– – OHDSI

Observational Health Data Sciences and Informatics (OHDSI)

2025-03-02

Contents

□□ **ix**
- □□□□□□ ix
- □□□□□ ix
- □□□ ix
- □□□□□□□□□□□□ x
- □□□□□ x
- □□□□□□□□□□ x

I OHDSI □□□□□□ **1**

1 –□□□□□– OHDSI □□□□□□ **3**
- 1.1 □□□□□□□□□□□□□ 3
- 1.2 □□□□□□□□□□□□□□□□□ 3
- 1.3 OHDSI□□□ 5
- 1.4 OHDSI□□□□□□ 6
- 1.5 □□□ 6

2 –□□□□□– □□□□□□□□□ **7**
- 2.1 □□□□□□□ 7
- 2.2 □□□□□□□□□□ 10
- 2.3 □□□ 11

3 –□□□□□– □□□□□□□□□ **13**
- 3.1 □□□□□□□□□ 13
- 3.2 □□□□□□□□□□□□: the Study-a-Thon 13
- 3.3 □□□□□□□□□□ 14
- 3.4 □□□□□□□ 14
- 3.5 □□□□□□□ 14
- 3.6 □□□□□□□□□□ 14
- 3.7 OHDSI□FAIR□□□□□□□□□□□□□ 14

□4□ –□□□□□– □□□□□□□□ **17**

4 □□□□□□□□ 19

4.1 □□□□ . . . 19
4.2 □□□□□□□□□ . . . 21
4.3 CDM□□□□□□□ . . . 23
4.4 □□□□ . . . 28
4.5 □□□ . . . 28
4.6 □□□□ . . . 28

5 □5□ –□□□□□– □□□□□□□□□ 31

5.1 □□□□□□□□□□□□□□□□□□□□□□ . . . 31
5.2 □□□□□ . . . 33
5.3 □□ . . . 37
5.4 □□ . . . 39
5.5 □□□□□□□□ . . . 40
5.6 □□□□□ . . . 40
5.7 □□□ . . . 41
5.8 □□ . . . 41

6 □6□ –□□□□□– □□□□□□□□□ 43

6.1 □□□□ . . . 43
6.2 □□□□2: □□□□□□□□□□□ . . . 49
6.3 □□□□ 3: ETL□□□ . . . 51
6.4 □□□□ 4: □□□□ . . . 51
6.5 ETL□□□□THEMIS . . . 51
6.6 CDM□□□ETL□□□□□□ . . . 52
6.7 ETL□□□□□□□□□□ . . . 52
6.8 □□□ . . . 52
6.9 □□ . . . 52

□7□ –□□□□□– □□□□□ 55

7 □□□□□□□□□ 57

7.1 □□□□ . . . 57
7.2 □□□□□□□□ . . . 58
7.3 □□□□□□□□ . . . 59
7.4 □□□□□□□□□□ . . . 59
7.5 □□□□□□□ . . . 60
7.6 □□□ . . . 60
7.7 □□ . . . 60

8 □8□ –□□□□□– OHDSI □□□□□ 63

8.1 □□□□□ . . . 63
8.2 □□□□ . . . 64
8.3 ATLAS . . . 64
8.4 □□□□□□□□□ . . . 66

8.5 □□□□ 68
8.6 □□□ 70

9 □9□ –□□□□□– SQL□R 71
9.1 SqlRender 72
9.2 DatabaseConnector 78
9.3 CDM□□□□□ 80
9.4 □□□□□□□□□□□□□□□□ 83
9.5 QueryLibrary 84
9.6 □□□□□□□□□□ 85
9.7 SQL□R□□□□□□□□□ 85
9.8 □□□ 90
9.9 □□□□ 90

10 □10□ –□□□□□– □□□□□□□ 93
10.1 □□□□□□□ 93
10.2 □□□□□□□□□□□□□ 94
10.3 □□□□□□□□ 95
10.4 □□□□□□□□□ 95
10.5 □□□□□□□□□□ 95
10.6 □□□□□□□□□□ 96
10.7 ATLAS□□□□□□□□□□□ 97
10.8 SQL□□□□□□□□□□□□ 101
10.9 □□□□ 109
10.10 □□ 109

11 □11□ –□□□□□– □□□□ 111
11.1 □□□□□□□□□□□□□□ 111
11.2 □□□□□□□□ 111
11.3 □□□□ 112
11.4 □□□ 112
11.5 □□□□□□□□□□ 113
11.6 ATLAS□□□□□□□□□□□□□□□ 113
11.7 ATLAS□□□□□□□□□□□□ 115
11.8 R□□□□□□□□□□□ 120
11.9 ATLAS□□□□□□□□□□□□□□□ 124
11.10 ATLAS □□□□□□□□□ 125
11.11 □□□ 131
11.12 □□ 131

12 □12□ –□□□□□– □□□□□□□□ 133
12.1 □□□□□□□□□□ 133
12.2 □□□□□□□□□□□□ 135
12.3 □□□□□□□□ 135
12.4 □□□□□□□□ 136

12.5 □□□□□□□□□□□□□□ 136
12.6 □□□□□□□□□□ 137
12.7 ATLAS□□□□□□□□□□ 139
12.8 R□□□□□□□□□□ 141
12.9 □□□ 147
12.10 □□ 147

13 □13□ –□□□□□– □□□□□□□ 151
13.1 □□□□ 151
13.2 □□□□□ 152
13.3 □□□□□□ 153
13.4 □□□□□□□□ 156
13.5 □□□□□□□□□□□□ 159
13.6 ATLAS□□□□□□□ 160
13.7 R□□□□□□ 164
13.8 □□□□□□□□□□□□ 176
13.9 □□□ 176
13.10 □□ 176

II –□□□□□– □□□□ 179

14 □□□□ 181
14.1 □□□□□□□□□□ 181
14.2 □□□□□□□ 182
14.3 □□□□□□□ 182

15 –□□□□□– □□□□□ 183
15.1 □□□□□□□□□□ 183
15.2 □□□□□□□□□ 187
15.3 □□□□□□ACHILLES 188
15.4 □□□□□□□□□□□□ 189
15.5 □□□ 191
15.6 □□ 191

16 –□□□□□– □□□□□□ 193
16.1 □□□□□□□□□□□ 193
16.2 □□ 199

17 –□□□□□– □□□□□□□□□□ 201
17.1 □□□□□□□□□ 201
17.2 □□□□□□□□□□□□□ 203
17.3 □□□ 204

18 –□□□□□– □□□□□□ 205
18.1 □□□□□□□□□ 205

18.2 □□□□□□□□ 205
18.3 □□□ 210

III OHDSI□□ **213**

19 –□□□□□– □□□□□ **215**

20 –□□□□□– OHDSI □□□□□□□□ **219**
20.1 OHDSI □□□□□□□□□□□ 219
20.2 OHDSI□□□□□□□□ 219
20.3 OHDSI□□□□□□□□□□□□□□□□□□ 222
20.4 □□□ 223

Appendix **223**

A –□□□□□– □□□ **225**

B –□□□□□– □□□□□□ **227**
B.1 ACE□□□ 227
B.2 ACE□□□□□□□□□□□□□ 228
B.3 □□□□□□□AMI□ 231
B.4 □□□□□ 232
B.5 □□□□□□□□□□□□□□□□□□□□□□□□ 233
B.6 □□□□□□□□□□□□□□□□□□□□ 235
B.7 □□□□□3□□□□□□□□□□□□□□□□□□□□□□□□□□ 238
B.8 ACE□□□□□□ 239
B.9 □□□□□□□□□□□□□□□ARB□□□□ 239
B.10 □□□□□□□□□□□□□□□□□□□□□□ 241
B.11 □□□□□□□□□□□□□□□□□□□□□□DCCB□□□□ 241
B.12 □□□□□□□□□□□□□□□□□□□□□□□NDCCB□□□□ 242
B.13 β□□□□□ 242
B.14 □□□□□□□□ 243
B.15 □□□□□□□□□□□□ 243
B.16 α1□□□□□ 243

C –□□□□□– □□□□ **245**
C.1 ACE□□□□□□□□□□□□□□□□□□□□□ 245

D –□□□□□– □□□□□□□□□□□ **249**

E –□□□□□– □□□ **251**
E.1 □□□□□□□□ 251
E.2 □□□□□□□□□ 254
E.3 □□□□□□□ 255
E.4 □□□□□□□□□□□□ 256

E.5 SQL□R 256
E.6 □□□□□□□□ 258
E.7 □□□□□ 262
E.8 □□□□□□□□ 264
E.9 □□□□□□□□ 272
E.10 □□□□□□ 274
E.11 275

Bibliography **277**

Index **287**

[illegible]OHDSI[illegible]OHDSI[illegible]OHD
http://book.ohdsi.org [illegible]
[illegible]

[illegible]OHDSI[illegible]OHDSI[illegible]OHDSI
[illegible]OHDSI[illegible]OHDSI[illegible]

[illegible]5[illegible]

I. OHDSI[illegible]
II. [illegible]
III. [illegible]
IV. [illegible]
V. OHDSI[illegible]

[illegible]

[illegible]1[illegible]

Hamed Abedtash	Mustafa Ascha	Mark Beno
Clair Blacketer	David Blatt	Brian Christian
Gino Cloft	Frank DeFalco	Sara Dempster
Jon Duke	Sergio Eslava	Clark Evans
Thomas Falconer	George Hripcsak	Vojtech Huser
Mark Khayter	Greg Klebanov	Kristin Kostka
Bob Lanese	Wanda Lattimore	Chun Li
David Madigan	Sindhoosha Malay	Harry Menegay
Akihiko Nishimura	Ellen Palmer	Nirav Patil
Jose Posada	Nicole Pratt	Dani Prieto-Alhambra
Christian Reich	Jenna Reps	Peter Rijnbeek
Patrick Ryan	Craig Sachson	Izzy Saridakis
Paola Saroufim	Martijn Schuemie	Sarah Seager
Anthony Sena	Sunah Song	Matt Spotnitz
Marc Suchard	Joel Swerdel	Devin Tian
Don Torok	Kees van Bochove	Mui Van Zandt
Erica Voss	Kristin Waite	Mike Warfe
Jamie Weaver	James Wiggins	Andrew Williams
Seng Chan You		

□□□□□□□□OHDSI□□□

- ACHILLES: □□□□□ 1.6.6
- ATLAS: □□□□□ 2.7.3
- EUNOMIA: □□□□□ 1.0.0
- □□□□□□□□□□□□: □ 1□□□

□□□□ RMarkdown □□□□□ bookdown □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
https://github.com/OHDSI/TheBookOfOhdsiInJapanese/ □□□□□□□□□□□□□□□□□□□□□
“travis” □□

Table 1: □□□□□□□□□□Methods Library□□□□□□□□□□□□

□□□□□	□□□□□
CaseControl	1.6.0
CaseCrossover	1.1.0
CohortMethod	3.1.0
Cyclops	2.0.2
DatabaseConnector	2.4.1
EmpiricalCalibration	2.0.0
EvidenceSynthesis	0.0.4
FeatureExtraction	2.2.4
MethodEvaluation	1.1.0
ParallelLogger	1.1.0
PatientLevelPrediction	3.0.6
SelfControlledCaseSeries	1.4.0
SelfControlledCohort	1.5.0
SqlRender	1.6.2

Part I

OHDSI

Chapter 1

– – OHDSI

&

1.1

10
(Olsen et al., 2007)
(Olsen et al., 2007)

1.2

Observational Medical Outcomes Partnership OMOP OMOP FDA
(Stang et al., 2010) OMOP

(Overhage et al., 2012) OMOP

OMOP
CDM

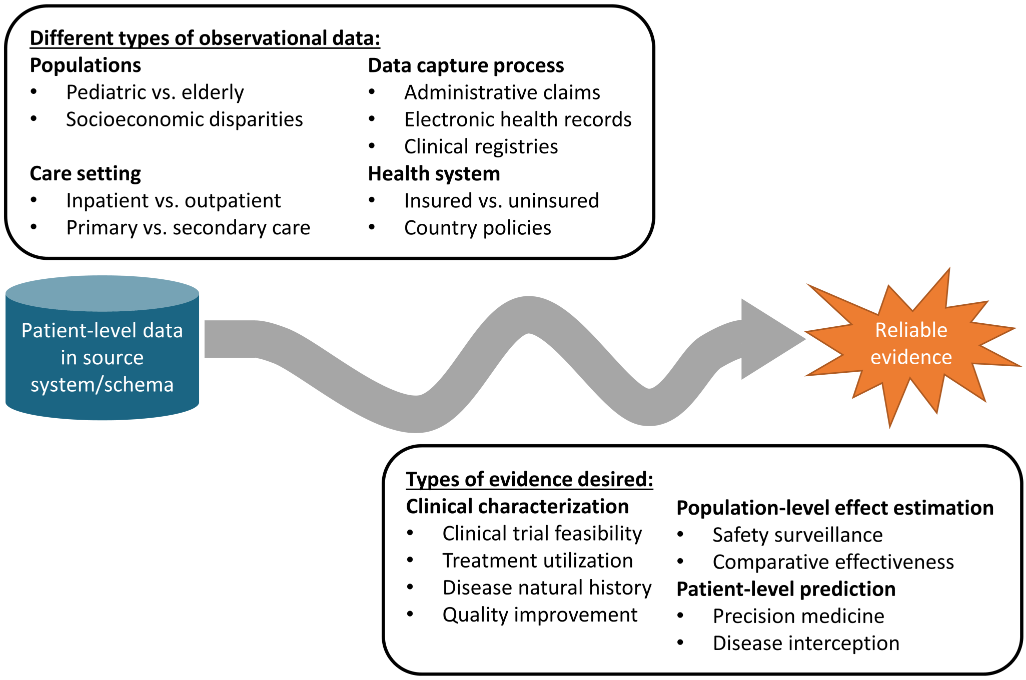


Figure 1.1: □□□□□□□□□□□□□

OMOP□□□□□□□□□□□□□(Ryan et al., 2012, 2013b) □□□□□□□(Schuemie et al., 2014)□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□(Madigan et al., 2013b,a)□□□□□□□□□OMOP□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□

OMOP□□□□□□□□□□FDA□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□-□□- □□
- □□
- □□

□□□□□□□□□□OHDSI□□□□□□□□

##□□□□□□□□□□□□□□□□□□OHDSI

Observational Health Data Sciences and Informatics□OHDSI□□□□□□□□□□□□□□□□□□□□□□ et al., 2015)□OHDSI□□□

1.2.1

□□

1.2.2

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.2.3

- □□□: □□
- □□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□
- □□□□□□□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□: □□
- □□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1.3 OHDSI

OHDSI□2014□□
OHDSI□□□□□□□□□1.2□□□□□□□□□□□□□□□□□□□□□□□□□□□□[1]

Figure 1.2: 2019□8□□□□OHDSI□□□□□□

2019□8□□□□OHDSI□20□□□□□□100□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CDM□□□OHDSI□□□

[1] https://www.ohdsi.org/who-we-are/collaborators/

OHDSI□OMOP CDM□ □2) □ □OHDSI□OMOP CDM□OHDSI□

OHDSI□ Academy of Science□5000□11□ □OHDSI□(Tian et al., 2018)□ et al., 2018a)□(Duke et al., 2017)□(Vashisht et al., 2018)□ et al., 2018b)□OHDSI□ et al., 2018)□(Johnston et al., 2019; Cepeda et al., 2018; Reps et al., 2019)

1.4 OHDSI

OHDSI□OHDSI□[2] □

1.5

- OHDSI□
- □
- OHDSI□

[2]https://github.com/OHDSI

Chapter 2

— —

&

-

OHDSI

2.1

OHDSI

2.1.1 OHDSI

OHDSI[1] OHDSI

OHDSI

- **:** OHDSI
- **:** OHDSI
- **:** OHDSI OMOP CDM
- **:** CDM OHDSI
- **CDM :** CDM
- **:**
- **:** OMOP OHDSI

to OHDSI! - Please introduce yourself OHDSI -
2)

[1] https://forums.ohdsi.org

□□□

2.1.2 OHDSI

OHDSI□□□

OHDSI□□
OHDSI □□

OHDSI□□
a-thon□□□OHD

OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□

2.1.3 OHDSI

OHDSI □□□□□□ □□□□□OHDSI □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

OHDSI□□

OHDSI □□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

2.1.4 OHDSI

OHDSI □□
□□
Wiki OHDSIWiki](https://www.ohdsi.org/web/wiki/doku.php?id=projects:overview)□□□□□□□□□□){.

□ 2.1.4 □□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□ □□□□□OHDSI□□□□□□

□□□□□□□□	□□	□□□□□
Atlas & WebAPI	Atlas□WebAPI□□OMOP□	□□□□□□□□□Atlas/WebAPI□□□□□□□ & JavaScript□□□□□□□□□□
CDM & Vocabulary	□□□	OMOP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OM
Genomics	OMOP CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ CDM □□□□□□□□□□□□□□□□□□	□□□□□□□□□□□
Population-Level Estimation	□□	

□□□□□□□□	□□	□□□□□
Natural Language Processing	OHDSI□□□□□□□□□□□□□□□□□(EHR)□□□□□□□□□□□□□□□□□	
□□□□□□□□	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	
Gold Standard Phenotype Library	OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	
FHIR Workgroup	OHDSI FHIR□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□EHR	□□□□□□□□□□□□□□□□□□□□
GIS	OMOP CD-M□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	□□□□□□□□□□□□□□□□□□□□
Clinical Trials	OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ OHDSI □□□□□□□□□□□□□□□□	
THEMIS	THEMIS □□□□□OMOP CDM □□□□□□□□□□□□□□□□ OMOP □□□□□□□□□ ETL □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	ETL□□□□□□□□□□□□□□□□□□□□□□
Metadata & Annotations	□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	
Patient Generated Health Data (PGHD)		
□□	□□□□□□□□□□	
□□□□□□□□□□□□□□□□□□□□□□□□/□□□/□□□□□□ □□□□□□□□□□□ PGHD □ ETL □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□		
Women of OHDSI	OHDSI	□□□□□□□□□□□□□□□□□□□□ (STEM) □□□□□□□□□□□□□□□□□□□□ □□□□□□□ STEM □□□□□□□□□□□□□□□□□□□□
Steering Committee	OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ OHDSI □□□□□□□□□□□□□□□□□□□□□□	

2.1.5 OHDSI

OHDSI □□□ □□□
OHDSI □□□OHDSI□□□[2]□□□[3]□□□[4]□□□
OHDSI □□□OHDSI Web□□□OHDSI□□□

2.1.6 OHDSI

OHDSI □□□OMOP□□□OHDS
CDM□□□

2.2

□□□ OHDSI □□□

□□□ □□□OHDSI□□□

□□**OHDSI**□□□ □□□
□□□
OHDSI □□□

□□□**OM**
CDM□**OHDSI**□□□ □□□/□□□OHDSI□□□
CDM□□□OHDSI□□□
7□7□□□OMOP CDM□OHDSI□□□

□□□**ETL**□□□**OMOP CDM**□□□
□□□OMOP□□□ETL□□□OHDSI□□
CDM□□□

□□□**OHDSI**□□□R□□□
Render□□□OHDSI/SqlRender□GitHub□□□

□□□**OHDSI**□□□
□□□OHDSI□□□Apache□□□

□□□**OHDSI**□□□ □□□
OHDSI □□□

□□□**OHDSI**□□□ □□□OHDSI□□□

[2]https://www.ohdsi-europe.org/
[3]https://forums.ohdsi.org/c/For-collaborators-wishing-to-communicate-in-Korean
[4]https://ohdsichina.org/

2.3

- OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□**OHDSI**□□□□□□□
- □□□ETL□□□□□□□OHDSI□□□□□□□□□□□□□

Chapter 3

– –

□□□□□*:* □□□□□□□□□□□

OHDSI□□□ □□O □□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□

3.1

□□□□□□□□□□□□□□□□□1990□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□20 (Wikipedia, 2019a) □□□

□□□□□□□□□□□□□□□□□□2.0□□□□□□□(Wikipedia, 2019b)□□□□□□□□□□□□□□ hacking□□□ 2019) □□ et al., 2016) □□□ et al., 2018)

OHDSI□□□ □□□ □□□

3.2 : the Study-a-Thon

□□□□□□□□□□□□□□□□□□□□study-a-thon□□□□□□□□□□□□□□□□OMOP□□□□□ a-thon□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ a-thon□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□study-a-thon□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 8□□□□ATLAS□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

3.3

OHDSI□OMOP□
4□5□
7□
(Garza et al., 2016)

3.4

OHDSI□
Methods Library□ATLAS□Athena□OHDSI□
□20□GitHub□
Library□

3.5

□
//howoften.org□http://data.ohdsi.org□
□20□OMOP□
CDM□OHDSI ETL□

3.6

□
wiki□GitHub □
□OHDSI □

3.7 OHDSI FAIR

3.7.1

□Wilkinson et al. (2016□) □FAIR□OHDSI□

3.7.2

OMOP□
2019□IMI EHDEN□

3.7.3

OMOP□OMOP CDM□SQL□OMOP□
EHDEN□LEGEND□http://how

3.7.4

□OMOP□OHDSI□
FHIR□HL7 CIMI□openEHR□OHDSI□
Athena□OMOP□

3.7.5

□FAIR□
□
ETL□
CDM□

- OHDSI□
- □

4 – –

Chapter 4

: *Clair Blacketer*

(ii)
- (iii)

CDM 5 CDM

CDM 4.1

4.1

CDM

-
-
- - 13
- - 12

CDM

- : CDM
- :
- :

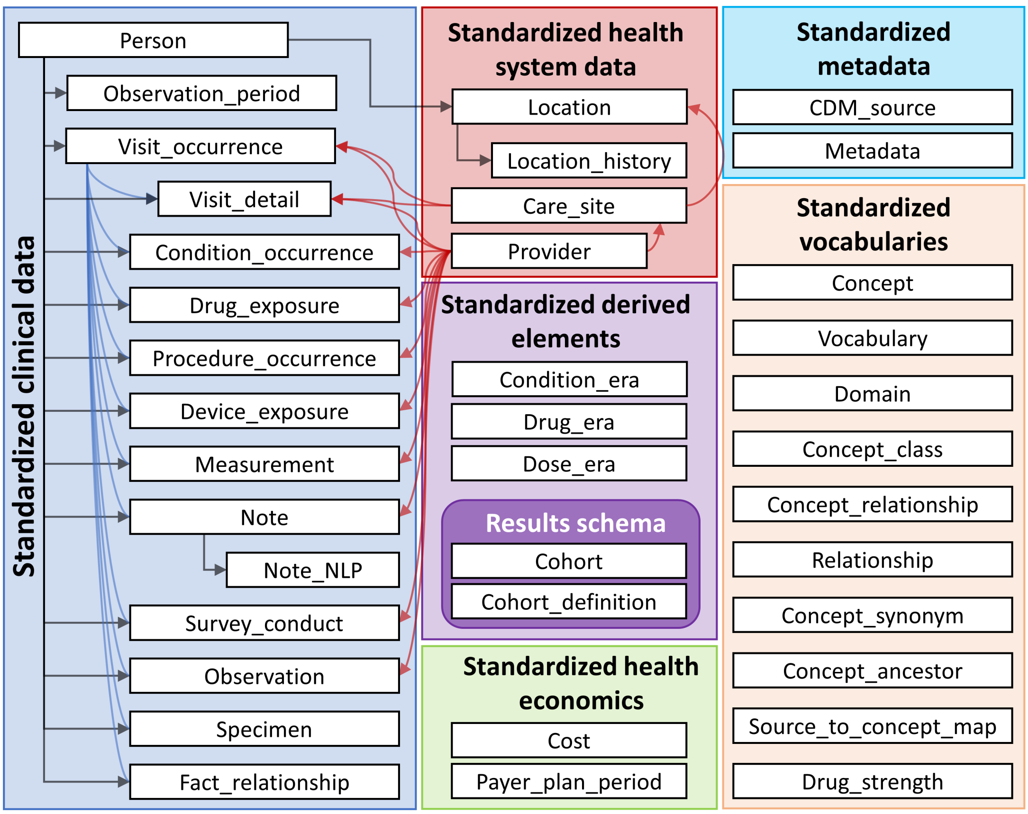


Figure 4.1: CDM□□□□□6.0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- [illegible]: [illegible]- [illegible]-[illegible]
- [illegible]: [illegible]CDM[illegible]
- [illegible]: [illegible]
- [illegible]: [illegible]
- [illegible]: CDM[illegible]Oracle[illegible]SQL Server[illegible]
- [illegible]: CDM[illegible]
- [illegible]: [illegible]CDM[illegible]github[illegible](https://github.com/OHDSI/Comm

4.2

CDM[illegible]CDM[illegible]

4.2.1

CDM[illegible]PERSON[illegible]

4.2.2

[illegible] 10[illegible]COHORT[illegible]

4.2.3

CDM[illegible]ANSI SQL[illegible]VARCHAR[illegible]INTEGER[illegible]FL

[illegible]: [illegible] 8[illegible]

4.2.4

[illegible] 5.2.3 [illegible]

[illegible]: [illegible]

[illegible]	[illegible]ID	[illegible]	[illegible]ID
1731378	Drug	183	Route
477597	Device	180	Currency
257000	Procedure	158	Payer
163807	Condition	123	Visit

