 'Submit Research Execution'. It starts with a 'Select CDM to execute research' section containing a table:

ID	ORGANIZATION	ALIAS	CDM 버전	권한	SELECT
43	IDC 1	IDC #1 CDM	CDM v5	실행 가능	<b>선택</b>
42	IDC 2	IDC #2 CDM	CDM v5	데이터 관리자 승인 필요	<b>선택</b>
41	IDC 3	IDC #3 CDM	CDM v5	실행 불가	선택불가

Below the table, there are fields for 'Selected CDM' (with a note: '목록에서 연구를 실행할 데이터 셋을 선택해주세요.'), 'Select Research Type' (set to 'Estimation (ATLAS)'), 'Select R version' (set to 'v3.5.1'), and 'Alias' (set to '[Matrix\_12] T2D vs. without T2D in PD'). At the bottom are 'Close' and 'Submit' buttons.

- Code 실행을 위한 정보를 선택하고 Submit 클릭

- 기관 **CDM SELECT** : 기관의 정책에 따라 설정한 연구(코드 실행) 권한이 다를 수 있음
- Research type** : 현재 Estimation 만 지원
- R version** : 현재 v.3.5.1 만 지원
- Atlas** : 소속 기관 Atlas 에서 설계한 연구를 선택

# 프로젝트

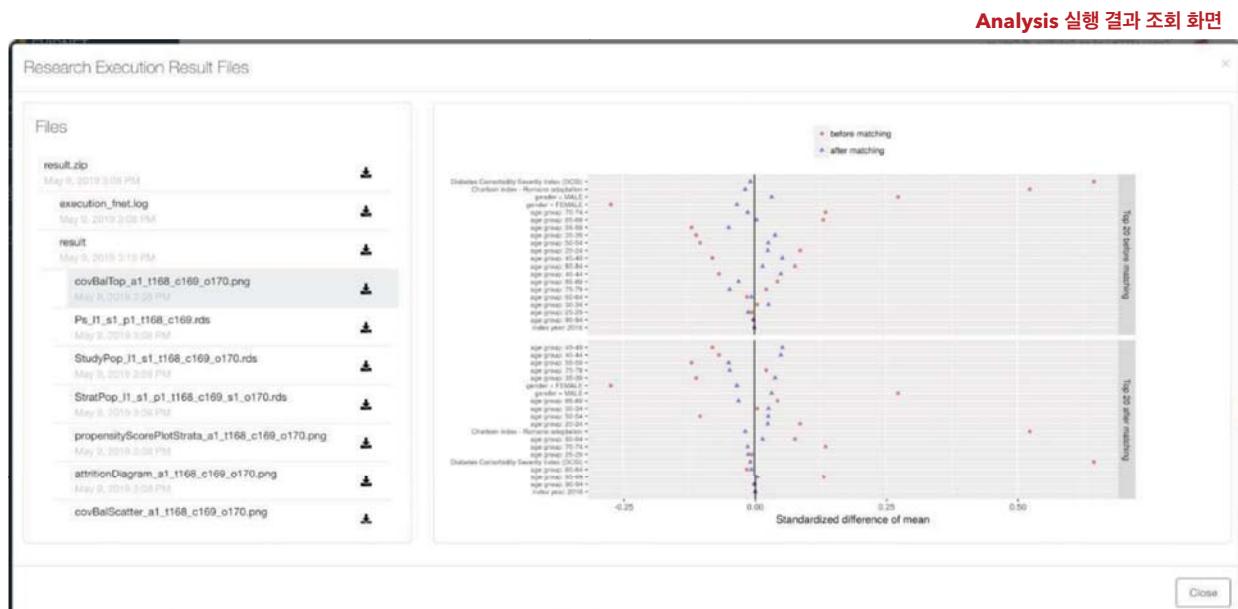
Analyses > 연구 코드 실행/결과 조회

Analysis 실행 화면									
Execution List									
ID	ALIAS	STATUS	CDM	INPUT	RESULT	QUEUED AT	STARTED AT	FINISHED AT	METRIC
323	execution 1	Success	EvidNet (CDM)	4 files	2 files	May 9, 2019 12:39 AM	May 9, 2019 12:39 AM	May 9, 2019 12:39 AM	Show metric

- Submit 완료 후 Execution List에 분석 현황 확인 가능

Execution List									
ID	ALIAS	STATUS	CDM	INPUT	RESULT	QUEUED AT	STARTED AT	FINISHED AT	METRIC
329	[Matrix_12] T2D vs. without T2D in PD	Success	idc2 (IDC2 DB)	4 files	1 files	May 9, 2019 3:06 PM	May 9, 2019 3:06 PM	May 9, 2019 3:08 PM	Show metric
328	[Matrix_12] T2D vs. without T2D in PD	승인 대기중	idc2 (IDC2 DB)	4 files	No result files.	May 9, 2019 2:59 PM			-

- 실행이 완료 되면 RESULT에 파일을 클릭하여 결과 조회 및 다운로드(기관 정책에 따라 데이터 담당자의 승인이 필요할 수 있음) 가능



# 프로젝트

IRB

IRB 화면

Test

Dashboard Analyses **IRB** Data Members Discussion Settings

IRB Status

List of IRB status of this project.

Search

Search

New IRB

기관명

IRB 상태

최신 IRB 내용

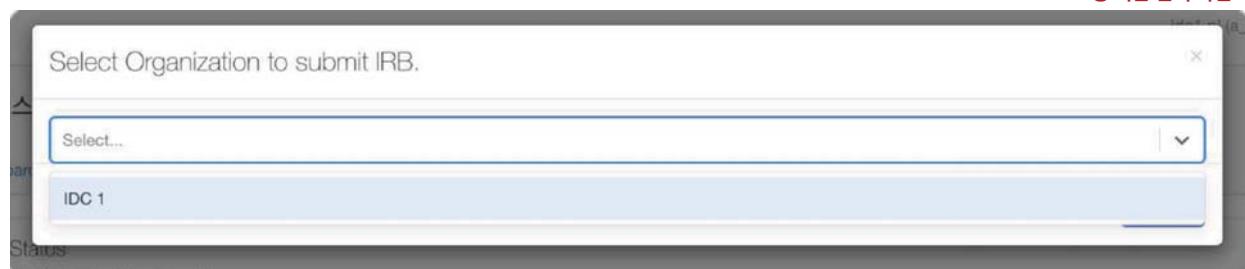
최신 IRB 상태 변경 시작

IRB 상태

No result.

- 해당 프로젝트 IRB에 대하여 기관 관리자에게 승인 유무 확인 절차
- 'New IRB' 클릭시 기관 관리자에게 IRB 승인 확인 요청 가능

IRB 요청 기관 선택 화면



- IRB 승인 확인 요청할 기관 선택

IRB 승인 확인 요청화면



- 기관 관리자에게 남길 해당 프로젝트 **IRB 승인 코드 또는 문서 파일 추가** 후 **Submit** 시 기관 관리자에게 해당 메모(요청) 전송

# 프로젝트

Data

Data 화면

A screenshot of a web application interface titled 'Test'. The top navigation bar includes links for Dashboard, Analyses, IRB, Data (which is highlighted in blue), Members, Discussion, and Settings. Below the navigation is a search bar with the placeholder 'Select...'. A dropdown menu is open, showing the option 'idc1 (권한 요청 필요)'.

- 확인 하고자하는 기관 선택

A screenshot of the same 'Test' application interface. The 'Data' tab is still selected. A message box displays the text '기관 데이터 담당자에게 권한을 요청합니다.' (Requesting access from the institution data manager).

- 코드 가져오기 및 완료 화면 (기관의 정책에 따라서 연구 권한이 달라질 수 있음 )

A screenshot of the 'Test' application interface under the 'Data' tab. The left sidebar shows 'IDC 1 - 접속 가능'. The main area is titled 'Data Sources reports' and lists 'CDM\_HIRA\_IDC1' and 'Dashboard'. The 'Dashboard' section is expanded, showing a 'Dashboard report (CDM\_HIRA\_IDC1)' with a 'CDM Summary' table and a 'Population by Gender' donut chart. The 'CDM Summary' table contains two rows: 'Source name' (cdmpv531) and 'Number of persons' (1.4680M). The 'Population by Gender' chart is divided into two equal halves, one blue (FEMALE) and one light blue (MALE).