□□□□□□□	□□□□ID	□□□□□□□	□□□□ID
145898	Observation	51	Cost
89645	Measurement	50	Race
33759	Spec Anatomic Site	13	Plan Stop Reason
17302	Meas Value	11	Plan
1799	Specimen	6	Episode
1215	Provider Specialty	6	Sponsor
1046	Unit	5	Meas Value Operator
944	Metadata	3	Spec Disease Status
538	Revenue Code	2	Gender
336	Type Concept	2	Ethnicity
194	Relationship	1	Observation Type

4.2.5

CDM□□

CONCEPT□□
5□□□□

4.2.6

□□□□□□□□□□□□□□□□□1□□□□□□□□□□

□: □□□□□□□□□□

□□	□□
[Event]_ID	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
[Event]_CONCEPT_ID	CONCEPT□□□□□□□□□□□□□□□□□□□□□□□□□ = 31967□□□□□□□□□SNOMED□□□□□□□□□□□□□□
[Event]_SOURCE_CONCEPT_ID	CONCEPT□□□□□□□□□□□□□□□□□□□□□□□□□ = 45431665□□Read□□□□□□□□□□□□□□□□□□□□□C CT□□□□□31967□□□□□□□□□□□□□□□□□□□□
[Event]_TYPE_CONCEPT_ID	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
[Event]_SOURCE_VALUE	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 9□□□787.02□□□□□□□□□□□□□□□

4.2.7

□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ICD9CM□NDC□Read□□□□
- □□□□□□□CDM□□

- □□□
- □□□

□□□□□□□□□□□□□□QA□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 4.2 □□□□□□□□□□□□□□□□□□□□□□□□TB□□ICD9CM□□□□011□□□

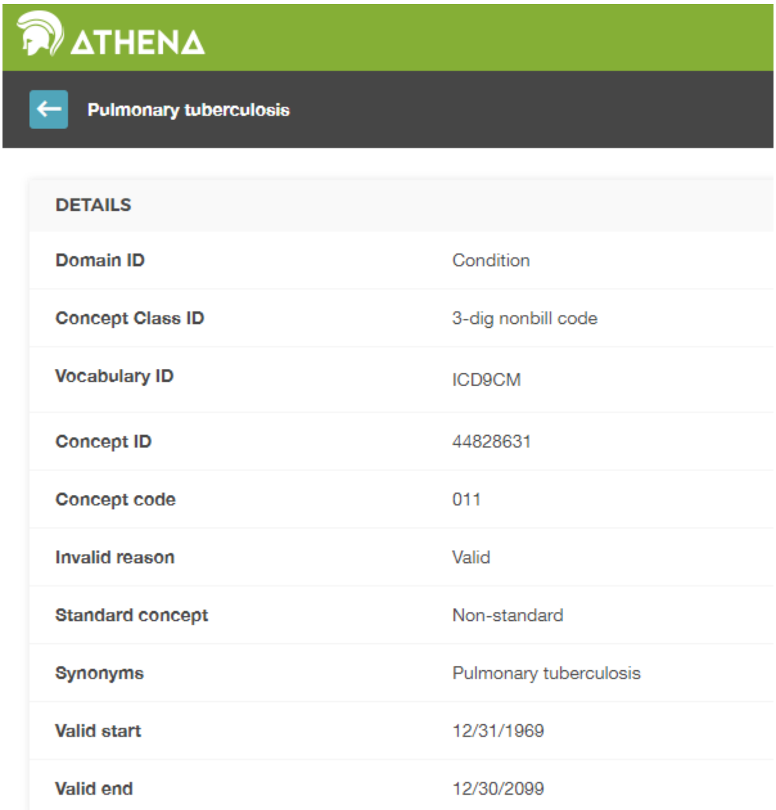


Figure 4.2: □□□□ICD9CM□□□

□□□□□□□□011□□□□□UB04□□□□□□□□□□□□□□□□□□A□□□□□□□□□DRG□□ TB□□□□□□□□□□Figure 4.3 □□□□□□□□□□□□□□□□□□□□□□□OMOP□□□□□□□ 253954□□□□□□□□□□□□□□□□□□□□□Read□ICD10□CIEL□MeSH□□□□□□□□□□

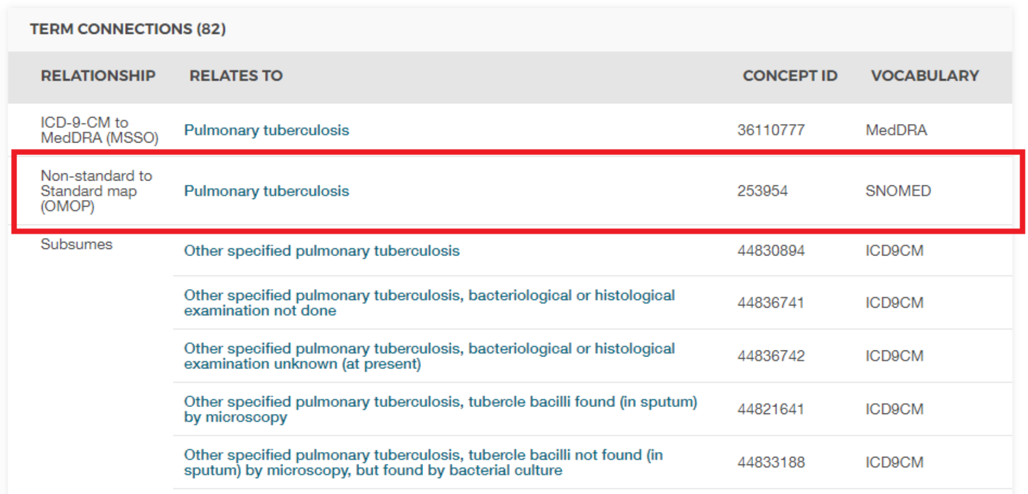
□□□□□□□□□□□□□□□□□□□□□□□□ 4.7□□□□

4.3 CDM

CDM□□16□□□□□□□□□□□□10□□□□□□□□□□□□2□□□□□□□□□□□□4□□□□□□ □□□□□□□□□CDM Wiki□□□□□□□□□□□□□[1]□

□□□1□□□□□□□□□

[1] https://github.com/OHDSI/CommonDataModel/wiki



TERM CONNECTIONS (82)

RELATIONSHIP	RELATES TO	CONCEPT ID	VOCABULARY
ICD-9-CM to MedDRA (MSSO)	Pulmonary tuberculosis	36110777	MedDRA
Non-standard to Standard map (OMOP)	Pulmonary tuberculosis	253954	SNOMED
Subsumes	Other specified pulmonary tuberculosis	44830894	ICD9CM
	Other specified pulmonary tuberculosis, bacteriological or histological examination not done	44836741	ICD9CM
	Other specified pulmonary tuberculosis, bacteriological or histological examination unknown (at present)	44836742	ICD9CM
	Other specified pulmonary tuberculosis, tubercle bacilli found (in sputum) by microscopy	44821641	ICD9CM
	Other specified pulmonary tuberculosis, tubercle bacilli not found (in sputum) by microscopy, but found by bacterial culture	44833188	ICD9CM

Figure 4.3: □□□□SNOMED□□□

4.3.1 :

□□

□□

Lauren□□□
//endometriosis-uk.org/laurens-story □□□□□□□□□□□

4.3.2 PERSON

Lauren□□□□□□□□□□□□

- □□□36□□□□□□
- □□□□□□□1982□3□12□□□
- □□□□□□□
- □□□□□□□□□□

□□□□□□□□□□□□PERSON□□□□□□□□□□□□□□□

Table 4.3: PERSON□□□□□

□□	□	□□
PERSON_ID	1	PERSON_ID□□□□□□□□□□□□□□□□□□□□□
GENDER_CONCEPT_ID	8532	□□□□□□□□□□□□□□ID□8532□□□
YEAR_OF_BIRTH	1982	
MONTH_OF_BIRTH	3	
DAY_OF_BIRTH	12	
BIRTH_DATETIME	1982-03-12 00:00:00	□□□□□□□□□□□□□□□□□□□□
DEATH_DATETIME		
RACE_CONCEPT_ID	8527	□□□□□□□□□□□□ID□8527□□□□□□□□□□□□
ETHNICITY_CONCEPT_ID	38003564	□□□□□□□□□□□□□□□□□□□□□□US□□□□
LOCATION_ID		□□□□□□□□□□□
PROVIDER_ID		□□□□□□□□□□□□□□□□□□□
CARE_SITE		□□□□□□□□□□□□□□□
PERSON_SOURCE_VALUE	1	□□□□□□□□□□□□□□□□□□□□□□□□□□□
GENDER_SOURCE_VALUE	F	□□□□□□□□□□□□□□□□□□□□□□□□□□□
GENDER_SOURCE_CONCEPT_ID	0	□□□□□□□□OHDSI□□□□□□□□□□□□□□□□ F□□□□□□□PCORNet□□□□□□□□□□□□□□□
RACE_SOURCE_VALUE	white	□□□□□□□□□□□□□□□□□□□□□□□□□□
RACE_SOURCE_CONCEPT_ID	0	□□□GENDER_SOURCE_CONCEPT_ID□□□□□□
ETHNICITY_SOURCE_VALUE	english	□□□□□□□□□□□□□□□□□□□□□□□□□□□
ETHNICITY_SOURCE_CONCEPT_ID	0	□□□GENDER_SOURCE_CONCEPT_ID□□□□□□

4.3.3 OBSERVATION_PERIOD

OBSERVATION_PERIOD□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Lauren□□□□□□□□□□□□□□□□□□□□□□□

Lauren□□□□Table 4.4□□□□□□□□□□EHR□□□□□□□□□□□□□□□□□□□□□□□□□□

Table 4.4: Lauren□□□□□□□□□□□□□□

□□□□□□□ID	□□□	□□□	□□□
70	2010-01-06	2010-01-06	□□□□

□□□□□□□ID	□□□	□□□	□□□
80	2011-01-06	2011-01-06	□□□□
90	2012-01-06	2012-01-06	□□□□
100	2013-01-07	2013-01-07	□□□□
101	2013-01-14	2013-01-14	□□□□
102	2013-01-17	2013-01-24	□□□□

□□□□□□□□□□□□□□□□□□□OBSERVATION_PERIOD□□□□□□□□□□□□□□□□□□□□

Table 4.5: OBSERVATION_PERIOD□□□□□

□□	□	□□
OBSERVATION_PERIOD_ID	1	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
PERSON_ID	1	□□□PERSON□□□□□Laura□□□□□□□□□□□□□□□□
OBSERVATION_PERIOD_START_DATE	2010-01-06	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
OBSERVATION_PERIOD_END_DATE	2013-01-24	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
PERIOD_TYPE_CONCEPT_ID	44814725	□Obs Period Type□□□□□□□□□□□□□□□□□□□□44814724□□□□□

4.3.4 VISIT_OCCURRENCE

VISIT_OCCURRENCE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□

Lauren□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□VISIT_OCCURRENCE□□□□□□□□□□□□□

Table 4.6: VISIT_OCCURRENCE□□□□□

□□	□	□□
VISIT_OCCURRENCE_ID	514	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
PERSON_ID	1	□□□PERSON□□□□□Laura□□□□□□□□□□□□□□□
VISIT_CONCEPT_ID	9201	□□□□□□□□□□□□□□9201□□□
VISIT_START_DATE	2013-01-17	□□□□□□□□□
VISIT_START_DATETIME	2013-01-17 00:00:00	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
VISIT_END_DATE	2013-01-24	□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□
VISIT_END_DATETIME	2013-01-24 00:00:00	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□	□	□□
VISIT_TYPE_ CONCEPT_ID	32034	□□□□□□□□□□□□□□□□□□□□□□□□□EHR□
PROVIDER_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CARE_SITE_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
VISIT_SOURCE_ VALUE	□□	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
VISIT_SOURCE_ CONCEPT_ID	NULL	□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□
ADMITTED_FROM_ CONCEPT_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 8536□□□□□□□□□□□
ADMITTED_FROM_ SOURCE_CONCEPT_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
DISCHARGE_TO_ CONCEPT_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 8615□□□□□□□□□□□□□□□
DISCHARGE_TO_ SOURCE_VALUE	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□ living facility□□□□
PRECEDING_VIS		

IT_OCCURRENCE_ID|NULL|□□□□□□□□□□□□□□□□□ADMITTED_FROM_CONCEPT

- □□□1□□ wiki□□□□□□□□□□□□

4.3.5 CONDITION_OCCURRENCE

CONDITION_OCCURRENCE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Lauren□□□□□□□□□□

□□□□□□□□□□□□□□□□□

> □3□□ 2□□□

□□□□□□□□□□□□SNOMED□□□□266599000□□□□□□□□CONDITION_OCCURREN

Table 4.7: CONDITION_OCCURRENCE□□□□□

□□	□	□□
CONDITION_ OCCURRENCE_ID	964	□□□□□□□□□□□□□□□□□□□□□□□□□□□□
PERSON_ID	1	□□□PERSON□□□□□Laura□□□□□□□□□□□□
CONDITION_ CONCEPT_ID	194696	SNOMED□□□266599000□□□□□□194696□□□

□□	□	□□
CONDITION_START_DATE	2010-01-06	□□□□□□□□□□□□□□□□
CONDITION_START_DATETIME	2010-01-06 00:00:00	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CONDITION_END_DATE	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CONDITION_END_DATETIME	NULL	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CONDITION_TYPE_CONCEPT_ID	32020	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□EHR□□□□□□□□□□□□□□□□□□□□□□ID 32020 □□□□□□□□□□□□□| |CONDITION_STATUS_CONCEPT_ID|NULL|□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 4203942 □□□□□□□□□□□□□□| |STOP_REASON|NULL|□□□□□□□□□□□□□□□□□□

4.4

□□□□□□□□□CDM□□□[2]□□□□□□□□□□

4.5

- CDM□□
- CDM□□□□□□□□□□□□□□□□
- CDM□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□

4.6

□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATHENA[3]□□□□ATLA

□□ **4.1.** □□□□□1974□8□4□□PERSON□□□

[2] https://github.com/OHDSI/CommonDataModel/wiki
[3] http://athena.ohdsi.org/
[4] http://atlas-demo.ohdsi.org

□□ **4.2.** □□□□2015□1□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□2019□7□1□□□□□□□□

□□ **4.3.** □□□□2019□5□1□□□□□□□□□□□□200 MG□□□□□NDC□□□□76168009520□□30□

□□□□

□□□3□□□□□□□□□□□□□@ref□installR□□□□□□□□□□□□□R□R-Studio□□□□Java□□□

```
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
```

Eunomia□□□□□□□□□□□□□R□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□

```
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM□□□□□□□□□□□□□main□□□□□□□□CONDITION_OCCURRENCE□□□□□□□□□□