- 기관 선택 후 연결된 ATLAS의 Achilles 화면 확인 가능

# 프로젝트

## Members

The screenshot shows a table of members for a project named '테스트프로젝트'. The columns are '이름' (Name), '프로젝트\_권한' (Project Permission), and '소속 기관' (Affiliated Organization). The rows list seven members: idc1 pl (관리자, IDC 1 - 데이터담당자), idc2 pl (관리자, IDC 2 - 데이터담당자), idc3 pl (관리자, IDC 3 - 데이터담당자), idc1 user (일반, IDC 1), idc2 user (일반, IDC 2), and idc3 user (일반, IDC 3). A red box highlights the 'Members' tab in the top navigation bar, and another red box highlights the '멤버 초대하기' (Invite Member) button in the top right corner.

이름	프로젝트_권한	소속 기관
idc1 pl	관리자	IDC 1 - 데이터담당자
idc2 pl	관리자	IDC 2 - 데이터담당자
idc3 pl	관리자	IDC 3 - 데이터담당자
idc1 user	일반	IDC 1
idc2 user	일반	IDC 2
idc3 user	일반	IDC 3

- '멤버 초대하기' 클릭시 해당 프로젝트에 원하는 기관의 연구자 또는 관리자를 추가할 수 있음

The image consists of two side-by-side screenshots. The left screenshot shows a search interface with a dropdown menu containing member names like 'user idc3 - c\_user@idc3.co.kr' and 'pl idc3 - c\_pl@idc3.co.kr'. The right screenshot shows a confirmation dialog titled '프로젝트에 초대' (Invite to Project) with a list of selected members and a '프로젝트 초대' (Invite to Project) button.

- 초대할 멤버 (이메일)검색 후 권한 (일반, 관리자) 선택 후 프로젝트에 초대

## Members (권한변경)화면

The screenshot shows a table of members for the '테스트프로젝트'. The columns are '이름' (Name), '프로젝트\_권한' (Project Permission), and '소속 기관' (Affiliated Organization). The rows list six members: idc1 pl (관리자, IDC 1 - 데이터담당자), idc2 pl (관리자, IDC 2 - 데이터담당자), idc3 pl (관리자, IDC 3 - 데이터담당자), idc1 user (일반, IDC 1), idc2 user (일반, IDC 2), and idc3 user (일반, IDC 3). In the '프로젝트\_권한' column for idc1 pl, the value '관리자' is highlighted with a red box, and the value '권한 변경 (일반)' is shown above it. Another red box highlights the '멤버 삭제' (Delete Member) link in the same row. The rest of the table and the top navigation bar are visible.

이름	프로젝트_권한	소속 기관
idc1 pl	관리자	IDC 1 - 데이터담당자
idc2 pl	관리자	IDC 2 - 데이터담당자
idc3 pl	관리자	IDC 3 - 데이터담당자
idc1 user	일반	IDC 1
idc2 user	일반	IDC 2
idc3 user	일반	IDC 3