```
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT *
FROM @cdm.condition_occurrence
LIMIT 1;"
result <- renderTranslateQuerySql(connection, sql, cdm = "main")
```

□□ **4.4.** SQL□R□□□□□□□□□□□□□□□□□□□□□□□□ID192671□□□□□□□□□□□□□□□□□□□□□

□□ **4.5.** SQL□R□□
10□□□□□□□□□□□□□□ICD-10□□□□□□K92.2□□□□

□□ **4.6.** SQL□R□□□□□□□PERSON_ID 61□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□@ref□Cdmanswers□□□□□□□□

Chapter 5

5 – –

: Christian Reich & Anna Ostropolets

OMOP OHDSI

5.1

of Mortality 5.1

OHDSI OMOP CDM CDM

-
-

OMOP CDM

5.1.1

CDM OHDSI
GitHub [3]

[1] https://github.com/OHDSI/Vocabulary-v5.0
[2] https://forums.ohdsi.org
[3] https://github.com/OHDSI/CommonDataModel/issues

1660.

A General BILL for this preſent Year,

Ending the 11th Day of December 1660.

According to the Report made to the King's moſt excellent Majeſty,

By the Company of Pariſh Clerks of LONDON, &c.

DISEASES and CASUALTIES.

Abortive and Stillborn	421	Flox and Small Pox	1523	Palſy	17
Aged	909	Found dead in the Streets, Fields, &c.	2	Plague	36
Ague and Fever	2303	French Pox	51	Pluriſy	12
Apoplexy and Suddenly	91	Gout	4	Quinſy and ſore Throat	21
Blaſted and Planet	3	Grief	13	Rickets	441
Bleeding and bloody Iſſue	7	Griping in the Guts	253	Riſing of the Lights	210
Bloody Flux, Scowring, and Flux	346	Hanged and made away themſelves	11	Rupture	12
Burnt and Scalded	6	Head-ach and Headmouldſhot	35	Scurvy	82
Cancer, Gangrene and Fiſtula	63	Jaundies	102	Shot	7
Canker, ſore Mouth and Thruſh	73	Impoſthume	105	Shingles	1
Childbed	226	Killed by ſeveral Accidents	55	Sores, Ulcers, broken and bruiſed Limbs	61
Chriſomes and Infants	858	King's Evil	28	Spleen	7
Cold, Cough and Hiccough	33	Lethargy	6	Spotted Fever and Purples	368
Colick and Wind	116	Livergrown	8	Starved	7
Conſumption and Tiſſick	2982	Lunatick and Frenzy	14	Strangury	22
Convulſion	742	Megrims	5	Stopping of the Stomach	186
Cut of the Stone and Stone	46	Meaſles	6	Surfeit	202
Dropſy and Tympany	646	Mother	1	Swine Pox	2
Drowned	57	Murthered	7	Teeth and Worms	839
Executed	7	Overlaid and Starved at Nurſe	46	Vomiting	8
Falling Sickneſs	4			Wen	1

Figure 5.1: 1660□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□62□□□□□□□□□□□□□□□□□□□□□□□□

5.1.2

Pallas ATHEN

OMOP CDM zip

5.1.3 :

OHDSI i
Extension 5.6.9

5.2

OMOP CDM

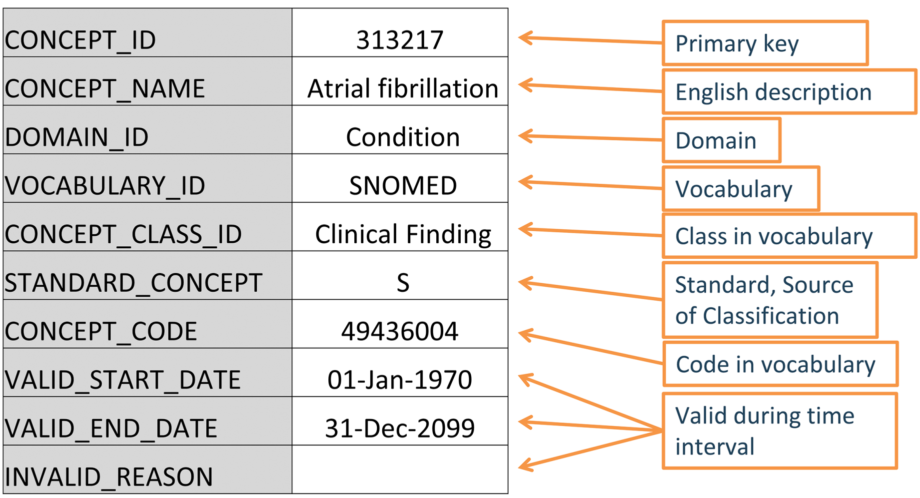


Figure 5.2: OMOP CDM

5.2.1 ID

ID ID

5.2.2

1

[4]http://athena.ohdsi.org

5.2.3

□□□□□□□□□DOMAIN_ID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CONCEPT_ID□□□
□□□□□□□□□□□□□Pallas□□OM

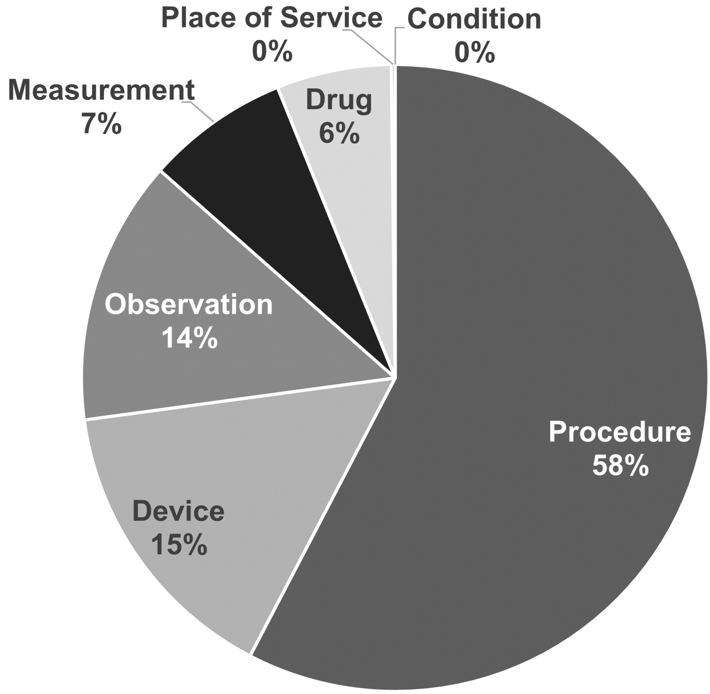


Figure 5.3: □□□□□□□□□□□CPT4□□□HCPCS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□CDM□□□□□□□□□□□□□□□□□□□□□□
4□□□5.6□□□□□□□□□□□□□□□□

5.2.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
9-CM□□ICD9CM□□□□□□□□□□□ID□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□111□□□□□

5.2.5

□□□ID□□□□□
Class ID□□□□□□□□□□□□

□: □□

□□□□□□□□□□□□□	□□□□□□□□□□□□□□□□□
□□	□□□□□□□□□□□□□□□□□□□ATC□CDT□Episode□HCPCS□HemOnc□ICDs□MedDR
□□	CIEL□HES□□□ICDO3□MeSH□NAACCR□NDFRT□OPCS4□PCORNET□Plan

[illegible]	[illegible]
[illegible]	CPT4 [illegible] ISBT [illegible] LOINC
[illegible]	APC [illegible] OXMIS [illegible]

[illegible] RxNo

5.2.6

[illegible] 1 [illegible]

5.2.7

[illegible]

5.2.8

[illegible]
- [illegible] 5.4 [illegible]

[illegible]

[illegible]: [illegible]/[illegible]/[illegible]

[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	SNOMED, ICDO3	SNOMED Veterinary	MedDRA
[illegible]	SNOMED, CPT4, HCPCS, ICD10PCS, ICD9Proc, OPCS4	SNOMED Veterinary, HemOnc, NAACCR	[illegible]
[illegible]	SNOMED, LOINC	SNOMED Veterinary, NAACCR, CPT4, HCPCS, OPCS4, PPI	[illegible]
[illegible]	RxNorm, RxNorm Extension, CVX	HCPCS, CPT4, HemOnc, NAACCR	ATC
[illegible]	SNOMED	[illegible]	[illegible]
[illegible]	SNOMED	[illegible]	[illegible]
[illegible]	CMS Place of Service, ABMT, NUCC	SNOMED, HCPCS, CPT4, UB04	[illegible]

5.2.9

[illegible] ICD9CM [illegible] NDC [illegible]

[illegible]: [illegible] 1001 [illegible]

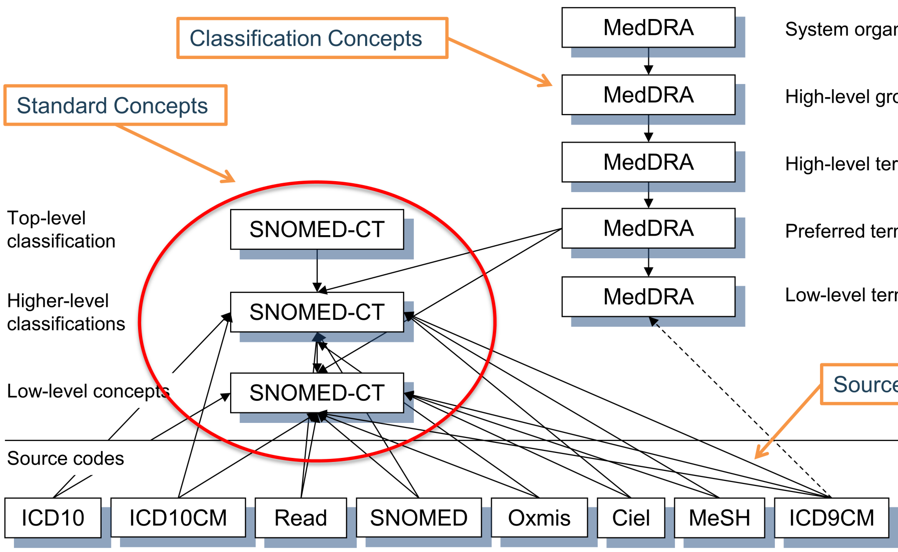


Figure 5.4: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SNOMED□□□□□□

□□□□□ID	□□□□□□□□□□□		□□□□ID	□□□□□□ID	□□□□□□□□
35803438	1001	□□□□□□□□□□□□		HemOnc	□□□□□□□□□□□
35942070	1001	AJCC TNM Clin T	□□	NAACCR	NAACCR□□
1036059	1001	□□□□□□	□□	RxNorm	□□
38003544	1001	□□□□□□□□□ - □□□	□□□□□	□□□□□	□□□□□
43228317	1001	□□□□□□□□□□□□□□		BDPM	□□
45417187	1001	□□□□□□□□□□□□□□□□□10 mg/ml□□□□□		Multum	Multum
45912144	1001	□□	□□	CIEL	□□

5.2.10

□□ CDM□□ - □□□□□□□□□□□□□□□□ - □□: □□□□□□□□□□ - VALID_START_DATE: □□□1970-1-1□ - VALID_END_DATE: □□ 12-31□□□□□□□□□□ - INVALID_REASON: NULL - □□□□□□□□□□□□□□□□ - □□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5.2.6□□□□ - VALID_START_DATE: □□ 1-1□ - VALID_END_DATE: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ - INVALID_REASON: "D" - □□□□□□□□□□□□□□□□□□□□□ - □□: □□

- VALID_START_DATE: □□□1
1-1□ - VALID_END_DATE: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- INVALID_REASON: “U” - □□□□□□□□□□□□□□□□□□□□□□□□□□□ - □□: □□□ - VALID_START_DATE: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

5.3

□□□2□□ to□□□□□□□□□□□□□□Mapped from□□□□□□□□□□

CONCEPT_RELATIONSHIP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□RELATIONSHIP_STA

5.3.1

□□□2□□□□ID□□□□□□□□□□□□□

Table 5.4: □□□□□□□□□□□□□□

□□ID□□	□□
“Maps to”□□□”Mapped from”	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
“Maps to value”□□□”Value mapped from”	MEASUREMENT□□□OBSERVATION□□□□□VALUE_AS_CONCEPT_I

□□□OMOP CDM□□□

□□ W61.51□□SNOMED□□□□□□□□□□□□ 217716004□□

□□□ICD9CM 070.43□□□□□□□□□□E□□□□□□□SNOMED 235867002□□□E□□□□□□SNOMED 72836002□□2□□

□Maps to value□□□□□□□□□□□□□□□□□□□□□□-□□-□□EAV□□□□□□□□□□□OMOP CD-M□□

- □□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□
- □□□□□□□□□□□□

□□
to□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Maps to value□□□□□□□□□□□□□□□□□□□□□□□

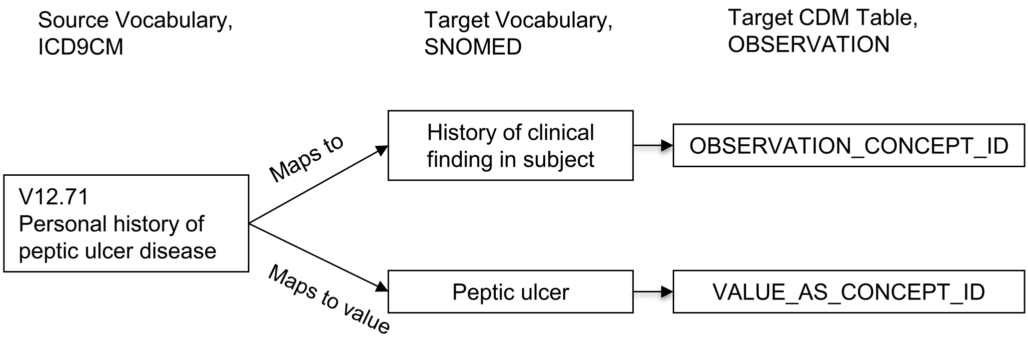


Figure 5.5: □□2□□
to□□Maps to
value□□□□□□□□□□□□□□□□□□□□□□

□□□OMOP
□□□□□□□□□□□□□□□OHDSI Wiki□□□□□□□□□□□□□[5]

5.3.2

□□□□□□□□□□Is a□-□Subsumes□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
49436004□□□□□□□□SNOMED 17366009□□□□□□□□□□Is a□□□□□□□□□□□□□□□□□□□
49436004□□□□□□□SNOMED 40593004□□□□□□□□□Is a□□□□□□□□

5.3.3

□□□□□□□□□□□Vocabulary A - Vocabulary B equivalent□□□□□□□□□□□□□□□□□□□□□□
- RxNorm equivalent□□□□□□Maps to□□□□□□□□□□□□□□

5.3.4

□□OHDSI
Wiki□□□□□□□□□□□□□□□□□□□□□□□□□□[6]

□□□
site of□□□□□□□□□□□□□□□□□□□□5.5□□□□□

[5]https://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary:mapping
[6]https://www.ohdsi.org/web/wiki/doku.php?id=documentation:vocabulary

Table 5.5: □□□□Finding site of□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

CONCEPT_ID_1	CONCEPT_ID_2
4000504 “Urethra part”	36713433 “□□□□□□□□”
4000504 “Urethra part”	433583 “□□□□□□□”
4000504 “Urethra part”	443533 “□□□□□□□□□”
4000504 “Urethra part”	4005956 “□□□□□□□□□”

□□□

5.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CONCEPT_ANCESTOR□□□□□□

CONCEPT_ANCESTOR□□□□□□CONCEPT_RELATIONSHIP□□□□□□□□□□□□□□□□ “Is a” - “Subsumes” □□□□□ 5.6 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

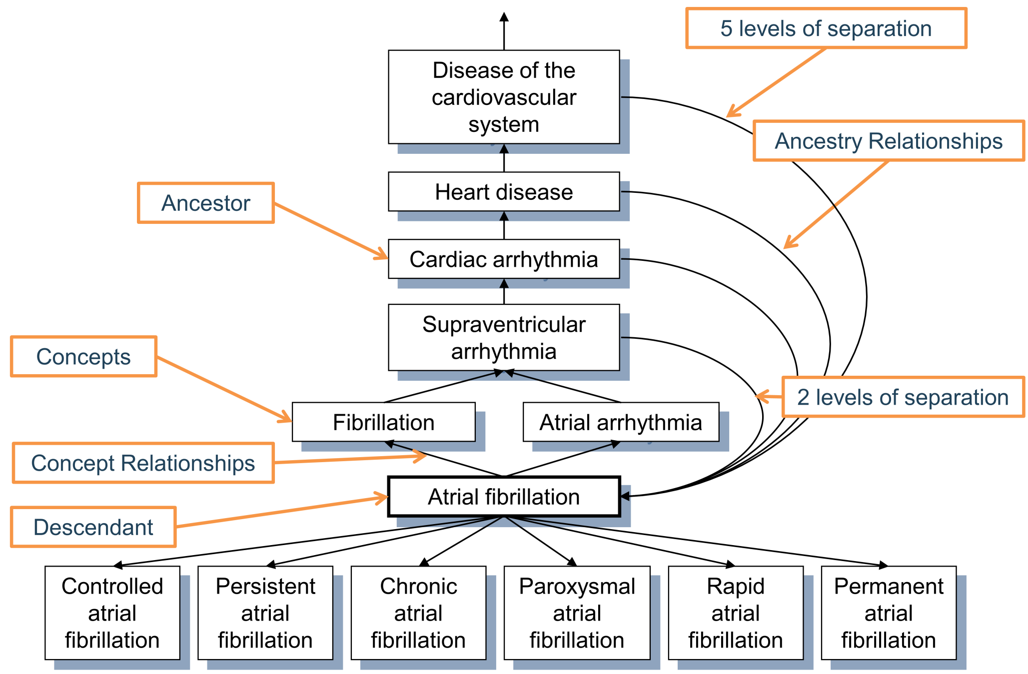


Figure 5.6: □□□□□□□□□□□□□□□□□□□Is a□□□Subsumes□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□MIN_LEVELS_OF_SEPARATION□□□MAX_L

□□□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□

5.5

DOMAIN_ID□VOCABULARY_ID□CONCEPT_CLASS_ID□□□□CONCEPT□□□□□□□□□□□CON

VOCABULARY□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□VOCABULARY_REF

5.6

5.6.1

OMOP CDM□□

5.6.2

□□□

5.6.3 OMOP

□□□□□□□□□□□□□□□□□□□□□□ICD-9□ICD-10□□□□□□□□□□□□□□□□□□□□□□□□

5.6.4

□□□□□HCPCS□CPT4□□□

5.6.5

□□□

5.6.6

□□□

5.6.7

□□□

5.6.8

□□□

5.6.9

□□□RxNorm□□
Extension□□□□□□□□□□□□□□□□□□□□

5.6.10 NULL

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5□□□□□□□□□□8507□□□□□8

5.7

- □□□□□□□□□□□□□□□□OMOP□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□OHD
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□pping□□□□□□□□□□
- □□□□□□□Maps to□□□□□Maps to value□□□□□□□□□□□□□□□□□□□□□□□
- □□□
- □□□□□□□□□□□□□□□□□□□□□□□□□
- □□□
- □□

5.8

□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATHENA[7]□□□ATLAS[8]□

□□ **5.1.** “□□□□”□□□□□□□□□ID□□□□□□

□□ **5.2.** “□□□□”□□□□□□□□□□□□□□□□□□□ICD-10CM□□□□□□□□□□□□□□□□□□□□□□□9CM□□□□□□□□□□□

□□ **5.3.** “□□□□”□□□□□□□□□□□□□□□□MedDRA□□□□□□□□□□□□□

□□□□□□□□□□□ E.2□□□□□□□□□□□□

[7]http://athena.ohdsi.org/
[8]http://atlas-demo.ohdsi.org

Chapter 6

6 – –

□□□□: *Clair Blacketer & Erica Voss*

6.1

□□□□□□/□□□□□□□OMOP□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□

ETL□□□4□□□□□□□□□□□□□□□□□□□□

1. □□□□□□□□□□□CDM□□□□□□□□□□□ETL□□□□□□□□□□□
2. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
3. □□□□□ETL□□□□□□□□
4. □□□□□□□□□□□□□□□□□□□

□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□
□□□□ 1: ETL□□□□□□

ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□

ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□White Rabbit□Rabbit-in-a-Hat□□□

6.1.1 White Rabbit

□□□□□□□□□ETL□□□
Rabbit□□□□□□□White Rabbit□□□□□□□□□□□□□□□□□□□□□□□□□OMOP CDM□ETL□□□□□□□□□□□□□□□□□
Rabbit□□□□□□□□□□□□□□□□ETL□□

[1] https://github.com/OHDSI/WhiteRabbit.

□□□□□

White Rabbit□□□ Server□Oracle□PostgreSQL□Microsoft APS□Microsoft Access□Amazon Redshift□□□□□□□□□□ In-a-Hat□□□□□□□□□□□□□□□□□□□White Rabbit□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1. □□□
2. □□□□□□□□□□□□CSV□□□□□□□□□□□□□□□□□□□□□□□□
3. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□
4. White Rabbit□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□

White Rabbit□□□ Rabbit□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□6.1□□□□□□□□Pick Folder□□

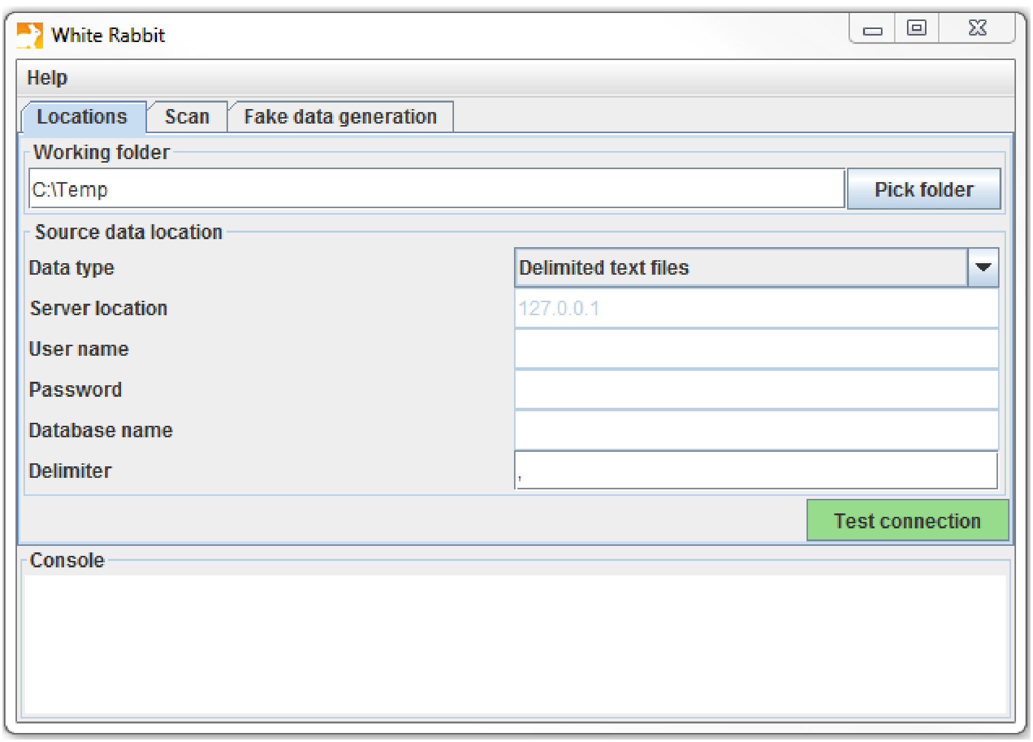


Figure 6.1: White Rabbit□□□□□□□□□□□□□□□□□□□□□□□□Pick Folder□□□□□

□□□□□□□□□□

White Rabbit□□□

□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□
+ □□
all in DB□□□□□□□□□□□□□□

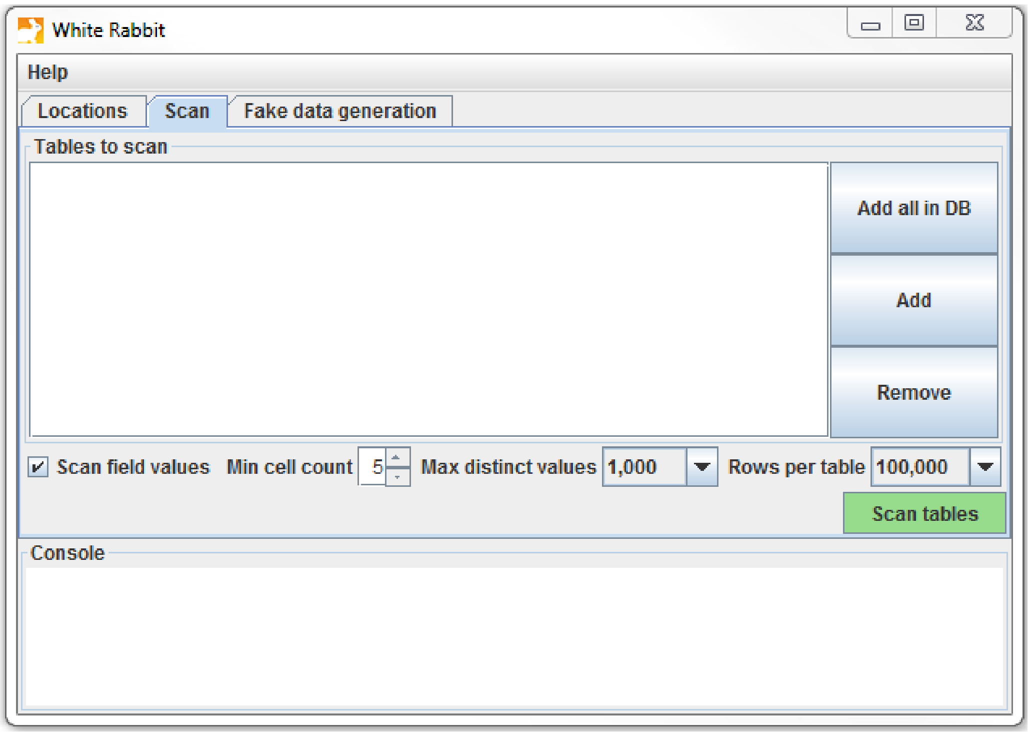


Figure 6.2: White Rabbit □□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□WhiteRabbit□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5□□□□□□□□□□□
- □□White Rabbit□□□□□□□100,000□□□□□□□□□□□□□□□□□□

□□□

□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Excel□□□□□□□□□□□□□□□□

□□□

	A	B	C	D	E	F	G
1	Table	Field	Type	Max length	N rows	N rows checked	Fraction empty
2	dbo.allergies	start	date	10	3184	3184	0
3	dbo.allergies	stop	date	10	3184	3184	0.725188442
4	dbo.allergies	patient	varchar	36	3184	3184	0
5	dbo.allergies	encounter	varchar	36	3184	3184	0
6	dbo.allergies	code	varchar	9	3184	3184	0
7	dbo.allergies	description	varchar	24	3184	3184	0
8							
9	dbo.careplans	id	varchar	36	30199	30199	0
10	dbo.careplans	start	date	10	30199	30199	0
11	dbo.careplans	stop	date	10	30199	30199	0.057849598
12	dbo.careplans	patient	varchar	36	30199	30199	0
13	dbo.careplans	encounter	varchar	36	30199	30199	0
14	dbo.careplans	code	varchar	15	30199	30199	0
15	dbo.careplans	description	varchar	62	30199	30199	0
16	dbo.careplans	reasoncode	varchar	9	30199	30199	0.050796384
17	dbo.careplans	reasondescription	varchar	56	30199	30199	0.050796384
18							

Overview | dbo.allergies | dbo.carepla ...

Figure 6.3: □□□□□□□□□□□□□□□□□□

	A	B
1	Sex	Frequency
2	2	61491
3	1	35401
4	List truncated...	

Figure 6.4: □□□□□□□□□□

□□6.4□□□□□□□
Rabbit□1□□□□2□□□

6.1.2 Rabbit-In-a-Hat

White Rabbit□□CDM□□□□□□□□□□ Rabbit□□□□□□□□□□□□□□□□□□□□□□Rabbit-in-a-Hat□□ in-a-Hat□□

□□□□□

Rabbit-In-a-Hat□White Rabbit□□□W Rabbit□□□□□□□□□□□□□□□□□□□□□□□□□□□□Rabbit-In-a-Hat□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ In-a-Hat□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□

□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1. White Rabbit□□□□□□□□□□□□□□□□□□□□□□□□□□□
2. □□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□
3. □□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□
4. □□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□
5. Rabbit-In-a-Hat□□□□□□□□□□□□□□□□□□□□□MS Word□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

ETL□□□□□□□□□

White Rabbit□□□□□□□□□□□□□□□□□Rabbit-In-a-Hat□□□□□□□□□□□□□□□□□□□□□□□□□OMOP CD-M□□□Synthea[2]□□□□□

ETL□□□□□□□□□□□

CDM□□□□□□□□□□□□□□□□□□□□□□□□□PERSON□□

CDM□□VISIT_OC

□□□□□□□□**Person**□□□□□

Synthea□□□□□□□□□□patients□□□□□□20□□□□□□□□□□□□□□□□□□□□□□6.6. □□□□□□□□□□□□□□□□□□□□□□□□ patients□□□□□□□□□□□□□□□□□□□□□□□□□□CDM PERSON□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□**??**□□□□Synthea patients□□□□□□CDM PERSON□□□ Field□□□□□□□□□□□□□□□□□□□□□CDM□□Source field□□□□□□□□□□□□□□□□□□□□□□□CDM□□□patients□ & comments□□

[2]Synthea™□□□

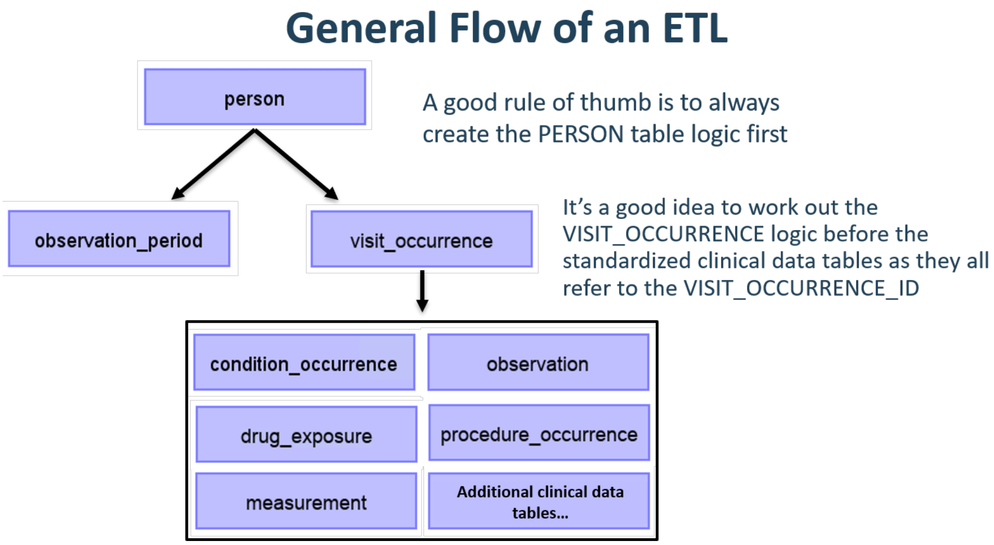


Figure 6.5: ETL□□□□□□□□□□□□□□□□□□□□□□□□□

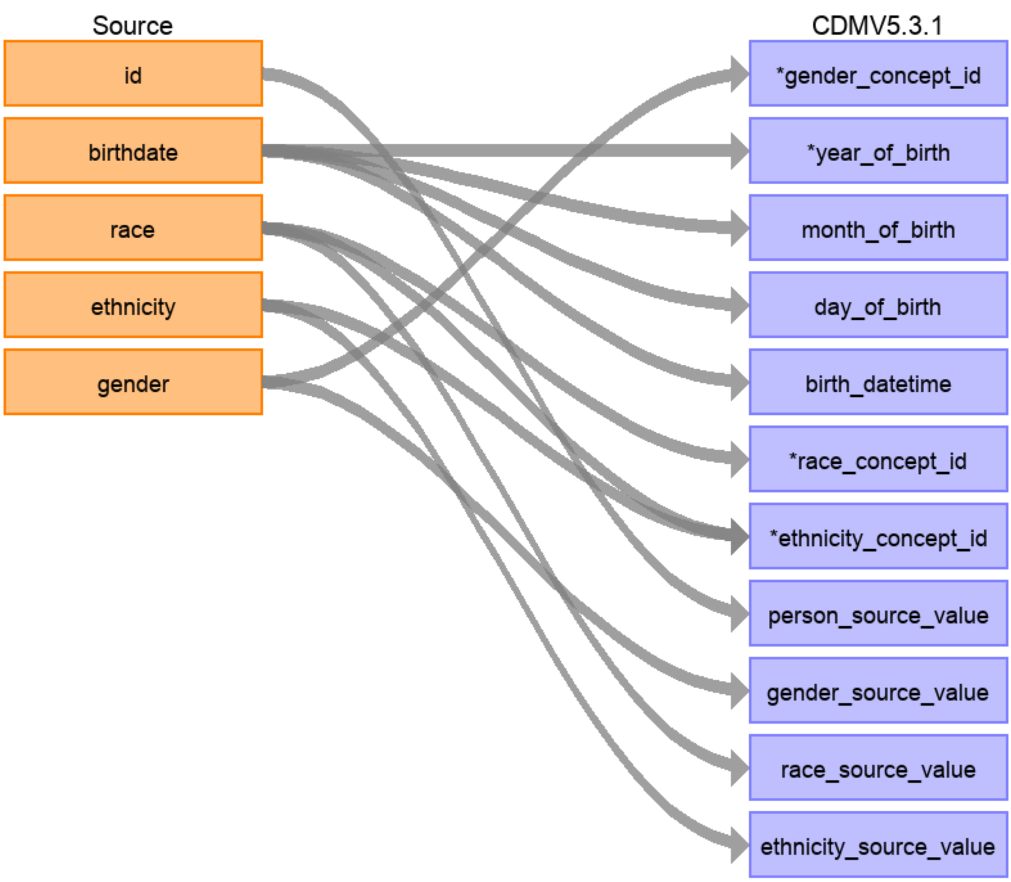


Figure 6.6: Synthea Patients□□□□□□CDM PERSON□□□□□□□□□□□□

□□(□tab:syntheaEtlPerson) Synthea Patients□□□□□CDM PERSON□□□□□□□□□□□□ETL

□□□□□□□	□□□□□□□	□□□□□□□□□□□□
PERSON_ID		□□□□□□ PER-SON_ID□□□□□□□□□□□□□□□□□□□□id□□var□□□□id□□□□□□□□□□□□□□□□□□□□□□
GENDER_CONCEPT_ID	gender	□□□□M□□□□□GENDER_CONCEPT_ID□8507□□
YEAR_OF_BIRTH	birthdate	□□□□□□

6.2 2:

OMOP□□CDM□□□□10CM□□□□□□□□□□□□□□□□SNOMED□□□□□□□□□□□□□□□□□□relationship_id =□Maps to□□□□CONCEPT_RELATIONSHIP□□□□□□□□□□□□□□□□□□□□□□ICD-10CM□□□□I21□□□□□□□□□□□□□□□□□□□ID□□□□□□□□□□□□SQL□□□□□□□

```sql
SELECT concept_id_2 AS standard_concept_id
FROM concept_relationship
INNER JOIN concept AS source_concept
  ON concept_id = concept_id_1
WHERE concept_code = 'I21'
  AND vocabulary_id = 'ICD10CM'
  AND relationship_id = 'Maps to';
```

STANDARD_CONCEPT_ID
312327

□□

- □□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATC□□□□□□□□
- Usagi□□□□

6.2.1 Usagi

Usagi□□Translate[3] □□□□□□□□□□□□□□□□□□□□□□□□□Usagi□□□□□□□□□□□□□□□□□□□□□□

□□□□□

□□□□□□□□□□□□□□□□Usagi□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

[3] https://translate.google.com/
[4] https://github.com/OHDSI/Usagi

□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
2. Usagi□□□
3. Usagi□□
4. □□□□□□□SOURCE_TO_CONCEPT_MAP□□□□□□□□□□□□□□□□□

□□□□□□□Usagi□□□□□□

□□□□□□□CSV□□□Excel (.xlsx) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ Translate□□□□□□□

□□□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□File → Import codes□□□□□□Usagi□□□□□□□□□□□□□□Import codes …□□□□□□□□□6.7□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

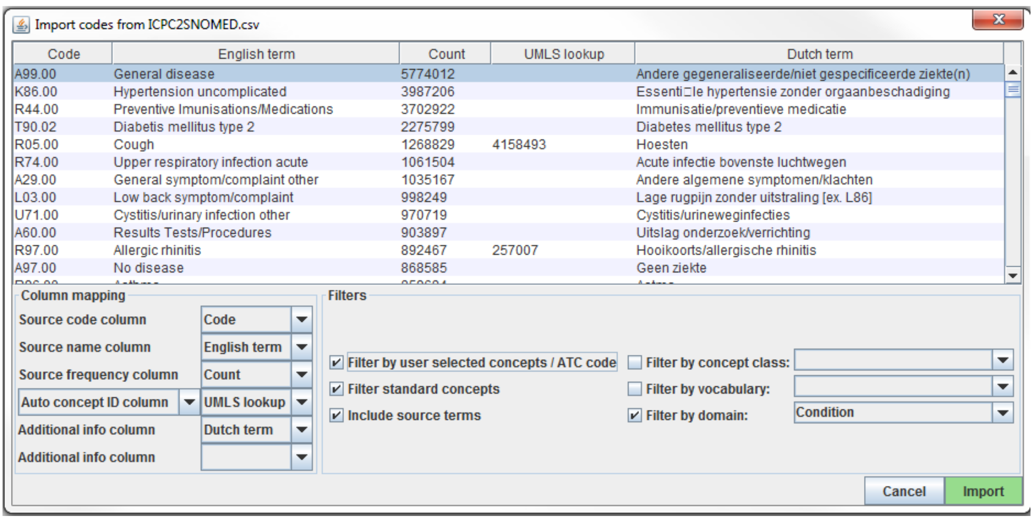


Figure 6.7: Usagi□□□□□□□□

□Column mapping□□□□□□□□□□□□□□□□□□□□□□□□□Usagi□□□□□□□□□□□□□□□□ info□□□

□□□□Filters□□□□□□□□□□□□Usagi□□□□□□□□□□□□□□□□□□□□□□□□6.7□□□□□ standard concepts□□□□□□□□□Usagi□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□Filter by automatically selected concepts / ATC code□□□□□□□□□□□□□□□□ concept ID□□□□□□□□□□□□□□□□CONCEPT_IDs□□□ATC□□□□□□□□□□□□□□□□□□□□

1. Column mapping□□□“Auto concept ID column”□□”ATC column”□□□□□□
2. Column mapping□□□ATC□□□□□□□”AT

6.3 3: ETL

ETL ETL
6.6

ETL

ETL

- ETL-Synthea - Synthea SQL
  - https://github.com/OHDSI/etl-synthea
- ETL-CDMBuilder - .NET
  - https://github.com/OHDSI/etl-cdmbuilder
- ETL-LambdaBuilder - AWS lambda
  - https://github.com/OHDSI/etl-lambdabuilder

” ”ETL

ETL

6.4 4:

-
> -> -> CDM

- ETL 1
  -
-
  - 1 1
-
  -
- CDM
  - CDM
- ETL
  - ETL

ETL OHDSI
15

6.5 ETL THEMIS

CDM ETL

OHDSI OHDSI
Wiki [5] CDM ETL

[5] https://github.com/OHDSI/CommonDataModel/wiki

□□
Wiki□□□□□□□□□□□□□□□□□□□□□□Themis□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

6.6 CDM ETL

ETL□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

□□

OMOP□□□1□□□□□□□□□□

CDM□□□ETL□□

6.7 ETL

ETL□□

- 80/20□□□
- □□
- CDM□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- OHDSI CDM□□[8]

6.8

- ETL□□□
  - □□□□□□□□□□CDM□□□□□□□□□□□□ETL□□□□□□
  - □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
  - □□□□□ETL□□□
  - □□□□□□□□□□□□□□□□□□□□
- OHDSI□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□

6.9

□□ **6.1.** ETL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

A) □□□□□□□□□□□CDM□□□□□□□□□□□□□ETL□□□□□□

B) □□□□□ETL□□□

[6] https://github.com/OHDSI/Themis
[7] http://forums.ohdsi.org/
[8] https://forums.ohdsi.org/c/implementers

C) □□□□□□□□□□□□□□□□□□□□□□□□□□□□

D) □□□□□□□□□□□□□□□□□□□□

□□ **6.2.** □□□□OHDSI□□□□□□□□□□□□6.2□□□PERSON□□□□□□□□□4□□□□□□□□□□

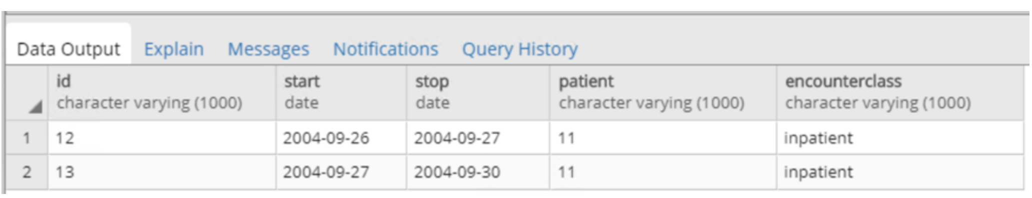
□: PERSON□□□□□

□	□
PERSON_ID	A123B456
GENDER_CONCEPT_ID	8532
YEAR_OF_BIRTH	NULL
MONTH_OF_BIRTH	NULL
DAY_OF_BIRTH	NULL
RACE_CONCEPT_ID	0
ETHNICITY_CONCEPT_ID	8527
PERSON_SOURCE_VALUE	A123B456
GENDER_SOURCE_VALUE	F
RACE_SOURCE_VALUE	WHITE
ETHNICITY_SOURCE_VALUE	□□□□□□□□□

□□ **6.3.** VISIT_OCCURRENCE□□□□□□□□□□□□□□□□□□□□□Synthea□□□□□□□□□□□□
PATIENT□START□END□□□□□□□□□□□□□□□□PERSON_ID□□□□□□□□END□□□□□

- MIN(START)□VISIT_START_DATE□□□□□
- MAX(END)□VISIT_END_DATE□□□□□
- “IP”□PLACE_OF_SERVICE_SOURCE_VALUE□□□□□

□□□□□□□□6.8□□□□□□□□□□□□□□□□□□□CDM□□□□□□VISIT_OCCURRENCE□



	id character varying (1000)	start date	stop date	patient character varying (1000)	encounterclass character varying (1000)
1	12	2004-09-26	2004-09-27	11	inpatient
2	13	2004-09-27	2004-09-30	11	inpatient

Figure 6.8: □□□□□□□□□

□□□□□□□□□□E.3□□□□□□□□□□□□□

7 – –

Chapter 7

□□□□□□□□*: David Madigan*

OHDSI□□□

- □□□□
- □□□□□□□□
- □□□□□□□□

□□;
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapters 14 - 18□□□□□□□□□□□

7.1

□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□

□□□

□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□
- 65□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□…□
- □□□□□□□□□…□
- □□□□□□…□
- □□□□□□□□□□□□□□□□□□□□…□
- □□□HbA1c□□□□…□
- □□□□□□□…□

- □□□□□□□□□□□□…□
- □□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□
- □□□□□□
- …□□□□□□□□□□□
- X□□□□□□□□□□□□
- …□□□□□□□□
- …□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- …□□□□□□□

□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□
- □□
- □□□□
- □□□
- □□□
- □□□□
- □□□□□□□□□□□□□
- □□□□
- □□□□□□
- □□□□□□□
- □□□□□□□□
- □□□□□□□
- □□□□□

7.2

□□

□□□□□□□□

□□

□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- …□□□□□
- □□…□□□□□□□□
- □□□□□□□□□□□
- Y□□□□X□□□□□□
- …□□□□□□□□□□□□

□□□□□□□□□□□□□□□

- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]

7.3

[illegible]

[illegible]

[illegible]

- [illegible]1[illegible]
- [illegible]1[illegible]

[illegible]

- [illegible]…[illegible]
- …[illegible]

[illegible]

- [illegible]
- [illegible]
- [illegible]/[illegible]
- [illegible]

[illegible]

> [illegible]

7.4

[illegible]ACE[illegible]

7.4.1

[illegible]
10[illegible]Chapter 11[illegible]

7.4.2

□□□□□□□□□□□□□Chapter 12□□□□□□□□□□□□□□□□□□□□□□□□□□ACE□□□□□□

7.4.3

□□□
13□□□□□□□□□ACE□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□

7.5

OHDSI□□□:

- □□□
- □□□□□□□□□□□□□□□
- □□□
- □□□

7.5.1

OHDSI□□□

7.5.2

OHDSI□□□
perkins2017principled□ □□□□□□□□□□□□□□□□□□□□□□□□□□

7.6

- □□□□□□□3□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□

7.7

□□ **7.1.** □□□□□□□□□□□□□□□□□□□□□□□□

1. □□NSAIDs□□□□□□□□□□□□□GI□□□□□□□□□□□

2. □□□□□□□□□□□□□□□□□□□□□□GI□□□□□□□□□□□□□□□□

3. □□□□□□□□□□□□□□□□□□□□□GI□□□□□□□□□□□□□□□

□□**7.2.** □□□□□□□□□□□□□□□□□□□□□□GI□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□Appendix E.4□□□□□□

Chapter 8

8 – – OHDSI

□□□□□*Lead: Martijn Schuemie & Frank DeFalco*

OHDSI□□□

□□

8.1

□8.1□□CDM□□

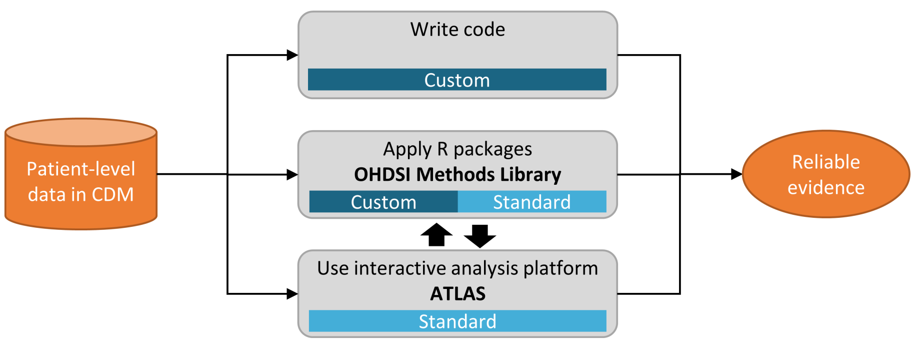


Figure 8.1: CDM□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□

2□□□□□□□R□□□□□□□□□OHDSI Methods Library□□□□□□□□□□□□□□□□□□□Cha
9□□□SqlRender
Library□□

3□□□ATLAS□□
Libraries□□□

Library□□□ATLAS□□□□□

ATLAS□Methods Library□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□Methods Library□□□□□□□□□□□□□□□□□□□□□□□□□□□Methods Library□□□□□□□□□□□□ATLAS

8.2

CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Methods Library□□□

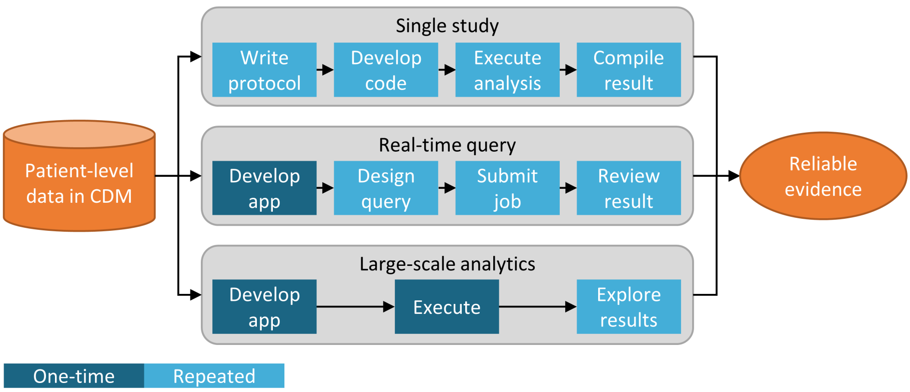


Figure 8.2: (□□□)□□□□□□□□□□□□□□□□□□□□

□□ et al., 2017,)□ □□□□□□□□□□□□□□□□□□□OHDSI Methods Library□□□□□□□□□□□□□□□□

□□

□□ et al., 2018b,)□ □□□□□□□□□□□□□□□□□□□4□□□□□□□□□□□□□□□□□□□□□□□□□□□ //data.ohdsi.org/SystematicEvidence/

8.3 ATLAS

ATLAS□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□ WebAPI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Apache Tomcat□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□

□8.3□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□

□□□□□□ □□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□ ATLAS□OMOP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

[1] http://www.ohdsi.org/web/atlas

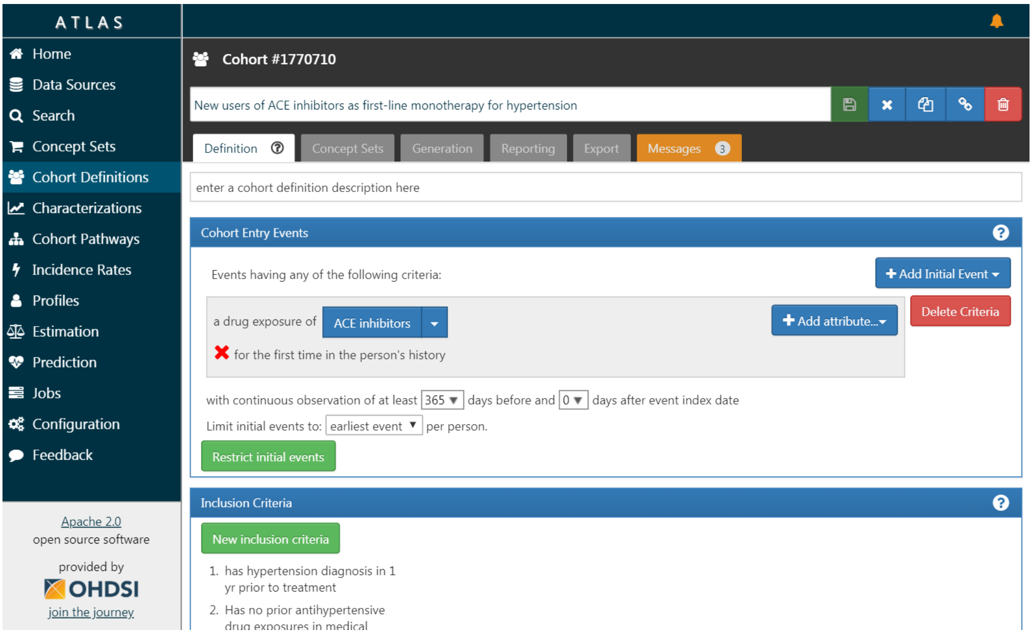


Figure 8.3: ATLAS□□□□□□□□□□

□□□□□□□□□ □□

□□□□□□ □□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□ □□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□ □□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□ □□□11□

□□□□□□ □□□

□□□□□□□ □□□

□□□□□□□ □□□

□□□ □□□□□□□□□□□□□□□WebAPI□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ □□□

□□□□□□□ □□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□

8.3.1

ATLAS□WebAPI□□□ Shiro□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□WebAPI□□□□

[2]https://github.com/OHDSI/WebAPI/wiki/Security-Configuration

8.3.2

ATLAS□□□□□□□□□□ATLAS GitHub□□□□□□□wiki□□□□□□□□□□□□□□□□[3] □□wiki□□

8.3.3

ATLAS□□□□□□□□□□OHDSI WebAPI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ GitHub□□□□□□□□□□□□□□□□□□[4]□□□WebAPI GitHub□□□□□□□□□□□□□□□□□□□□□□□□□□□□

8.4

OHDSI□□□□□□□□□□□□□□8.4 □□□□□□□□□□□□□□□□R□□□□□□□□□□□□□□□□

□□□□□□□□□□□□CDM□□ 9□Chapter 11□Chapter 12□Chapter 13□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ (Boland et al., 2017, ; Duke et al., 2017, ; Ramcharran et al., 2017, ; Weinstein et al., 2017, ; Wang et al., 2017, ; Ryan et al., 2017, ; Ryan et al., 2018, ; Vashisht et al., 2018, ; Yuan et al., 2018, ; Johnston et al., 2019,) □□□□□□□□□□□□□□□□□□□□□□□□□□ (Schuemie et al., 2014, ; Schuemie et al., 2016, ; Reps et al., 2018, ; Tian et al., 2018, ; Schuemie et al., 2018a, ; Schuemie et al., 2018b, ; Reps et al., 2019,)□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 17□□□□□□□□□□□□

8.4.1

□□□

□□Chapter 18□□

8.4.2

□□□

1. □□□
2. □□□
3. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Cyclops□□□□□□□□□□□□□□□□

[3] https://github.com/OHDSI/ATLAS/wiki
[4] https://github.com/OHDSI/Atlas/wiki/Atlas-Setup-Guide

Figure 8.4: OHDSI□□□□□□□□□□□□□□□□□

8.4.3

R□□ Library□□□□□□□[5]□□□□□□□GitHub□□□□□□CRAN□□□□□□□□□□□□R□□□□□□□□□□□

□□□ Library□□□□□□□□□□□□□GitHub□□□□□□□□□CRAN□□□□□□□□□□□□□□

8.4.4

□□□

□□□□□□□□□□□CDM□□ SQL Server□Oracle□□□□□□□□□□□□□□□□Microsoft APS□IBM Netezza□Amazon Redshift□□□□□□□□□□□□□□□□□□□□□Impala□□□□Hadoop□Google Big-Query□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□ Server□□□□□□□□□□□□□□□□□□□□□R□□□□□□□□□□□□□□□□□□□□□□□□□RStu Library□Java□□

8.4.5

OHDSI R□□

1. **R**□□
2. **Rtools**□□Windows□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
3. **RStudio**□□R□□□□□□□□IDE□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

8.5

ATLAS□Methods Library□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□ Web Services (AWS)□□□

8.5.1 Broadsea

Broadsea[7]□Docker□□□□□□□□□□□□□□[^^dockerUrl] OHDSI□□□□□□□□□□□□□□Docker□□□ Windows□MacOS□Linux□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Broadsea Docker□□□□□□□Methods Library□ATLAS□□□□□□OHDSI□□□□□□□□□□□□□

[5] https://ohdsi.github.io/MethodsLibrary
[6] https://ohdsi.github.io/CohortMethod/articles/MultipleAnalyses.html
[7] https://github.com/OHDSI/Broadsea

8.5.2 Amazon AWS

Amazon□□AWS□□
in-a-Box[8]□OHDSIonAWS□□□[^^ohdsiOnAwsUrl]

OHDSI-in-a-Box□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□
CDM□□□□□□□□□□□ETL□□□□OHDSI-in-a-Box□□□□□□□□□OHDSI-in-a-Box□□□□□□□□□□8.5□□□□□□□□□□

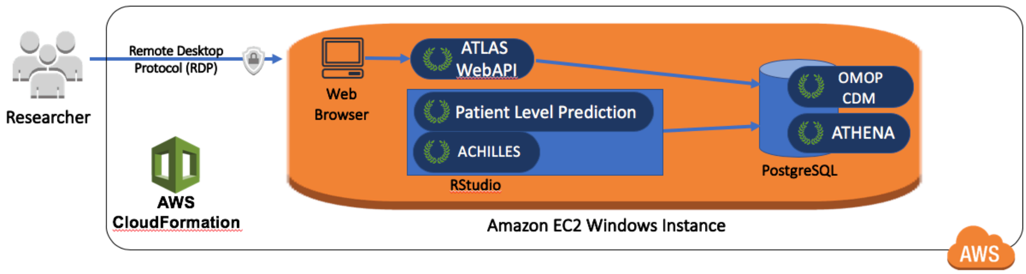


Figure 8.5: OHDSI-in-a-Box□Amazon Web Services□□□□□□□□

OHDSIonAWS□□□OH
Redshift□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□ATLAS□□□□
Server□□□□□RStudio□□OHDSI Methods Library□□□□□□□□□□□□□□□□□□□□□□□

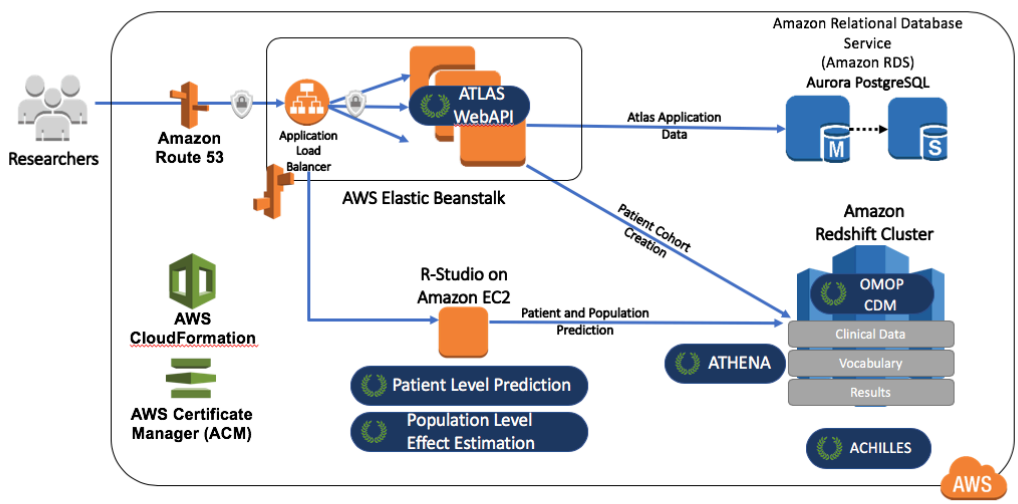


Figure 8.6: OHDSIonAWS□Amazon Web Services□□□□□□□□

[8] https://github.com/OHDSI/OHDSI-in-a-Box

8.6

- CDM□□□□□□□□□□□□□□□□
  - □□□□□□□□□□
  - OHDSI Methods Library□R□□□□□□□□□□□□□□□
  - □□□□□□□□□□□□□□□□□□ATLAS□□□
- OHDSI□□□□□□□□□□□□□□□□□
  - □□□□
  - □□□□□□□□□
  - □□□□□□□□□□
- □□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□
  - □□□□□□□□□□□□□□□□□□ATLAS
  - OHDSI Methods Library□R□□□□□
- OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□

Chapter 9

9 □□ – □□ – SQL R

□□□□□□□□: *Martijn Schuemie & Peter Rijnbeek*

□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
SQL Server□□□ATI
Library□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□SQL□Structured Query Language□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Server□PERSON□□□□□□□10□□□□□□□□□□□□□□□□□□□

```sql
SELECT TOP 10 * FROM person;
```

□□□PostgreSQL□□□□□□□□□□□□□□□□□□

```sql
SELECT * FROM person LIMIT 10;
```

OHDSI□□OHDSI□□
- **OHDSI SQL** - □□□SQL Server SQL□□□□□□□□□□□□□□□□□□SQL□□□□□OHDSI SQL□□□□□□□

□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□CDM□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□SqlRender□Database
SQL□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□OHDSI□□□□□□□□□□□□

9.1 SqlRender

SqlRender CRAN Comprehensive R Archive Network

```r
install.packages("SqlRender")
```

SqlRender PostgreSQL Microsoft SQL Server SQLite Oracle APS IBM Netezza Amazon Redshift Impala Hadoop Goo BigQuery R

9.1.1 SQL

SQL

@
a render

```r
sql <- "SELECT * FROM concept WHERE concept_id = @a;"
render(sql, a = 123)
```

```
## [1] "SELECT * FROM concept WHERE concept_id = 123;"
```

```r
sql <- "SELECT * FROM @x WHERE person_id = @a;"
render(sql, x = "observation", a = 123)
```

```
## [1] "SELECT * FROM observation WHERE person_id = 123;"
```

```r
sql <- "SELECT * FROM concept WHERE concept_id IN (@a);"
render(sql, a = c(123, 234, 345))
```

```
## [1] "SELECT * FROM concept WHERE concept_id IN (123,234,345);"
```

If-Then-Else

{Condition} ? {if true} : {if false} *condition*
true 1 *if true* *if false*

```
sql <- "SELECT * FROM cohort {@x} ? {WHERE subject_id = 1}"
render(sql, x = FALSE)
```

```
## [1] "SELECT * FROM cohort "
```

```
render(sql, x = TRUE)
```

```
## [1] "SELECT * FROM cohort WHERE subject_id = 1"
```

□□□□□□□□□□□□□□□□□

```
sql <- "SELECT * FROM cohort {@x == 1} ? {WHERE subject_id = 1};"
render(sql, x = 1)
```

```
## [1] "SELECT * FROM cohort WHERE subject_id = 1;"
```

```
render(sql, x = 2)
```

```
## [1] "SELECT * FROM cohort ;"
```

`IN` □□□□□□□□□□□□□□□

```
sql <- "SELECT * FROM cohort {@x IN (1,2,3)} ? {WHERE subject_id = 1};"
render(sql, x = 2)
```

```
## [1] "SELECT * FROM cohort WHERE subject_id = 1;"
```

9.1.2 SQL

SqlRender □□□□□□□□□□□□□□□OHDSI SQL□□□□SQL□□□□□□□□□□□□□□□□

```
sql <- "SELECT TOP 10 * FROM person;"
translate(sql, targetDialect = "postgresql")
```