- '프로젝트 권한'란에 권한 (관리자 또는 일반) 클릭시 해당 멤버에 대하여 권한 변경 및 해당 프로젝트에서 멤버 삭제 가능

# 프로젝트

## Permission

- 기관 담당자로부터 데이터 담당 권한을 위임받은 연구원만 해당 메뉴 진입 가능

Permissions 화면

ID	요청 연구원	기관	요청내용	처리
34	idc1 user	IDC 1	데이터 통계 조회 요청	승인됨 (승인자 : idc1 pl) May 7, 2019 11:54 PM
32	idc1 user	IDC 2	연구 결과 조회 요청 분석명 (Analysis Name) : a기관 Analyses 연구명 (Execution Name) : [282] [Matrix_12] T2D vs. without T2D in PD 분석 실행자 : idc1 user	승인됨 (승인자 : idc2 pl) May 3, 2019 9:44 AM
31	idc3 user	IDC 2	연구 실행 요청 분석명 (Analysis Name) : C 기관 연구원 연구명 (Execution Name) : [287] ACEi vs. ARB(angioedema) 분석 실행자 : idc3 user	승인됨 (승인자 : idc2 pl) May 3, 2019 9:42 AM
30	idc1 user	IDC 2	연구 실행 요청 분석명 (Analysis Name) : a기관 Analyses 연구명 (Execution Name) : [282] [Matrix_12] T2D vs. without T2D in PD 분석 실행자 : idc1 user	승인됨 (승인자 : idc2 pl) May 3, 2019 9:40 AM
29	idc1 pl	IDC 2	데이터 통계 조회 요청	승인됨 (승인자 : idc2 pl) May 3, 2019 9:36 AM
28	idc3 user	IDC 2	데이터 통계 조회 요청	승인됨 (승인자 : idc2 pl) May 3, 2019 9:36 AM

- 승인 요청 건 및 히스토리 확인 가능

Permissions 화면

ID	요청 연구원	기관	요청내용	처리
35	idc1 user	IDC 3	데이터 통계 조회 요청	승인됨 (승인자 : idc1 pl)
34	idc1 user	IDC 1	데이터 통계 조회 요청	승인됨 (승인자 : idc1 pl) May 7, 2019 11:54 PM
32	idc1 user	IDC 2	연구 결과 조회 요청 분석명 (Analysis Name) : a기관 Analyses 연구명 (Execution Name) : [282] [Matrix_12] T2D vs. without T2D in PD 분석 실행자 : idc1 user	승인됨 (승인자 : idc2 pl) May 3, 2019 9:44 AM
31	idc3 user	IDC 2	연구 실행 요청 분석명 (Analysis Name) : C 기관 연구원 연구명 (Execution Name) : [287] ACEi vs. ARB(angioedema) 분석 실행자 : idc3 user	승인됨 (승인자 : idc2 pl) May 3, 2019 9:42 AM

- 타기관, 소속기관 연구자들의 데이터 조회, 연구결과 조회 등 요청이 오는 것에 대하여 승인 / 거절 선택

## 프로젝트

## Discussion

- 프로젝트관련하여 참여자와 공유하거나 의견을 주고 받을때 사용가능한 보드

**Discussion 화면**

ARTICLE ID	TITLE	LABEL	DATE	PARTICIPANTS
24	의견 부탁드립니다.		2019-05-03T00:21:13.000Z	5

- '글 작성'을 누르면 새로운 글을 생성 가능

**Discussion 글 작성 화면**

New Discussion

Title

Font: Normal, 14pt, 16pt, 18pt, 20pt, 22pt, 24pt, 26pt, 28pt, 30pt, 32pt, 34pt, 36pt, 38pt, 40pt

Label

Code Bug  
Help  
Notice  
Want

Cancel Save

- 제목, 내용 기입 후 **SAVE** 클릭 시 해당 프로젝트 멤버들에게 작성한 게시글 공유

# Appendix

용어	설명
연구원	FeederNet®에서 연구활동을 하는 사용자
프로젝트 리더	프로젝트를 생성하고 프로젝트 멤버 초대 및 관리함
기관 관리자	기관을 관리 프로젝트별 IRB 승인상태를 확인함 기관의 정책에 따라 필요한 경우 프로젝트 별로 동일 기관 소속의 프로젝트 연구원에게 데이터 담당 권한을 부여함
연구 권한 설정	기관 관리자가 프로젝트 별로 설정함 권한은 소속 기관, 타기관 구분하여 연구 활동에 대한 접근을 제어(허용, 데이터 담당자 승인 필요, 불가)를 함
데이터 담당자	기관 관리자의 프로젝트 권한 설정에 따라 데이터 접근이 필요한 연구 활동에 대한 제어(승인, 반려)를 함
연구 활동	기관 데이터 통계 조회, 연구 코드 실행, 연구 결과 조회