```
## [1] "SELECT  * FROM person LIMIT 10;"
## attr(,"sqlDialect")
## [1] "postgresql"
```

`targetDialect` □□□□□□□□□□□□□□□□□□“oracle”, “postgresql”, “pdw”, “redshift”, “impala”, “netezza”, “bigquery”, “sqlite”, “sql server”□

> □□□□SQL□□
> SQL□□□□□□□SQL□□
> Server□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Translate□□□□□□□□□□□□□□□□

□□□□SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□: supported by translate

□□	□□	□□
ABS	EXP	RAND
ACOS	FLOOR	RANK
ASIN	GETDATE	RIGHT
ATAN	HASHBYTES*	ROUND
AVG	ISNULL	ROW_NUMBER
CAST	ISNUMERIC	RTRIM
CEILING	LEFT	SIN
CHARINDEX	LEN	SQRT
CONCAT	LOG	SQUARE
COS	LOG10	STDEV
COUNT	LOWER	SUM
COUNT_BIG	LTRIM	TAN
DATEADD	MAX	UPPER
DATEDIFF	MIN	VAR
DATEFROMPARTS	MONTH	YEAR
DATETIMEFROMPARTS	NEWID	
DAY	PI	
EOMONTH	POWER	

* Oracle□□□□□□□□□□□□□□SQLite□□□□□□□□□□□□□□

□□□□□□□SQL□□□

```sql
-- Simple selects:
SELECT * FROM table;

-- Selects with joins:
SELECT * FROM table_1 INNER JOIN table_2 ON a = b;

-- Nested queries:
```

```
SELECT * FROM (SELECT * FROM table_1) tmp WHERE a = b;

-- Limiting to top rows:
SELECT TOP 10 * FROM table;

-- Selecting into a new table:
SELECT * INTO new_table FROM table;

-- Creating tables:
CREATE TABLE table (field INT);

-- Inserting verbatim values:
INSERT INTO other_table (field_1) VALUES (1);

-- Inserting from SELECT:
INSERT INTO other_table (field_1) SELECT value FROM table;

-- Simple drop commands:
DROP TABLE table;

-- Drop table if it exists:
IF OBJECT_ID('ACHILLES_analysis', 'U') IS NOT NULL
  DROP TABLE ACHILLES_analysis;

-- Drop temp table if it exists:
IF OBJECT_ID('tempdb..#cohorts', 'U') IS NOT NULL
  DROP TABLE #cohorts;

-- Common table expressions:
WITH cte AS (SELECT * FROM table) SELECT * FROM cte;

-- OVER clauses:
SELECT ROW_NUMBER() OVER (PARTITION BY a ORDER BY b)
  AS "Row Number" FROM table;

-- CASE WHEN clauses:
SELECT CASE WHEN a=1 THEN a ELSE 0 END AS value FROM table;

-- UNIONs:
SELECT * FROM a UNION SELECT * FROM b;

-- INTERSECTIONs:
SELECT * FROM a INTERSECT SELECT * FROM b;

-- EXCEPT:
SELECT * FROM a EXCEPT SELECT * FROM b;
```

□□□□□□

□□□□□□□SQL Server□□□□□□□□□□□□□□□□□1□□□□SQL Server□□□SELECT first_name + ' ' + last_name AS full_name FROM table □□□□□□□□□□ PostgreSQL □ Oracle □□ SELECT first_name || ' ' || last_name AS full_name FROM table □□□□□□□□□□□SqlRender□□□□□□□□□□□□□□□□□□□□□□□□ SELECT first_name + last_name AS full_name FROM table □□□□□□□SqlRender□2□□□□ last_name + CAST(age AS VARCHAR(3)) AS full_name FROM table □□□□□□□□□□□□□□ □□□□□□□□□□□□□□□

□□□□□□□□□□AS□□□□□

□□□SQL□□□□□□□□□□□□□□□□□□□□AS □□□□□□□□□□□□□□□□□□□□□□□□□ Server□PostgreSQL□Redshift□□□□□□□□□□□□□□□□□□□□

```sql
-- Using AS keyword
SELECT *
FROM my_table AS table_1
INNER JOIN (
  SELECT * FROM other_table
) AS table_2
ON table_1.person_id = table_2.person_id;

-- Not using AS keyword
SELECT *
FROM my_table table_1
INNER JOIN (
  SELECT * FROM other_table
) table_2
ON table_1.person_id = table_2.person_id;
```

□□□□Oracle□□ AS □□ AS □□□□□□□□□□□□□□□□□□□□□□SqlRender□□□□□□□□□□□□□□□□□Oracle□□□□ AS □□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□

□□□

1. □□□
2. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□

```r
sql <- "SELECT * FROM #children;"
translate(sql, targetDialect = "oracle", oracleTempSchema = "temp_schema")
```

```
## Warning: The 'oracleTempSchema' argument is deprecated. Use 'tempEmulatio
## This warning is displayed once every 8 hours.
```

```
## [1] "SELECT * FROM temp_schema.n0hfc14ochildren ;"
## attr(,"sqlDialect")
## [1] "oracle"
```

□□□□□ `temp_schema` □□□□□□□□□□□□□□□□□□□□□

□□□Oracle□□□□□□□□□□□30□□□□□□□□□□□□□□□□□□□□□□□**22**□□□□□□□□

□□□□Oracle□□□ `TRUNCATE` □□□ `DROP` □□□□□□□□□□□□Oracle□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□

SQL Server□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□ Server□□□□□□□

```sql
CREATE TABLE #temp (txt VARCHAR);

INSERT INTO #temp
SELECT '1';

SELECT * FROM #temp WHERE txt = 1;
```

`txt` □VARCHAR□□□□□□□□□□□□□□□□□□□□□□□□□□SQL Server□□□□□□□

□□

```sql
SELECT * FROM #temp WHERE txt = CAST(1 AS VARCHAR);
```

□□□

```sql
SELECT * FROM #temp WHERE CAST(txt AS INT) = 1;
```

□□□□□□□□□□□□□□□□□□

SQL Server□□□□□□DBMS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```sql
SELECT * FROM concept WHERE concept_class_id = 'Clinical Finding'
```

□□□□□□□□□□□□□□□□□□□□□□□□

```sql
SELECT * FROM concept WHERE LOWER(concept_class_id) = 'clinical finding'
```

□□□□□□□□□□□

SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□cdm_data.dbo.pe
□ cdm_data □□□□□□□□ dbo □□□□□□ person □□□□□□□□□□□□□□□□□□□□□□□□□
Server□□□□□□□□□□□□□□□1□□□□□□□□□□□ dbo □□□□□□□□□□□□□□□□□□□□□□□□
PostgreSQL□□□SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□SQL Server□□
@databaseSchema □□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□

```
SELECT * FROM @databaseSchema.person
```

SQL Server□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□databaseSchema
= "cdm_data.dbo"□□
= "cdm_data"□

□□□□□□□□□□□□□□□USE □□□□□□□USE cdm_data.dbo; □□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□**SQL**□□□□□

□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□

□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□Shiny□□□□
SqlRender □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
launchSqlRenderDeveloper()
```

□□□□□□□??□□

“‘{r sqlDeveloper, fig.cap=‘The SqlDeveloper’

9.2 DatabaseConnector

DatabaseConnector□□Java□JDBC□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
R Archive Network□□□□□□□□□□□□□□□□□□□□□□□□

```
install.packages("DatabaseConnector")
```

DatabaseConnector□□□□□□□□□□□□□□□□PostgreSQL□Microsoft SQL Server□SQLite□□□□Oracle□□□□□□□□□□□□□□Microsoft APS□IBM Netezza□□□□Amazon□□□□□□□□□□□□□□□□□□□□□□□Hadoop□□□□Impala□□□□Google BigQuery□□

9.2.1

□□□

```r
conn <- connect(dbms = "postgresql",
                server = "localhost/postgres",
                user = "joe",
                password = "secret",
                schema = "cdm")
```

```
## Connecting using PostgreSQL driver
```

□□□□□□□□□□□□□□□□□□□□□?connect□□□□□□□□□□□□□□□□□□□□□□□□

```r
disconnect(conn)
```

□□□□□□□□□□□□□□JDBC□□□□□□□□□□□□□□□□□□

```r
connString <- "jdbc:postgresql://localhost:5432/postgres"
conn <- connect(dbms = "postgresql",
                connectionString = connString,
                user = "joe",
                password = "secret",
                schema = "cdm")
```

```
## Connecting using PostgreSQL driver
```

□□□

```r
details <- createConnectionDetails(dbms = "postgresql",
                                   server = "localhost/postgres",
                                   user = "joe",
                                   password = "secret",
                                   schema = "cdm")
conn <- connect(details)
```

```
## Connecting using PostgreSQL driver
```

9.2.2

□□□□□□□□□□□□□□□□□□□□□□□□querySql□executeSql□□□querySql□□□□□□

□□□□□□□

```
querySql(conn, "SELECT TOP 3 * FROM person")
```

```
##   person_id gender_concept_id year_of_birth
## 1         1              8507          1975
## 2         2              8507          1976
## 3         3              8507          1977
```

```
executeSql(conn, "TRUNCATE TABLE foo; DROP TABLE foo;")
```

□□

9.2.3 ffdf

□□8.4.2□□□□□□□

```
x <- querySql.ffdf(conn, "SELECT * FROM person")
```

x□□□ffdf□□□□□□□□□

9.2.4 SQL

SqlRender□□□□□□render□□□translate□□□□□□□□□□□□□□□□□□□□□renderTransla

```
x <- renderTranslateQuerySql(conn,
                             sql = "SELECT TOP 10 * FROM @schema.person",
                             schema = "cdm_synpuf")
```

SQL Server□□□□TOP 10□□□□□PostgreSQL□□□LIMIT 10□□□□□□□□□SQL□□□□□□@sche

9.2.5

□□□□□□□□□□□□□□□□□executeSql□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□

```
data(mtcars)
insertTable(conn, "mtcars", mtcars, createTable = TRUE)
```

□□□□□□mtcars□□□□□□□□□□□□□□□mtcars□□□□□□□□□□□□□□□□□□□□□□□□□□

9.3 CDM

□□□□□□□OHDSI SQL□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□C

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```sql
SELECT COUNT(*) AS person_count FROM @cdm.person;
```

PERSON_COUNT
26299001

□□□□□□□□□□□□□□□□□□□□□□□□□□

```sql
SELECT AVG(DATEDIFF(DAY,
                    observation_period_start_date,
                    observation_period_end_date) / 365.25) AS num_years
FROM @cdm.observation_period;
```

NUM_YEARS
1.980803

□□□

```sql
SELECT MAX(YEAR(observation_period_end_date) -
           year_of_birth) AS max_age
FROM @cdm.person
INNER JOIN @cdm.observation_period
  ON person.person_id = observation_period.person_id;
```

MAX_AGE
90

□□PERSON
`...` AS□□□□□□“ages”□□□□□□□□□□□□ages□□□□□□□□□□□□□□□□□□□□□□□□□
`< .50 * n`□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```sql
WITH ages
AS (
    SELECT age,
        ROW_NUMBER() OVER (
            ORDER BY age
            ) order_nr
    FROM (
        SELECT YEAR(observation_period_start_date) - year_of_birth AS age
        FROM @cdm.person
```

```
        INNER JOIN @cdm.observation_period
            ON person.person_id = observation_period.person_id
        ) age_computed
    )
SELECT MIN(age) AS min_age,
    MIN(CASE
            WHEN order_nr < .25 * n
                THEN 9999
            ELSE age
            END) AS q25_age,
    MIN(CASE
            WHEN order_nr < .50 * n
                THEN 9999
            ELSE age
            END) AS median_age,
    MIN(CASE
            WHEN order_nr < .75 * n
                THEN 9999
            ELSE age
            END) AS q75_age,
    MAX(age) AS max_age
FROM ages
CROSS JOIN (
    SELECT COUNT(*) AS n
    FROM ages
    ) population_size;
```

MIN_AGE	Q25_AGE	MEDIAN_AGE	Q75_AGE	MAX_AGE
0	6	17	34	90

□□□□□□□□□SQL□□□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□R□□□□□

```
sql <- "SELECT YEAR(observation_period_start_date) -
               year_of_birth AS age
FROM @cdm.person
INNER JOIN @cdm.observation_period
  ON person.person_id = observation_period.person_id;"
age <- renderTranslateQuerySql(conn, sql, cdm = "cdm")
quantile(age[, 1], c(0, 0.25, 0.5, 0.75, 1))
```

```
##    0%   25%   50%   75%  100%
##     0     6    17    34    90
```

□□

□□□□□CDM□□□

```sql
SELECT TOP 10 condition_source_value,
  COUNT(*) AS code_count
FROM @cdm.condition_occurrence
GROUP BY condition_source_value
ORDER BY -COUNT(*);
```

CONDITION_SOURCE_VALUE	CODE_COUNT
4019	49094668
25000	36149139
78099	28908399
319	25798284
31401	22547122
317	22453999
311	19626574
496	19570098
I10	19453451
3180	18973883

□□□□□CONDITION_OCCURRENCE□□□□□□CONDITION_SOURCE_VALUE□□□□□□

9.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□S

```sql
SELECT COUNT(*) AS subject_count,
  concept_name
FROM @cdm.person
INNER JOIN @cdm.concept
  ON person.gender_concept_id = concept.concept_id
GROUP BY concept_name;
```

SUBJECT_COUNT	CONCEPT_NAME
14927548	FEMALE
11371453	MALE

□□□

```sql
SELECT COUNT(*) AS prescription_count
FROM @cdm.drug_exposure
INNER JOIN @cdm.concept_ancestor
```

```
  ON drug_concept_id = descendant_concept_id
INNER JOIN @cdm.concept ingredient
  ON ancestor_concept_id = ingredient.concept_id
WHERE LOWER(ingredient.concept_name) = 'ibuprofen'
  AND ingredient.concept_class_id = 'Ingredient'
  AND ingredient.standard_concept = 'S';
```

PRESCRIPTION_COUNT
26871214

9.5 QueryLibrary

QueryLibrary□□CDM□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□[1]□□□□□

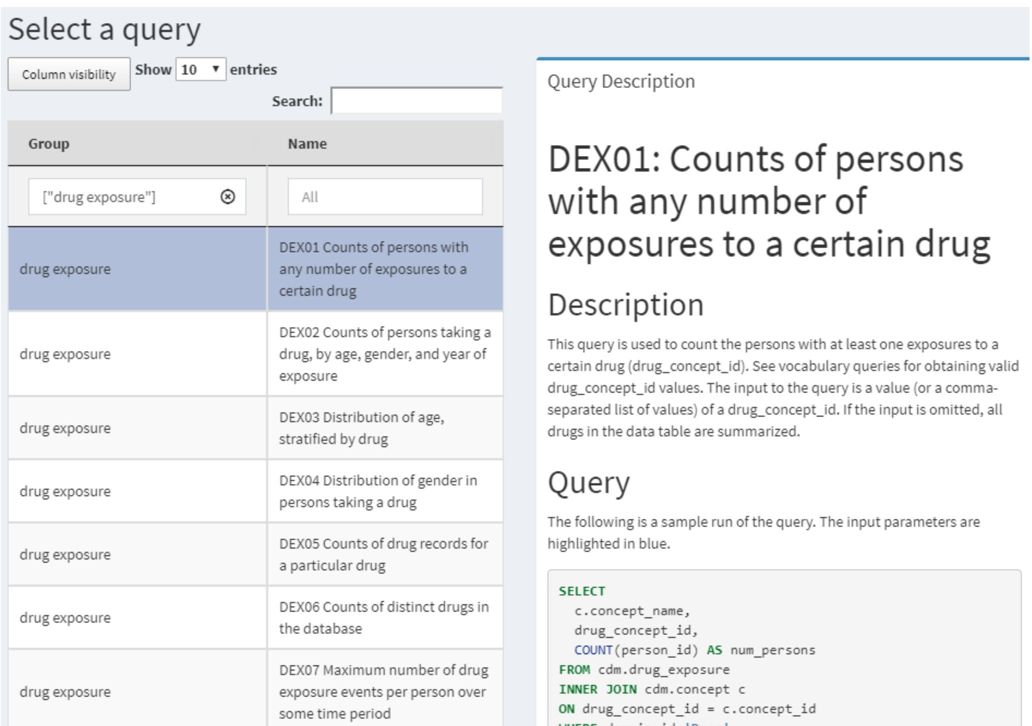


Figure 9.1: □□□□□□□□□CDM□□□□SQL□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

QueryLibrary□□SqlRender□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□CDM□□

[1]http://data.ohdsi.org/QueryLibrary
[2]https://github.com/OHDSI/QueryLibrary

9.6

9.6.1

□□□□□□ACE□□□□ACEi□□□□□□□□□□□□□□Slater et al. (1988) □□□□□ACEi□□□□

□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

9.6.2

□□□

9.6.3

□□□□□□□□□□□□□□□□□□□ER□□□□□□□□□□□□□□□□□□□□□□□□□□

9.6.4

□□□□□□□□□1□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□

9.7 SQL R

OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□

```r
library(DatabaseConnector)
conn <- connect(dbms = "postgresql",
                server = "localhost/postgres",
                user = "joe",
                password = "secret")
cdmDbSchema <- "cdm"
cohortDbSchema <- "scratch"
cohortTable <- "my_cohorts"

sql <- "
CREATE TABLE @cohort_db_schema.@cohort_table (
  cohort_definition_id INT,
  cohort_start_date DATE,
  cohort_end_date DATE,
  subject_id BIGINT
);
"
renderTranslateExecuteSql(conn, sql,
                          cohort_db_schema = cohortDbSchema,
                          cohort_table = cohortTable)
```

□□□

9.7.1

□□□□□□□□□□□□□□COHORT□□□□□□□□□□□

```
sql <- "
INSERT INTO @cohort_db_schema.@cohort_table (
  cohort_definition_id,
  cohort_start_date,
  cohort_end_date,
  subject_id
)
SELECT 1 AS cohort_definition_id,
  cohort_start_date,
  cohort_end_date,
  subject_id
FROM (
  SELECT drug_era_start_date AS cohort_start_date,
    drug_era_end_date AS cohort_end_date,
    person_id AS subject_id
  FROM (
    SELECT drug_era_start_date,
      drug_era_end_date,
      person_id,
      ROW_NUMBER() OVER (
        PARTITION BY person_id
            ORDER BY drug_era_start_date
      ) order_nr
    FROM @cdm_db_schema.drug_era
    WHERE drug_concept_id = 1308216 --
  ) ordered_exposures
  WHERE order_nr = 1
) first_era
INNER JOIN @cdm_db_schema.observation_period
  ON subject_id = person_id
    AND observation_period_start_date < cohort_start_date
    AND observation_period_end_date > cohort_start_date
WHERE DATEDIFF(DAY,
               observation_period_start_date,
               cohort_start_date) >= 365;
"

renderTranslateExecuteSql(conn, sql,
                          cohort_db_schema = cohortDbSchema,
                          cohort_table = cohortTable,
                          cdm_db_schema = cdmDbSchema)
```

□□□□□CDM□□□□□□□□□□DRUG_ERA□□□□□□□□□□□□□□□□□□DRUG_EXPOSURE

9.7.2

□□□□□□□□□□□□□□□□□□□□□□□□□□

```
sql <- "
INSERT INTO @cohort_db_schema.@cohort_table (
 cohort_definition_id,
 cohort_start_date,
 cohort_end_date,
subject_id
)
SELECT 2 AS cohort_definition_id,
  cohort_start_date,
  cohort_end_date,
  subject_id
FROM (
  SELECT DISTINCT person_id AS subject_id,
    condition_start_date AS cohort_start_date,
    condition_end_date AS cohort_end_date
  FROM @cdm_db_schema.condition_occurrence
  INNER JOIN @cdm_db_schema.concept_ancestor
    ON condition_concept_id = descendant_concept_id
  WHERE ancestor_concept_id = 432791 --
) distinct_occurrence
INNER JOIN @cdm_db_schema.visit_occurrence
  ON subject_id = person_id
  AND visit_start_date <= cohort_start_date
  AND visit_end_date >= cohort_start_date
WHERE visit_concept_id IN (262, 9203,
    9201) --    ER;
"

renderTranslateExecuteSql(conn, sql,
                          cohort_db_schema = cohortDbSchema,
                          cohort_table = cohortTable,
                          cdm_db_schema = cdmDbSchema)
```

□□□□□CONDITION_OCCURRENCE□□□□□CONCEPT_ANCESTOR□□□□□□□□□□□□

9.7.3

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
sql <- "
WITH tar AS (
  SELECT concept_name AS gender,
    FLOOR((YEAR(cohort_start_date) -
          year_of_birth) / 10) AS age,
```

```
    subject_id,
    cohort_start_date,
    CASE WHEN DATEADD(DAY, 7, cohort_start_date) >
      observation_period_end_date
    THEN observation_period_end_date
    ELSE DATEADD(DAY, 7, cohort_start_date)
    END AS cohort_end_date
  FROM @cohort_db_schema.@cohort_table
  INNER JOIN @cdm_db_schema.observation_period
    ON subject_id = observation_period.person_id
      AND observation_period_start_date < cohort_start_date
      AND observation_period_end_date > cohort_start_date
  INNER JOIN @cdm_db_schema.person
    ON subject_id = person.person_id
  INNER JOIN @cdm_db_schema.concept
    ON gender_concept_id = concept_id
  WHERE cohort_definition_id = 1 --
)
SELECT days.gender,
    days.age,
    days,
    CASE WHEN events IS NULL THEN 0 ELSE events END AS events
FROM (
  SELECT gender,
    age,
    SUM(DATEDIFF(DAY, cohort_start_date,
      cohort_end_date)) AS days
  FROM tar
  GROUP BY gender,
    age
) days
LEFT JOIN (
  SELECT gender,
      age,
      COUNT(*) AS events
  FROM tar
  INNER JOIN @cohort_db_schema.@cohort_table angioedema
    ON tar.subject_id = angioedema.subject_id
      AND tar.cohort_start_date <= angioedema.cohort_start_date
      AND tar.cohort_end_date >= angioedema.cohort_start_date
  WHERE cohort_definition_id = 2 --
  GROUP BY gender,
    age
) events
ON days.gender = events.gender
  AND days.age = events.age;
"

results <- renderTranslateQuerySql(conn, sql,
```

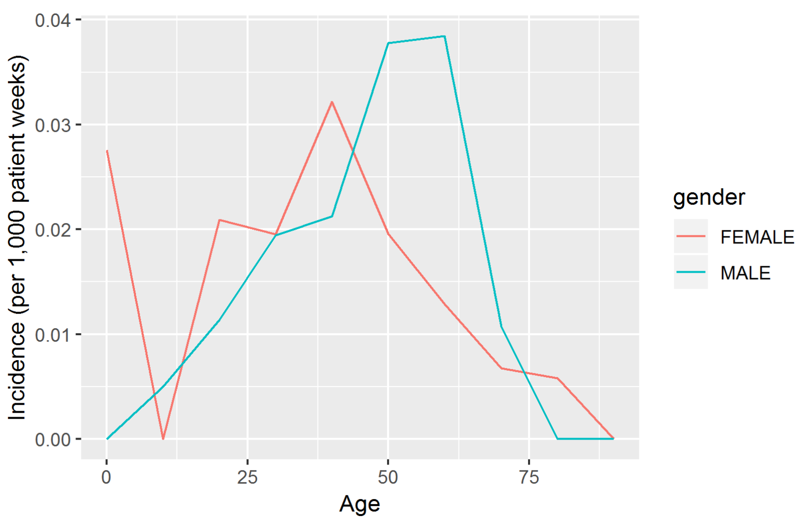
```
                                 cohort_db_schema = cohortDbSchema,
                                 cohort_table = cohortTable,
                                 cdm_db_schema = cdmDbSchema,
                                 snakeCaseToCamelCase = TRUE)
```

□□□CTE□tar□□□□□□□□□□□□□□□□□□□□□□□□□□□□OBSERVATION_PERIOD_

SQL□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□R□□□□□

ggplot2□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
#   IR
results$ir <- 1000 * results$events / results$days / 7

#
results$age <- results$age * 10

library(ggplot2)
ggplot(results, aes(x = age, y = ir, group = gender, color = gender)) +
  geom_line() +
  xlab("  ") +
  ylab("  1,000     ")
```



9.7.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
sql <- "
TRUNCATE TABLE @cohort_db_schema.@cohort_table;
DROP TABLE @cohort_db_schema.@cohort_table;
```

```
"
renderTranslateExecuteSql(conn, sql,
                          cohort_db_schema = cohortDbSchema,
                          cohort_table = cohortTable)

disconnect(conn)
```

9.7.5

OHDSI SQL□DatabaseConnector□□□SqlRender□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHD

□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□
SQL□□□□□SqlRender□□□DatabaseConnector□□□□□□□□□□□□□□

9.8

- **SQL**□Structured Query Language□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□
- **SqlRender**□□□**DatabaseConnector**R□□□□□□□CDM□□□□□□□□□□□□□□□□□□□
- R□SQL□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□
- **QueryLibrary**□□CDM□□□□□□□□SQL□□□□□□□□□□□□□□□□

9.9

□□□□

□□□□□□□□□□□□□□□□ 8.4.5 □□□□□□□□□□□R□R-Studio□Java
□□□□□□□□□□□□□□□□□□□□□□□□□SqlRender□DatabaseConnector□□□□
Eunomia □□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
```

Eunomia □□□□□□□CDM □□□□□□ R □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM □□□□□□□□□□□□□main□□□□

□□ **9.1.** SQL □□□ R □□□□□□□□□□□□□□□□□□□□□□□□□□□

□□**9.2.** SQL □□□R □□□□□□□□□□□1 □□□□□□□□□□□□□□□□□□□□□□□□□□□

□□**9.3.** SQL □□□R □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□ID □ 192671 □□)

□□□□□□□□□□E.5 □□□□□□

Chapter 10

10 – –

□…: *Kristin Kostka!*

□…

□…

□…ATLAS□…SQL□…

10.1

OHDSI□…

□…

OHDSI□…
9/ICD-10□NDC□HCPCS□…
9/ICD-10□… □…

OHDSI□…:

- □…
- □…
- □…
- □…

□…

1. □… □…
2. □… □…0□…100%□…

□…

10.2

□

□

- □CDM□
- □
- □DRUG_EXPOSURE□DAYS_SUPP
- □365□

□10.1□

Figure 10.1: □

□

- □
- □
- □

□

□

□

OHDSI□

10.3

□□□

- □□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□??□□□□

□□□#tab:conceptSetExpression□□□□□□□□□□□□□□□

□□□□□ID	□□□□□□	□□	□□	□□□□□□□
4329847	□□□□	□□□	□□	□□□
314666	□□□□□□	□□	□□	□□□

□□□□□□□□□10.2□□
9□□□ICD-10□□□□□□□□□□□

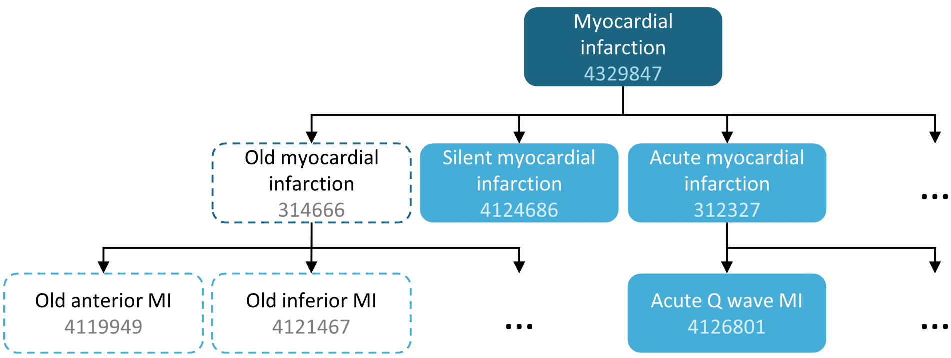


Figure 10.2: □□

10.4

□□□

□□□□□□□□CDM□□□□□□□□□□APHRODITE□□□□□□□□□□□□□□□□□□□□□□
(Banda et al., 2017) □

10.5

□□
□□□□ □□□□□□□□□□□□□□□□□□□□□□

[1] https://github.com/OHDSI/Aphrodite

□□□

□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 16□□□□□□□□□□□□□□

10.5.1 OHDSI

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□APHRODITE (Banda et al., 2017) □□□PheValuator□□□ (Swerdel et al., 2019) □□□OHDSI□□□□□□□□□□□□□□□□□□□□□ eMERGE Phenotype Library □□□□□□□□□□□□□□□ (Hripcsak et al., 2019)□□□□□□□□□□□□□

10.6

□□□

□□□
10.3 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

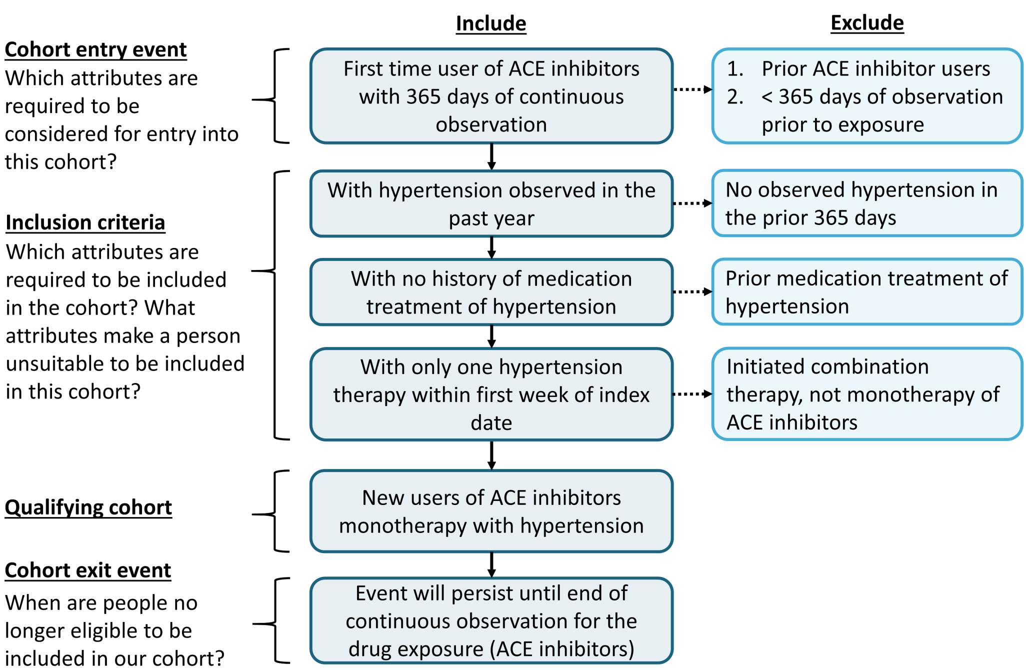


Figure 10.3: □□□□□□□□□□□□□□

□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□

[2] https://www.ohdsi.org/web/wiki/doku.php?id=projects:workgroups:gold-library-wg

10.7 ATLAS

□□ATLAS□□□□□□□ Cohort Definitions □□□□□□□□□□□□□□□□□□□□□□□□□□□
cohort□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10.4□□□□□□□□□□□□

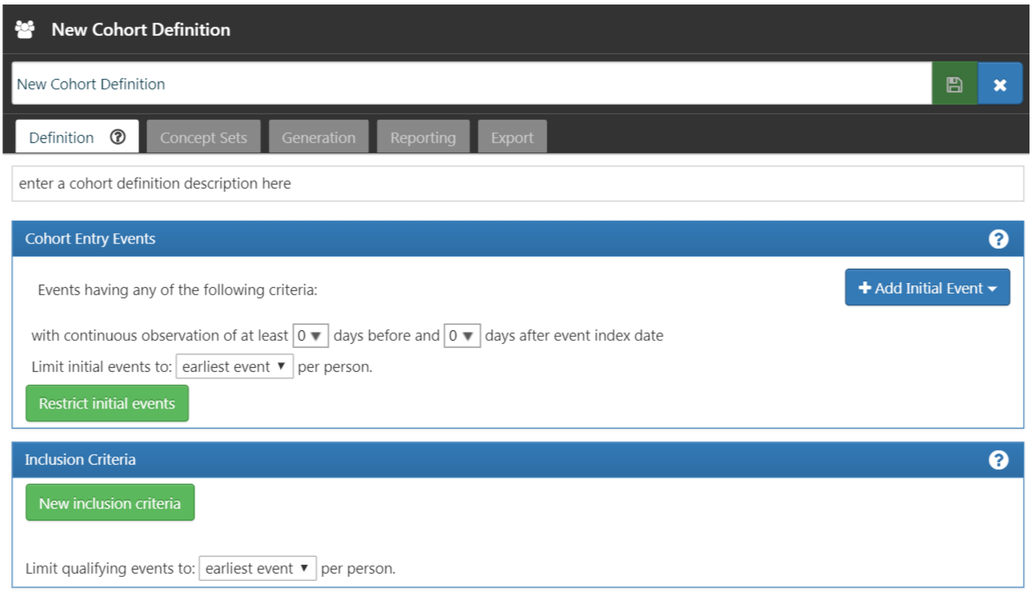


Figure 10.4: □□□□□□□□□

□□□□□□□New Cohort Definition□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□

□□□□□□□□ □□□□□□□□□□□□□□□□□□

10.7.1

□□□□□□□□□□□□□□□□□□□□□□□Add initial event□□□□□□□□□□□□□□□□□□□

□10.5□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□CONDITION_OCCURREN
Drug Exposure□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□10.6□□□□□□□ATLAS

10.7.2

□□□□□□□□□□□□□□□□□□ □□□□□□□□ACE□□□□□□□□□□□□□□□□□□□□□

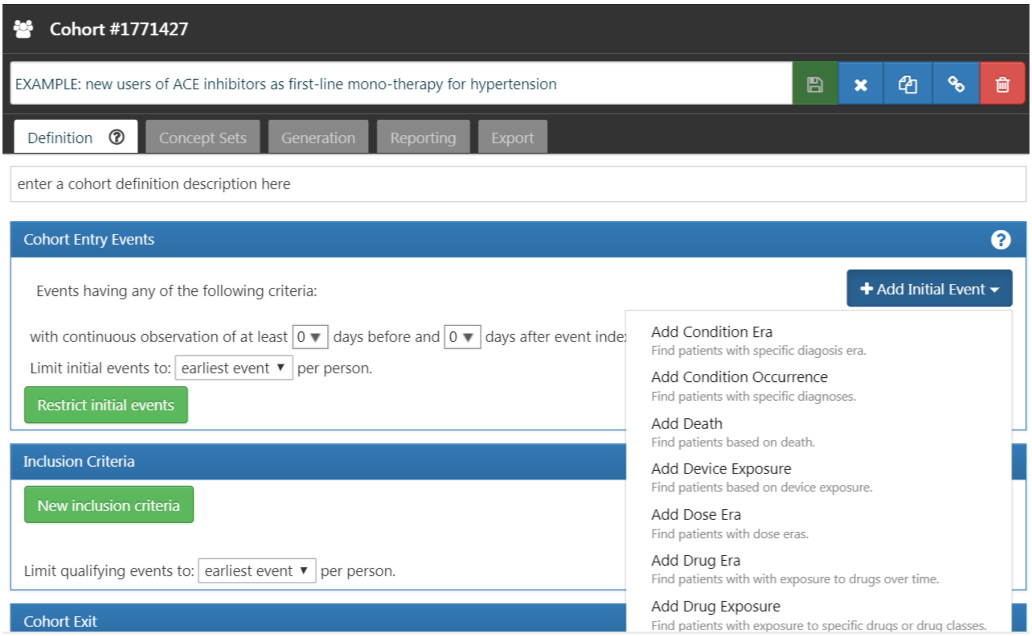


Figure 10.5: □□□□□□□□□

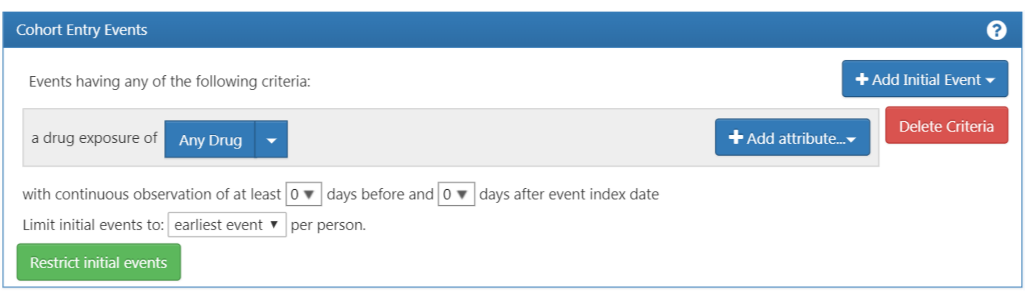


Figure 10.6: □□□□□□□

□□□□**1:** □□□□□□□□□□□□□□□□□□□□□□

□□□ set□□□□□□□□□New Concept Set□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Un Concept Set□□□□□□□□□□□□□□□□□□□□□□□□□ Search □□□□□□□□□□ACE□

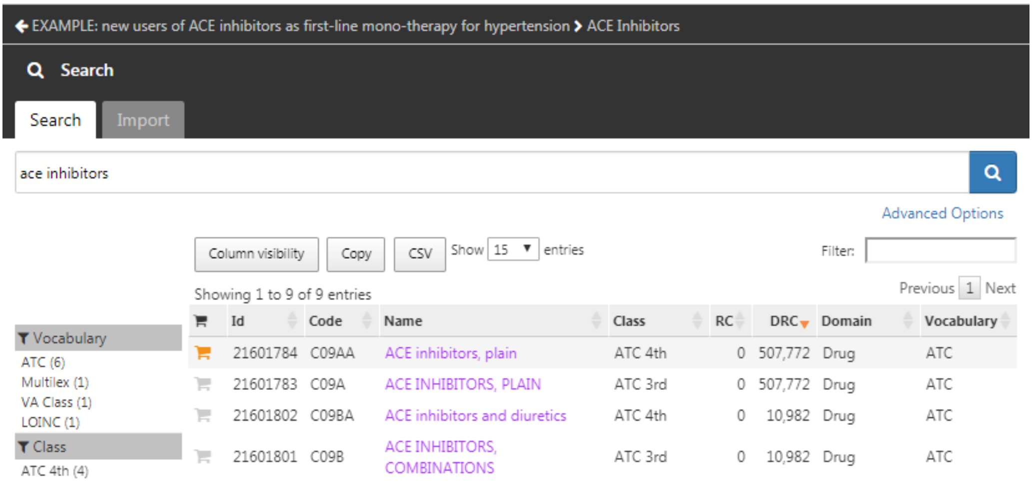


Figure 10.7: □□□□□ - ACE□□□

□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□10.7□□□□□□□□□□□□□

□10.8□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ACE□□□□□□□□□□□□□

□□□□**2:** □□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□Import Concept Set□□□□□□□□□ inhibitors□□□ □□Any Drug□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

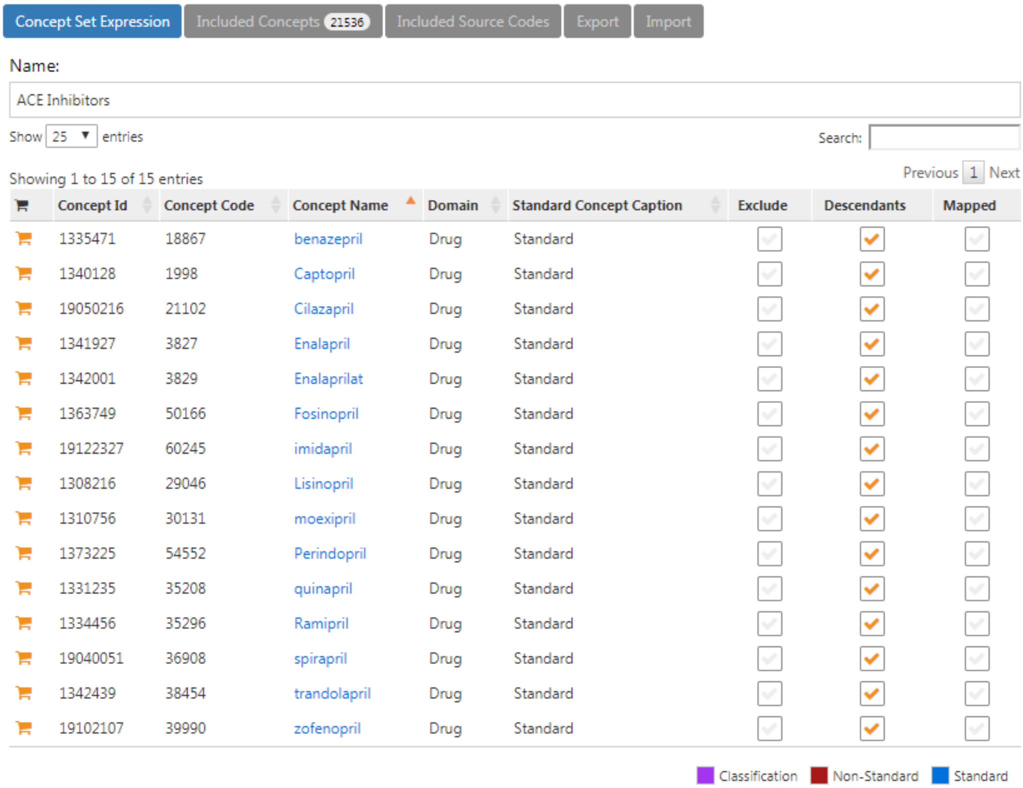
10.7.3

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ACE□□□□□□□□□□□□□□□ attribute□□□□□□□□□□□□Add first exposure criteria□□□□□□□□□□□□□□□□□□□□

□□□10.10

□□□ATLAS□□No□□□□

□□ event□□□□□□□□□□10.10□□□□□□□□□□□□□□□□□□□□



Concept Set Expression | Included Concepts 21536 | Included Source Codes | Export | Import

Name:

ACE Inhibitors

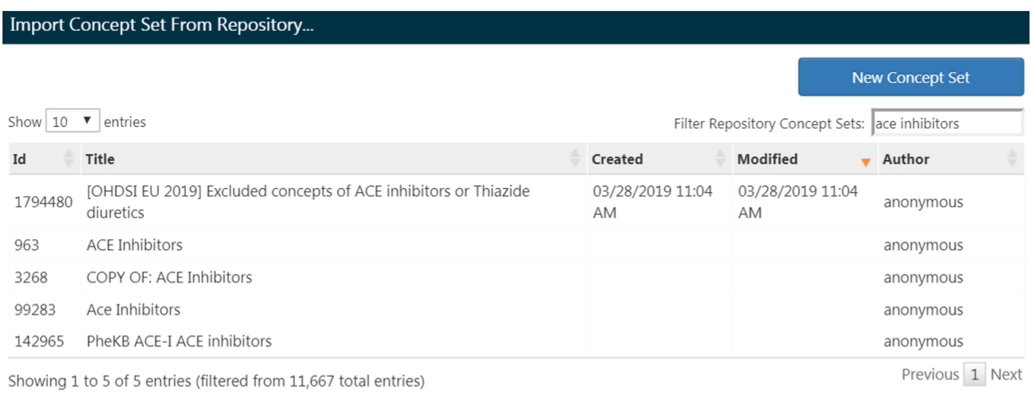
Show 25 entries Search:

Showing 1 to 15 of 15 entries Previous 1 Next

	Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
	1335471	18867	benazepril	Drug	Standard		✓	
	1340128	1998	Captopril	Drug	Standard		✓	
	19050216	21102	Cilazapril	Drug	Standard		✓	
	1341927	3827	Enalapril	Drug	Standard		✓	
	1342001	3829	Enalaprilat	Drug	Standard		✓	
	1363749	50166	Fosinopril	Drug	Standard		✓	
	19122327	60245	imidapril	Drug	Standard		✓	
	1308216	29046	Lisinopril	Drug	Standard		✓	
	1310756	30131	moexipril	Drug	Standard		✓	
	1373225	54552	Perindopril	Drug	Standard		✓	
	1331235	35208	quinapril	Drug	Standard		✓	
	1334456	35296	Ramipril	Drug	Standard		✓	
	19040051	36908	spirapril	Drug	Standard		✓	
	1342439	38454	trandolapril	Drug	Standard		✓	
	19102107	39990	zofenopril	Drug	Standard		✓	

Classification Non-Standard Standard

Figure 10.8: ACE□□□□□□□□□□□□□□□



Import Concept Set From Repository...

New Concept Set

Show 10 entries Filter Repository Concept Sets: ace inhibitors

Id	Title	Created	Modified	Author
1794480	[OHDSI EU 2019] Excluded concepts of ACE inhibitors or Thiazide diuretics	03/28/2019 11:04 AM	03/28/2019 11:04 AM	anonymous
963	ACE Inhibitors			anonymous
3268	COPY OF: ACE Inhibitors			anonymous
99283	Ace Inhibitors			anonymous
142965	PheKB ACE-I ACE inhibitors			anonymous

Showing 1 to 5 of 5 entries (filtered from 11,667 total entries) Previous 1 Next

Figure 10.9: ATLAS□□□□□□□□□□□□□□□□□□□□□□

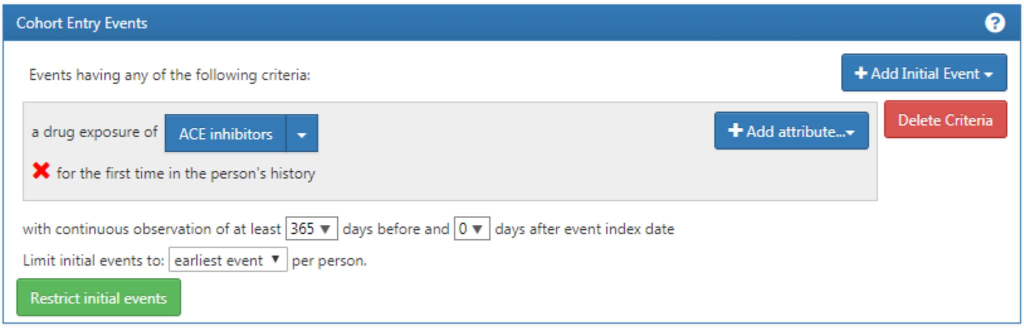


Figure 10.10: □□□□□□□□□□□□□□□□□□□□□□□□

□□

□10.11□□□

10.7.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Restrict initial events□□□□□New inclusion criteria□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATL initial events□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□ inclusion criteria□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□New inclusion criteria□□□ inclusion criteria□□□

□□□□□□□□□□□□□□□□+Add criteria to group□□□□□□□□□□□□□□□□□□□□□□□□ Initial Event□□ criteria to group□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CONDITION_O condition occurrence□□□□□□□

10.8 SQL

□□□□□SQL□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□9□□□□□□□□□□OH

□□□□□□□□□□□SQL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

10.8.1

□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DatabaseConnector□□□□□□

```r
library(CohortMethod)
connDetails <- createConnectionDetails(dbms = "postgresql",
                                       server = "localhost/ohdsi",
                                       user = "joe",
```

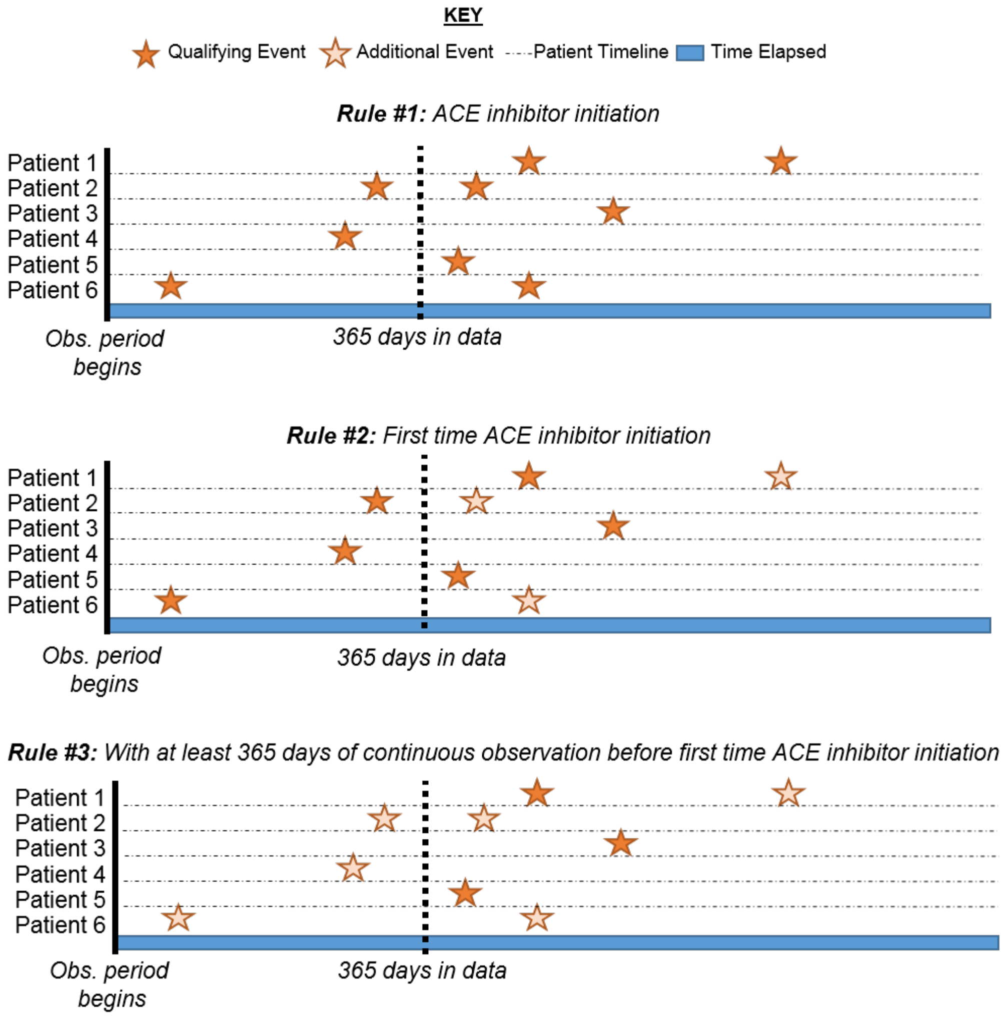

KEY
Qualifying Event
Additional Event
Patient Timeline
Time Elapsed
Rule #1: ACE inhibitor initiation
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Obs. period begins
365 days in data
Rule #2: First time ACE inhibitor initiation
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Obs. period begins
365 days in data
Rule #3: With at least 365 days of continuous observation before first time ACE inhibitor initiation
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Obs. period begins
365 days in data

Figure 10.11: □□□□□□□□□□□□□□□□□

```
                                                password = "supersecret")

cdmDbSchema <- "my_cdm_data"
cohortDbSchema <- "scratch"
cohortTable <- "my_cohorts"
```

□□□3□□□cdmDbSchema□cohortDbSchema□□□□cohortTable□□□□□□□□□□□□□□□
SQL Server□□□
<- "my_cdm_data.dbo"□

10.8.2

□□□□□□□□□□□□□□□□□□□□□ID□R□□□□□□□□□SQL□□□□□□

```
aceI <- c(1308216, 1310756, 1331235, 1334456, 1335471, 1340128, 1341927,
          1342439, 1363749, 1373225)

hypertension <- 316866

allHtDrugs <- c(904542, 907013, 932745, 942350, 956874, 970250, 974166,
                978555, 991382, 1305447, 1307046, 1307863, 1308216,
                1308842, 1309068, 1309799, 1310756, 1313200, 1314002,
                1314577, 1317640, 1317967, 1318137, 1318853, 1319880,
                1319998, 1322081, 1326012, 1327978, 1328165, 1331235,
                1332418, 1334456, 1335471, 1338005, 1340128, 1341238,
                1341927, 1342439, 1344965, 1345858, 1346686, 1346823,
                1347384, 1350489, 1351557, 1353766, 1353776, 1363053,
                1363749, 1367500, 1373225, 1373928, 1386957, 1395058,
                1398937, 40226742, 40235485)
```

10.8.3

□□□□□□□ACE□□□□□□□□□□□□□□□□□□

```
conn <- connect(connDetails)

sql <- "SELECT person_id AS subject_id,
  MIN(drug_exposure_start_date) AS cohort_start_date
INTO #first_use
FROM @cdm_db_schema.drug_exposure
INNER JOIN @cdm_db_schema.concept_ancestor
  ON descendant_concept_id = drug_concept_id
WHERE ancestor_concept_id IN (@ace_i)
GROUP BY person_id;"

renderTranslateExecuteSql(conn,
```

```
                          sql,
                          cdm_db_schema = cdmDbSchema,
                          ace_i = aceI)
```

DRUG_EXPOSURE□□□□□CONCEPT_ANCESTOR□□□□□□□□□□□□□ACE□□□□□□□□□

10.8.4 365

□□□OBSERVATION_PERIOD□□□□□□□□□365□□□□□□□□□□□□□□□□□□

```
sql <- "SELECT subject_id,
  cohort_start_date
INTO #has_prior_obs
FROM #first_use
INNER JOIN @cdm_db_schema.observation_period
  ON subject_id = person_id
    AND observation_period_start_date <= cohort_start_date
    AND observation_period_end_date >= cohort_start_date
WHERE DATEADD(DAY, 365, observation_period_start_date) < cohort_start_date;"

renderTranslateExecuteSql(conn, sql, cdm_db_schema = cdmDbSchema)
```

10.8.5

365□□□□□□□□□□□□□□□

```
sql <- "SELECT DISTINCT subject_id,
  cohort_start_date
INTO #has_ht
FROM #has_prior_obs
INNER JOIN @cdm_db_schema.condition_occurrence
  ON subject_id = person_id
    AND condition_start_date <= cohort_start_date
    AND condition_start_date >= DATEADD(DAY, -365, cohort_start_date)
INNER JOIN @cdm_db_schema.concept_ancestor
  ON descendant_concept_id = condition_concept_id
WHERE ancestor_concept_id = @hypertension;"

renderTranslateExecuteSql(conn,
                          sql,
                          cdm_db_schema = cdmDbSchema,
                          hypertension = hypertension)
```

□□□SELECT DISTINCT□□□□□□□□□□□□□□□□□□□

10.8.6

□□□□□□□□□□□□□□□□□□□□

```
sql <- "SELECT subject_id,
  cohort_start_date
INTO #no_prior_ht_drugs
FROM #has_ht
LEFT JOIN (
  SELECT *
  FROM @cdm_db_schema.drug_exposure
  INNER JOIN @cdm_db_schema.concept_ancestor
    ON descendant_concept_id = drug_concept_id
  WHERE ancestor_concept_id IN (@all_ht_drugs)
) ht_drugs
  ON subject_id = person_id
    AND drug_exposure_start_date < cohort_start_date
WHERE person_id IS NULL;"

renderTranslateExecuteSql(conn,
                          sql,
                          cdm_db_schema = cdmDbSchema,
                          all_ht_drugs = allHtDrugs)
```

LEFT JOIN□□□□□□□□DRUG_EXPOSURE□□□□□□□person_id□NULL□□□□□□□□□□

10.8.7

□□□□□□□□□□□□□7□□□□□□□□□□□□□□□□□□□□□□□

```
sql <- "SELECT subject_id,
  cohort_start_date
INTO #monotherapy
FROM #no_prior_ht_drugs
INNER JOIN @cdm_db_schema.drug_exposure
  ON subject_id = person_id
    AND drug_exposure_start_date >= cohort_start_date
    AND drug_exposure_start_date <= DATEADD(DAY, 7, cohort_start_date)
INNER JOIN @cdm_db_schema.concept_ancestor
  ON descendant_concept_id = drug_concept_id
WHERE ancestor_concept_id IN (@all_ht_drugs)
GROUP BY subject_id,
  cohort_start_date
HAVING COUNT(*) = 1;"

renderTranslateExecuteSql(conn,
                          sql,
                          cdm_db_schema = cdmDbSchema,
```

```
                          all_ht_drugs = allHtDrugs)
```

10.8.8

[illegible]

```
sql <- "
  SELECT person_id,
    CAST(1 AS INT) AS concept_id,
    drug_exposure_start_date AS exposure_start_date,
    drug_exposure_end_date AS exposure_end_date
  INTO #exposure
  FROM @cdm_db_schema.drug_exposure
  INNER JOIN @cdm_db_schema.concept_ancestor
    ON descendant_concept_id = drug_concept_id
  WHERE ancestor_concept_id IN (@ace_i);"
renderTranslateExecuteSql(conn,
                          sql,
                          cdm_db_schema = cdmDbSchema,
                          ace_i = aceI)
```

[illegible]

```
sql <- "
SELECT ends.person_id AS subject_id,
    ends.concept_id AS cohort_definition_id,
  MIN(exposure_start_date) AS cohort_start_date,
  ends.era_end_date AS cohort_end_date
INTO #exposure_era
FROM (
  SELECT exposure.person_id,
    exposure.concept_id,
    exposure.exposure_start_date,
    MIN(events.end_date) AS era_end_date
  FROM #exposure exposure
  JOIN (
--cteEndDates
    SELECT person_id,
      concept_id,
      DATEADD(DAY, - 1 * @max_gap, event_date) AS end_date
    FROM (
      SELECT person_id,
        concept_id,
        event_date,
        event_type,
        MAX(start_ordinal) OVER (
```

```
          PARTITION BY person_id ,concept_id ORDER BY event_date,
              event_type ROWS UNBOUNDED PRECEDING
          ) AS start_ordinal,
        ROW_NUMBER() OVER (
          PARTITION BY person_id, concept_id ORDER BY event_date,
            event_type
          ) AS overall_ord
      FROM (
-- select the start dates, assigning a row number to each
        SELECT person_id,
          concept_id,
          exposure_start_date AS event_date,
          0 AS event_type,
          ROW_NUMBER() OVER (
            PARTITION BY person_id, concept_id ORDER BY exposure_start_date
            ) AS start_ordinal
        FROM #exposure exposure

        UNION ALL
-- add the end dates with NULL as the row number, padding the end dates by
-- @max_gap to allow a grace period for overlapping ranges.

        SELECT person_id,
          concept_id,
          DATEADD(day, @max_gap, exposure_end_date),
          1 AS event_type,
          NULL
        FROM #exposure exposure
        ) rawdata
    ) events
  WHERE 2 * events.start_ordinal - events.overall_ord = 0
  ) events
  ON exposure.person_id = events.person_id
      AND exposure.concept_id = events.concept_id
      AND events.end_date >= exposure.exposure_end_date
  GROUP BY exposure.person_id,
      exposure.concept_id,
      exposure.exposure_start_date
  ) ends
GROUP BY ends.person_id,
  concept_id,
  ends.era_end_date;"

renderTranslateExecuteSql(conn,
                          sql,
                          cdm_db_schema = cdmDbSchema,
                          max_gap = 30)
```

□□□□□□□□□□□□□□□□max_gap□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□ACE□□

```
sql <- "SELECT ee.subject_id,
  CAST(1 AS INT) AS cohort_definition_id,
  ee.cohort_start_date,
  ee.cohort_end_date
INTO @cohort_db_schema.@cohort_table
FROM #monotherapy mt
INNER JOIN #exposure_era ee
  ON mt.subject_id = ee.subject_id
    AND mt.cohort_start_date = ee.cohort_start_date;"

renderTranslateExecuteSql(conn,
                          sql,
                          cohort_db_schema = cohortDbSchema,
                          cohort_table = cohortTable)
```

□□1□□□□

10.8.9

□□□

```
sql <- "TRUNCATE TABLE #first_use;
DROP TABLE #first_use;

TRUNCATE TABLE #has_prior_obs;
DROP TABLE #has_prior_obs;

TRUNCATE TABLE #has_ht;
DROP TABLE #has_ht;

TRUNCATE TABLE #no_prior_ht_drugs;
DROP TABLE #no_prior_ht_drugs;

TRUNCATE TABLE #monotherapy;
DROP TABLE #monotherapy;

TRUNCATE TABLE #exposure;
DROP TABLE #exposure;

TRUNCATE TABLE #exposure_era;
DROP TABLE #exposure_era;"

renderTranslateExecuteSql(conn, sql)

disconnect(conn)
```

10.9

- □□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□
- □□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□
- □□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□ATLAS□□□SQL□□□□□□□□□□□

10.10

□□□□

□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□ http://atlas-demo.ohdsi.org □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ **10.1.** □□□□□□□□□ATLAS□□□□□□□□□□□□

- □□□□□□□□□□□□□□
- 16□□□
- □□□□□□□□□365□□□□□□□□□□□□□□
- □□□NSAID□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□
- □□□□□□□□□□□□□30□□□□□□□□□□□□□□□□□□□□

□□□□

□□□□□□□□R□R-Studio□□□Java□□□□□□□□□□□□□□□□□□□□□□□□□ 8.4.5□□□□□□□□□□□□SqlRender□DatabaseConnector□□□□Eunomia □□□□□□□□□□□

```
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
```

Eunomia□□□□□□□CDM□□□□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM□□□□□□□□□□□□main□□□□□

□□ **10.2.** SQL□□□R□□□□□□□□□COHORT□□□□□□□□□□□□AMI□□□□□□□□□

- □□□□□□□□□□□□□□□□4329847□□□□□□□□□□□□□□□□□□□□□□314666□

- □□□□□ER□□□□□□□□□9201□9203□262□□□□□□□□□□ □□□□□□□□□□□□□□□□

□□□□□□□□□Appendix E.6 □□□□□□□□□

Chapter 11

11 – –

: Anthony Sena & Daniel Prieto-Alhambra

11.1

- {}

15

11.2

OHDSI

the Reporting of Observation Studies in Epidemiology STROBE (

Elm et al., 2008)□

11.3

□□
et al. (2016) □□OHDSI□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□□□□□□□□

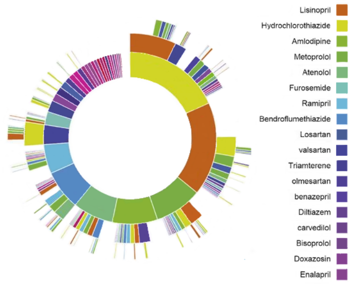


Figure 11.1: □□□□□□□OHDSI□□□□□□□□□□□□□□□

□□□□□□ 11.1 □□

□□
et al. (2016) □□□

□□□DUS□□□

11.4

□□
11.2 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

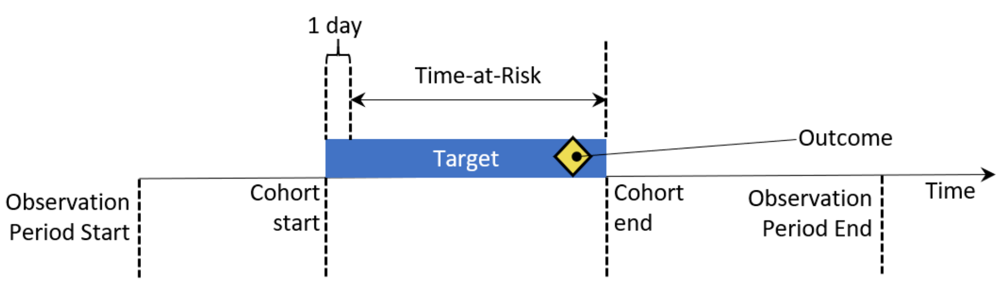


Figure 11.2: □□□

□ 11.2 □□□

□□□□□□□□□□□□2□□□□□□□□□□□

$$= \frac{\#}{\quad}$$

□□

$$= \frac{\#}{\quad}$$

□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DUS□□□□□□□□

11.5

□□□□□□□WHO□□□□□□□□□□□□□□□□(Who, 2013) □□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□WHO□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.6 ATLAS

□□□□□ACHILLES □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□
Data Sources □□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Occurrence□□□

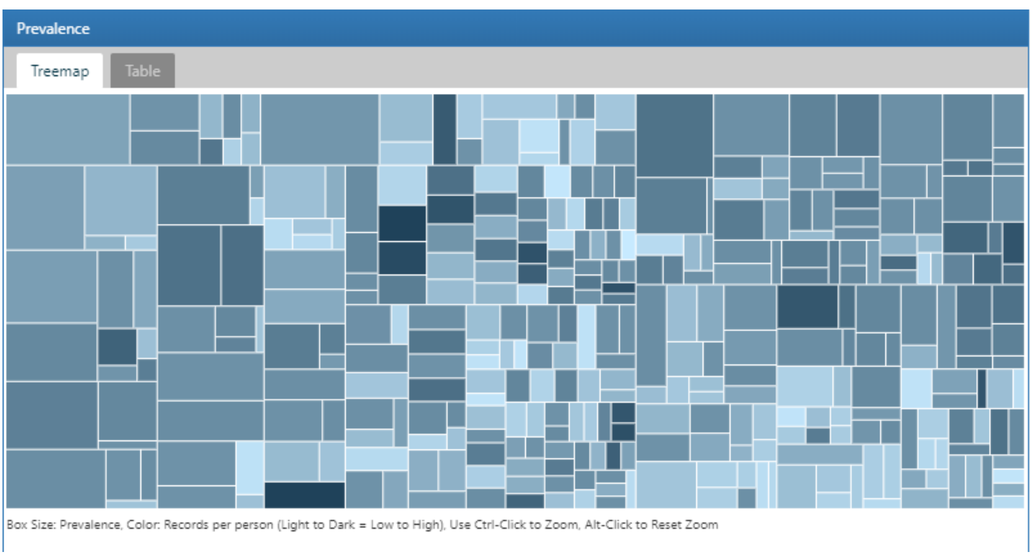
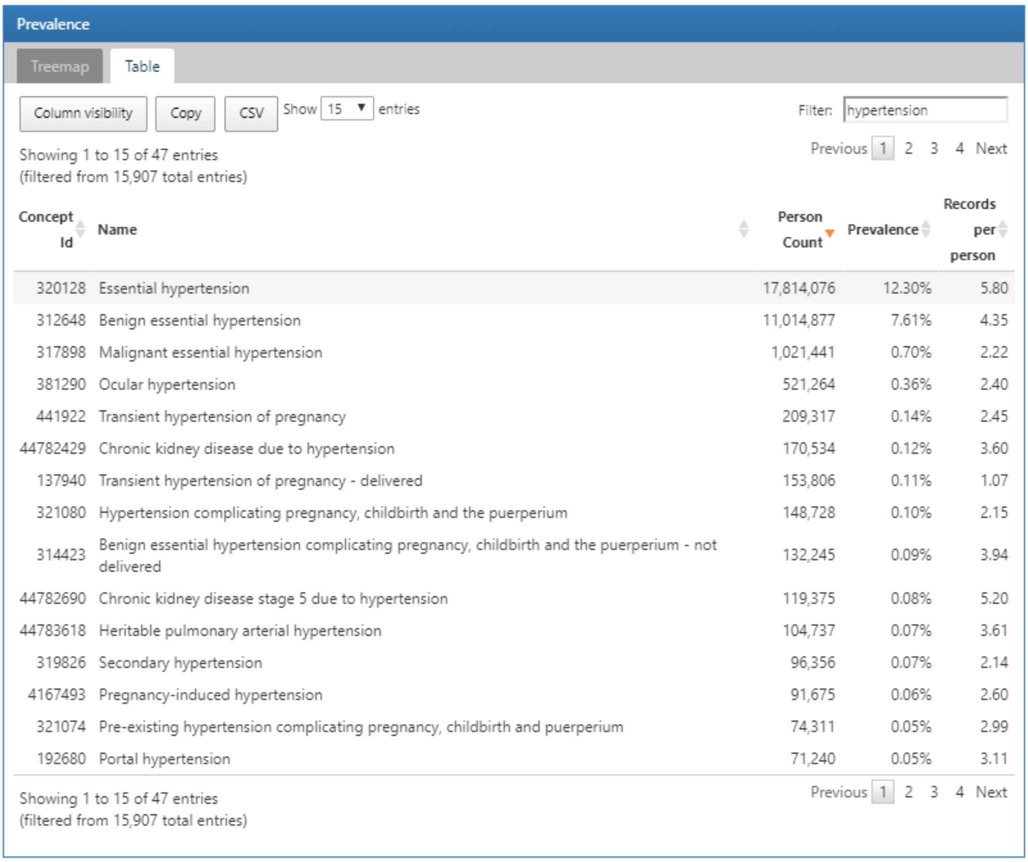


Figure 11.3: Atlas Data Sources: Condition Occurrence Treemap

□□


Prevalence

Treemap | Table

Column visibility | Copy | CSV | Show 15 entries | Filter: hypertension

Showing 1 to 15 of 47 entries
(filtered from 15,907 total entries)

Previous 1 2 3 4 Next

Concept Id	Name	Person Count	Prevalence	Records per person
320128	Essential hypertension	17,814,076	12.30%	5.80
312648	Benign essential hypertension	11,014,877	7.61%	4.35
317898	Malignant essential hypertension	1,021,441	0.70%	2.22
381290	Ocular hypertension	521,264	0.36%	2.40
441922	Transient hypertension of pregnancy	209,317	0.14%	2.45
44782429	Chronic kidney disease due to hypertension	170,534	0.12%	3.60
137940	Transient hypertension of pregnancy - delivered	153,806	0.11%	1.07
321080	Hypertension complicating pregnancy, childbirth and the puerperium	148,728	0.10%	2.15
314423	Benign essential hypertension complicating pregnancy, childbirth and the puerperium - not delivered	132,245	0.09%	3.94
44782690	Chronic kidney disease stage 5 due to hypertension	119,375	0.08%	5.20
44783618	Heritable pulmonary arterial hypertension	104,737	0.07%	3.61
319826	Secondary hypertension	96,356	0.07%	2.14
4167493	Pregnancy-induced hypertension	91,675	0.06%	2.60
321074	Pre-existing hypertension complicating pregnancy, childbirth and puerperium	74,311	0.05%	2.99
192680	Portal hypertension	71,240	0.05%	3.11

Showing 1 to 15 of 47 entries
(filtered from 15,907 total entries)

Previous 1 2 3 4 Next

Figure 11.4: ATLAS□□□□□□: □□□□□□□□□□□□□□□□□□□□

□□□essential hypertension□□

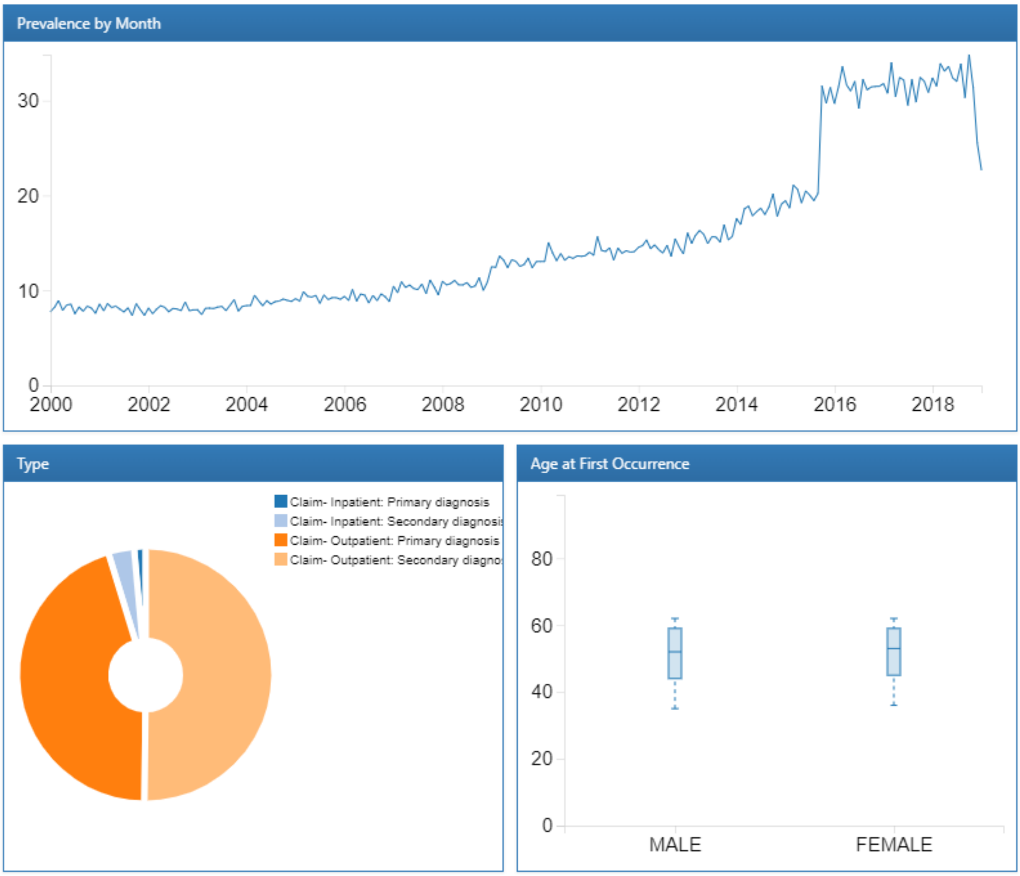


Figure 11.5: ATLAS□□□□□□: □□□□□□□□□□□□□□□□

□□□ Era□□

11.7 ATLAS

□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.7.1

□□□□□□□□□□□□□□1□□□□□□□□□□□□1□□□□□□□□□□□□□□□□2□□□□□□□□ B.6□□2□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□3□□□□□□□□□ B.7□□

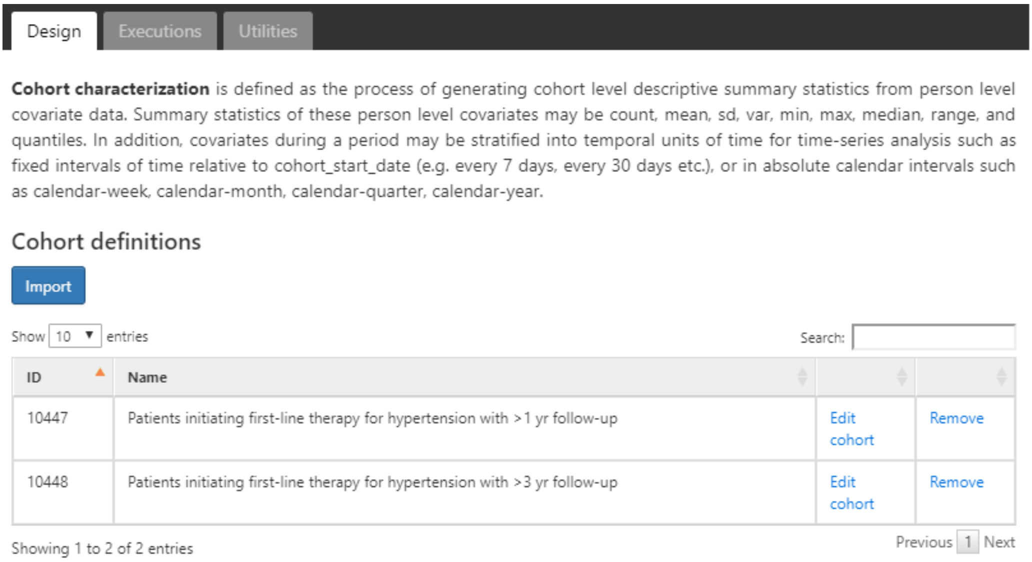


Figure 11.6: □□□□□□□ - □□□□□□□□□□

□□□□□□□

□□□□□□□ATLAS□□□□□□□□□□□□□□□□Chapter 10□□□□□ Import □□□□□□□□□ 11.6□□□

□□□□

ATLAS□□OMOP CDM□□□□□□□□□□□□□□□□□□□□□□□□□□100□□□□□□□□□□□□ R□□□□□□□□□□□□□□□□□□□□□□□□FeatureExtraction□R□□□□□□□□□□□□□□□□

Import □□□

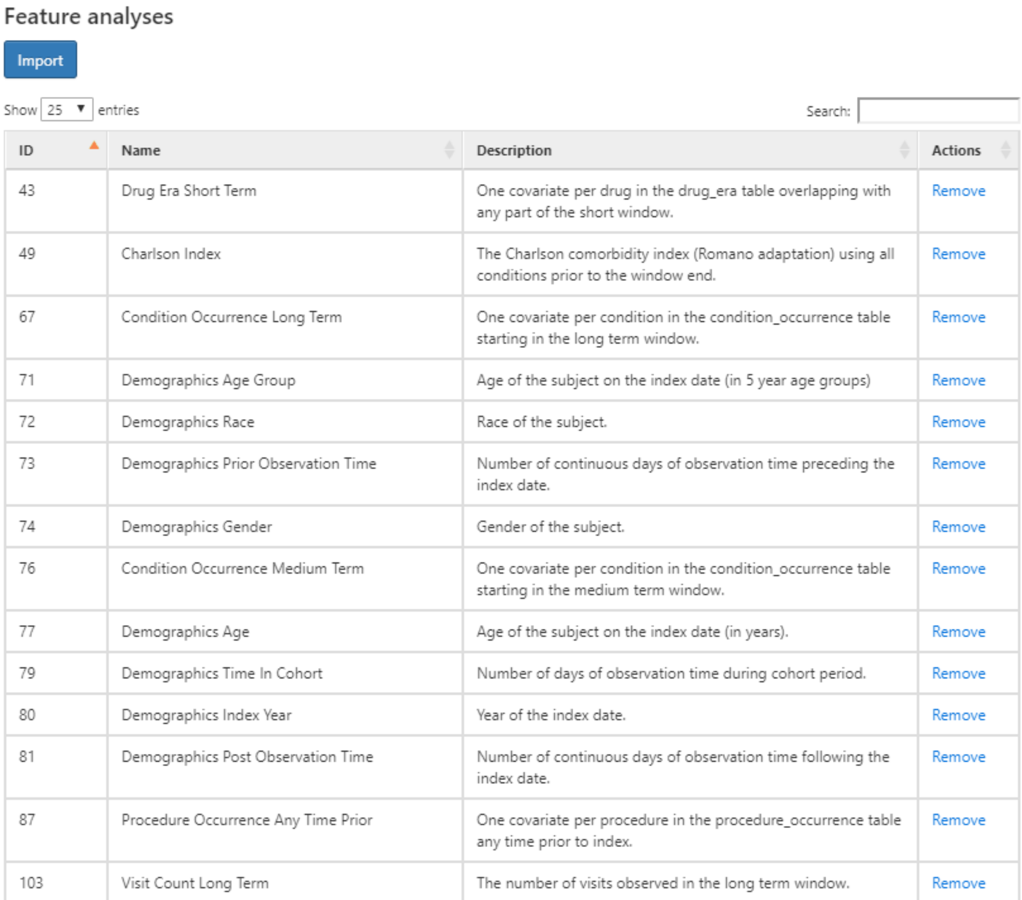
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Demographics□□□□□□□□□□□□□□ Procedure, Condition, Drug□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- **Any time prior**: □□□□□□□□□□□□□□□□□□□□□□□□□
- **Long term**: □□□□□□□□□□365□□
- **Medium term**: □□□□□□□□□□180□□
- **Short term**: □□□□□□□□□□30□□

□□□□□□□□

□□

□□



Feature analyses

Import

Show 25 entries Search:

ID	Name	Description	Actions
43	Drug Era Short Term	One covariate per drug in the drug_era table overlapping with any part of the short window.	Remove
49	Charlson Index	The Charlson comorbidity index (Romano adaptation) using all conditions prior to the window end.	Remove
67	Condition Occurrence Long Term	One covariate per condition in the condition_occurrence table starting in the long term window.	Remove
71	Demographics Age Group	Age of the subject on the index date (in 5 year age groups)	Remove
72	Demographics Race	Race of the subject.	Remove
73	Demographics Prior Observation Time	Number of continuous days of observation time preceding the index date.	Remove
74	Demographics Gender	Gender of the subject.	Remove
76	Condition Occurrence Medium Term	One covariate per condition in the condition_occurrence table starting in the medium term window.	Remove
77	Demographics Age	Age of the subject on the index date (in years).	Remove
79	Demographics Time In Cohort	Number of days of observation time during cohort period.	Remove
80	Demographics Index Year	Year of the index date.	Remove
81	Demographics Post Observation Time	Number of continuous days of observation time following the index date.	Remove
87	Procedure Occurrence Any Time Prior	One covariate per procedure in the procedure_occurrence table any time prior to index.	Remove
103	Visit Count Long Term	The number of visits observed in the long term window.	Remove

Figure 11.7: □□□□□□□ - □□□□

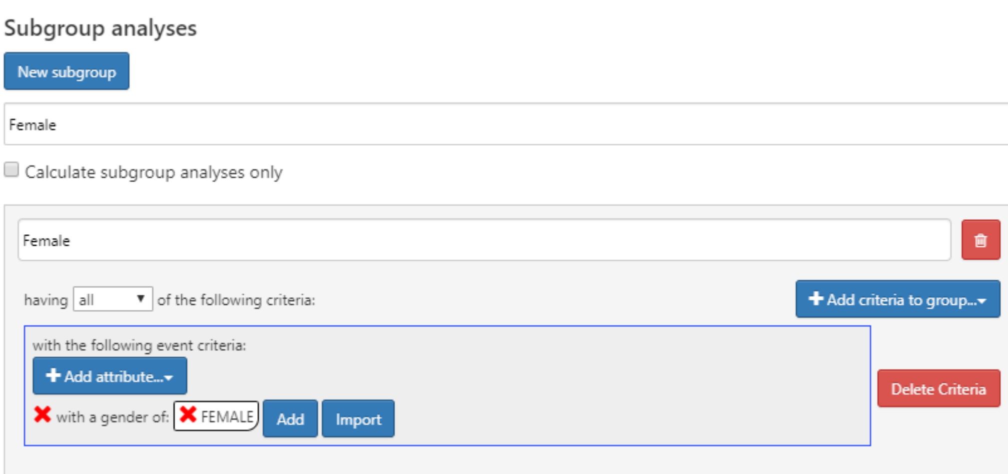


Figure 11.8: □□□□□□□ - □□□□□□□□□□

ATLAS□□□

11.7.2

□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

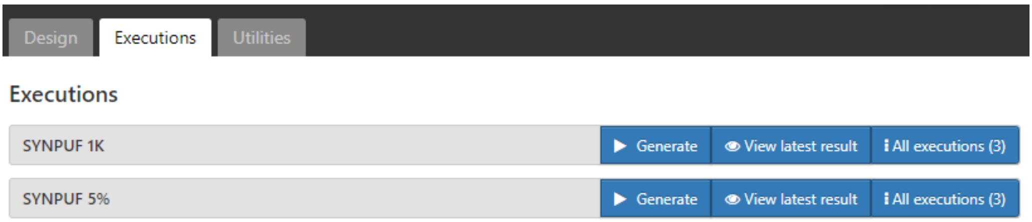
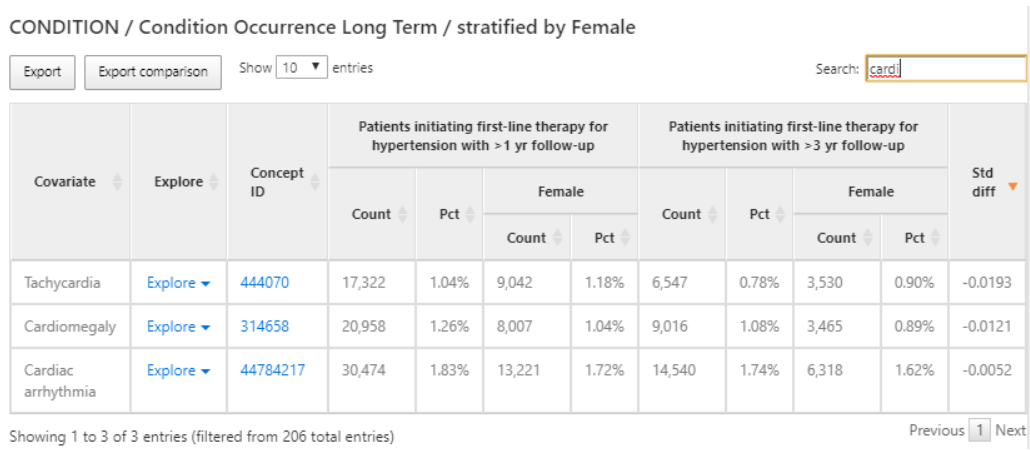


Figure 11.9: □□□□□□□□□ - CDM□□□□□

□□□□□□□□□□All Executions□□□□□□□□□□□□□□□□□□□□□□□□□□View Reports□□□□□□□□□□□□□□View latest result□□□□□□□□□□□□□□□□□□□□□□□□□□

11.7.3


CONDITION / Condition Occurrence Long Term / stratified by Female

Export | Export comparison | Show 10 entries | Search: cardi

Covariate	Explore	Concept ID	Patients initiating first-line therapy for hypertension with >1 yr follow-up				Patients initiating first-line therapy for hypertension with >3 yr follow-up				Std diff
			Count	Pct	Female		Count	Pct	Female		
					Count	Pct			Count	Pct	
Tachycardia	Explore	444070	17,322	1.04%	9,042	1.18%	6,547	0.78%	3,530	0.90%	-0.0193
Cardiomegaly	Explore	314658	20,958	1.26%	8,007	1.04%	9,016	1.08%	3,465	0.89%	-0.0121
Cardiac arrhythmia	Explore	44784217	30,474	1.83%	13,221	1.72%	14,540	1.74%	6,318	1.62%	-0.0052

Showing 1 to 3 of 3 entries (filtered from 206 total entries) Previous 1 Next

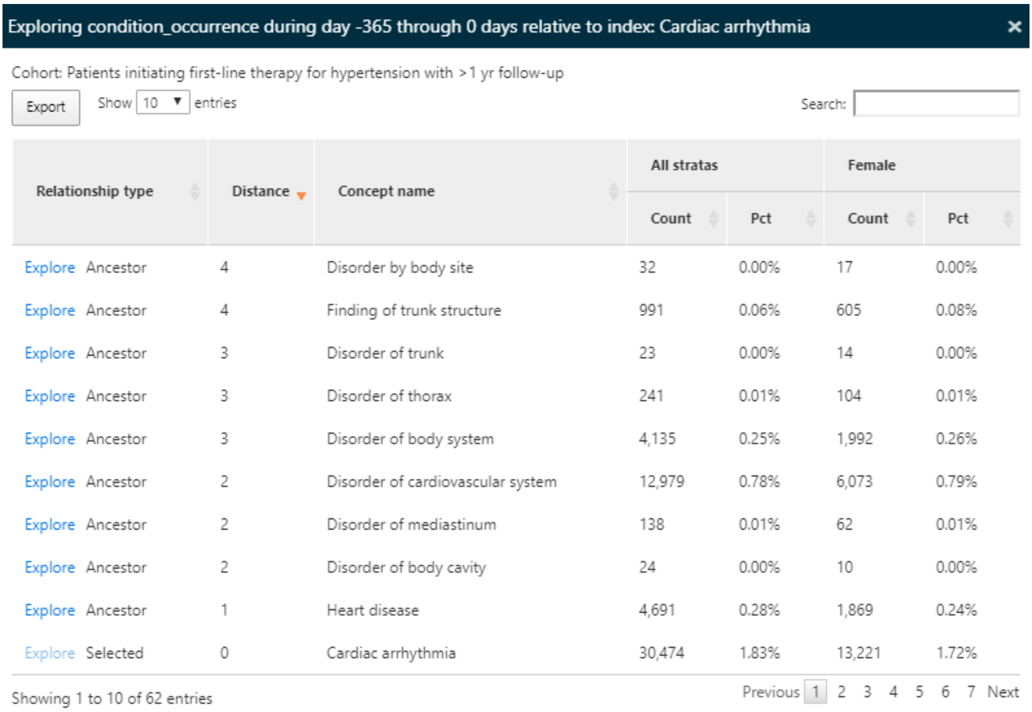
Figure 11.10: □□□□□□□ - □□1□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□11.10□□□□□□□□□□□□□□365

□□ @ref(fig:atlasCharacterizationResultsExplore□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□explore□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ @ref(fig:atlasCharacterizationResultsContra□□□□□□



Exploring condition_occurrence during day -365 through 0 days relative to index: Cardiac arrhythmia ×

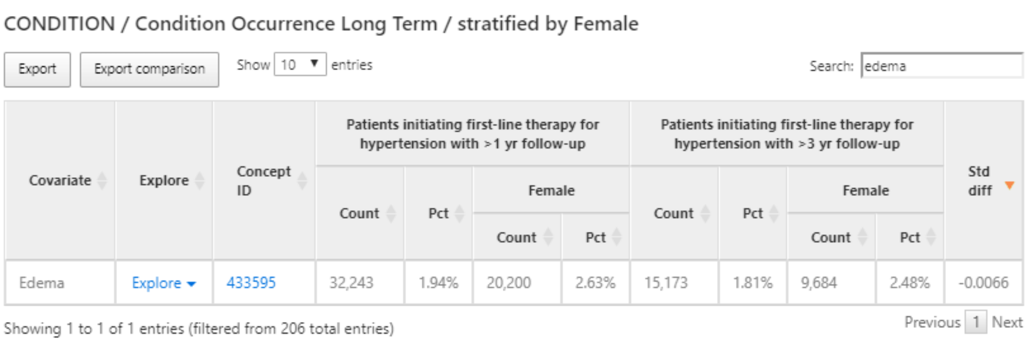
Cohort: Patients initiating first-line therapy for hypertension with >1 yr follow-up

Export Show 10 entries Search:

Relationship type		Distance	Concept name	All stratas		Female	
				Count	Pct	Count	Pct
Explore	Ancestor	4	Disorder by body site	32	0.00%	17	0.00%
Explore	Ancestor	4	Finding of trunk structure	991	0.06%	605	0.08%
Explore	Ancestor	3	Disorder of trunk	23	0.00%	14	0.00%
Explore	Ancestor	3	Disorder of thorax	241	0.01%	104	0.01%
Explore	Ancestor	3	Disorder of body system	4,135	0.25%	1,992	0.26%
Explore	Ancestor	2	Disorder of cardiovascular system	12,979	0.78%	6,073	0.79%
Explore	Ancestor	2	Disorder of mediastinum	138	0.01%	62	0.01%
Explore	Ancestor	2	Disorder of body cavity	24	0.00%	10	0.00%
Explore	Ancestor	1	Heart disease	4,691	0.28%	1,869	0.24%
Explore	Selected	0	Cardiac arrhythmia	30,474	1.83%	13,221	1.72%

Showing 1 to 10 of 62 entries Previous 1 2 3 4 5 6 7 Next

Figure 11.11: □□□□□□□ - □□□□□□□□□□



CONDITION / Condition Occurrence Long Term / stratified by Female

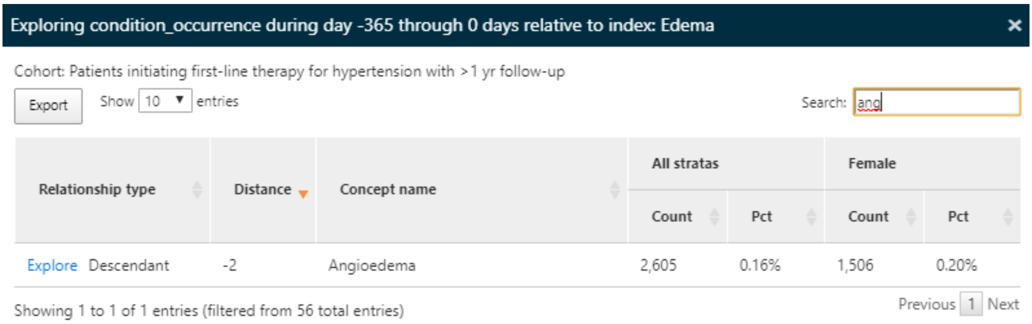
Export Export comparison Show 10 entries Search: edema

Covariate	Explore	Concept ID	Patients initiating first-line therapy for hypertension with >1 yr follow-up				Patients initiating first-line therapy for hypertension with >3 yr follow-up				Std diff
			Count	Pct	Female		Count	Pct	Female		
					Count	Pct			Count	Pct	
Edema	Explore	433595	32,243	1.94%	20,200	2.63%	15,173	1.81%	9,684	2.48%	-0.0066

Showing 1 to 1 of 1 entries (filtered from 206 total entries) Previous 1 Next

Figure 11.12: □□□□□□□ - □□□□□□□

□□□□explore□□


Exploring condition_occurrence during day -365 through 0 days relative to index: Edema

Cohort: Patients initiating first-line therapy for hypertension with >1 yr follow-up

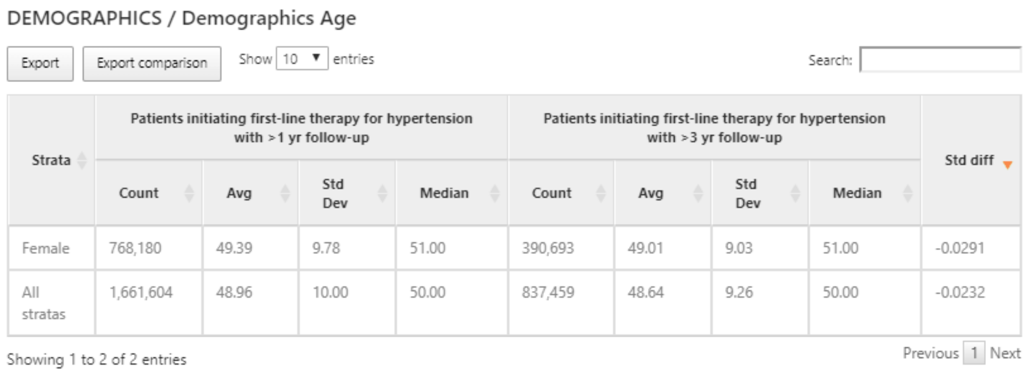
Export Show 10 entries Search: ang

	Relationship type	Distance	Concept name	All stratas		Female	
				Count	Pct	Count	Pct
Explore	Descendant	-2	Angioedema	2,605	0.16%	1,506	0.20%

Showing 1 to 1 of 1 entries (filtered from 56 total entries) Previous 1 Next

Figure 11.13: □□□□□□□□ - □□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□


DEMOGRAPHICS / Demographics Age

Export Export comparison Show 10 entries Search:

Strata	Patients initiating first-line therapy for hypertension with >1 yr follow-up				Patients initiating first-line therapy for hypertension with >3 yr follow-up				Std diff
	Count	Avg	Std Dev	Median	Count	Avg	Std Dev	Median	
Female	768,180	49.39	9.78	51.00	390,693	49.01	9.03	51.00	-0.0291
All stratas	1,661,604	48.96	10.00	50.00	837,459	48.64	9.26	50.00	-0.0232

Showing 1 to 2 of 2 entries Previous 1 Next

Figure 11.14: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

11.7.4

□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ **Analysis**□□□□□□□□□□□□**New Feature Analysis**□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□ACE□□□

□□ Analysis□□□□□□□□□□□□□□ Import □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.8 R

□□□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI R□□□□□□□□FeatureExtraction□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

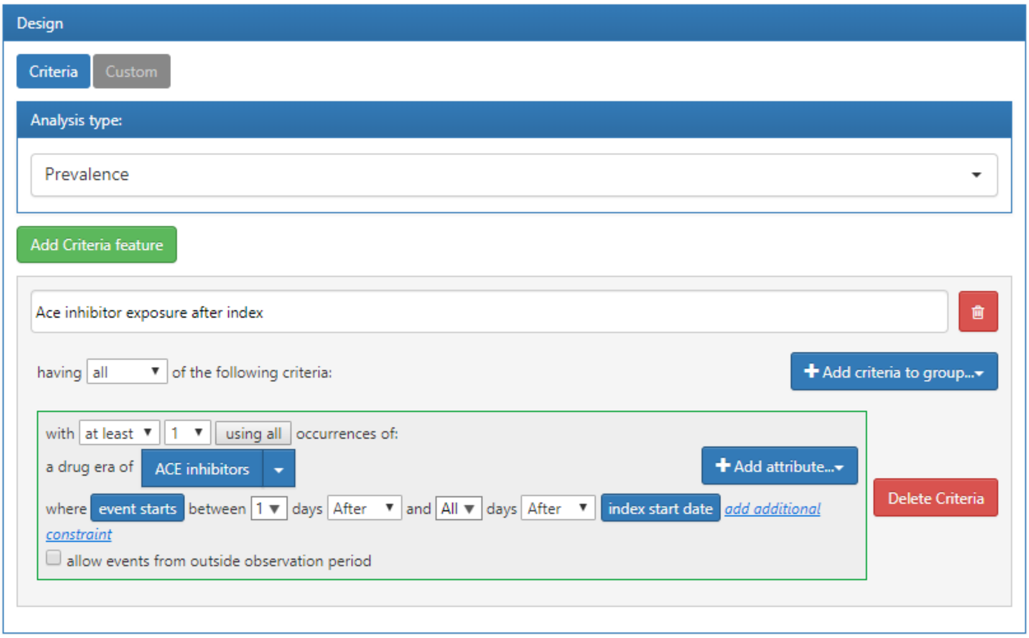
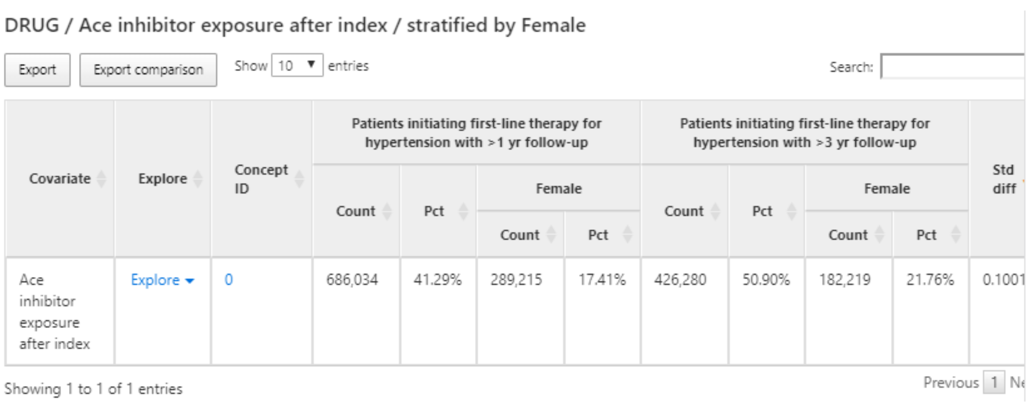


Figure 11.15: ATLAS□□□□□□□□□□



Covariate	Explore	Concept ID	Patients initiating first-line therapy for hypertension with >1 yr follow-up				Patients initiating first-line therapy for hypertension with >3 yr follow-up				Std diff
			Count	Pct	Female		Count	Pct	Female		
					Count	Pct			Count	Pct	
Ace inhibitor exposure after index	Explore	0	686,034	41.29%	289,215	17.41%	426,280	50.90%	182,219	21.76%	0.1001

Figure 11.16: □□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□

FeatureExtraction□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□

11.8.1

□□Chapter 10□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□ B.6□□□□□□□□□□□□□□Appendix B□□□□□□□□□□□□□scratch.my_cohorts□□□□□□□

11.8.2

□□□R□□□□□□□□□□□□□□□□□□□□□□□FeatureExtraction□DatabaseConnector□□□□□□

```r
library(FeatureExtraction)
connDetails <- createConnectionDetails(dbms = "postgresql",
                                       server = "localhost/ohdsi",
                                       user = "joe",
                                       password = "supersecret")

cdmDbSchema <- "my_cdm_data"
cohortsDbSchema <- "scratch"
cohortsDbTable <- "my_cohorts"
cdmVersion <- "5"
```

□□□4□□□cdmDbSchema□cohortsDbSchema□cohortsDbTable□□□□□□CDM□□□□□□□□□□ SQL Server□□□ <- "my_cdm_data.dbo"□□□□□□

11.8.3

createCovariateSettings□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```r
settings <- createCovariateSettings(
  useDemographicsGender = TRUE,
  useDemographicsAgeGroup = TRUE,
  useConditionOccurrenceAnyTimePrior = TRUE)
```

□□□□□□□□□□□□□5□□

□□

- □□: □□□□□□□□□□365□□□□
- □□: □□□□□□□□□□180□□□□
- □□: □□□□□□□□□□30□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```r
settings <- createCovariateSettings(useConditionEraLongTerm = TRUE,
                                    useConditionEraShortTerm = TRUE,
                                    useDrugEraLongTerm = TRUE,
                                    useDrugEraShortTerm = TRUE,
                                    longTermStartDays = -180,
                                    shortTermStartDays = -14,
                                    endDays = -1)
```

□□□□□□□□□□□□□□□□□□□□□□□□□180□□□□□□□□□□□□□□□□□□□□□□□□14□□□

□□□□□□□□□□ID□□□□□□□□□□□□□□□□□□□□□□

```r
settings <- createCovariateSettings(useConditionEraLongTerm = TRUE,
                                    useConditionEraShortTerm = TRUE,
                                    useDrugEraLongTerm = TRUE,
                                    useDrugEraShortTerm = TRUE,
                                    longTermStartDays = -180,
                                    shortTermStartDays = -14,
                                    endDays = -1,
                                    excludedCovariateConceptIds = 1124300,
                                    addDescendantsToExclude = TRUE,
                                    aggregated = TRUE)
```

□□□□□□□□□□□□□aggregated = TRUE□□□□□□FeatureExtraction□□□□□□□□□□

11.8.4

□□□□□□□□□□□□□□□□□□□□□□□□□□□

```r
covariateSettings <- createDefaultCovariateSettings()

covariateData2 <- getDbCovariateData(
  connectionDetails = connectionDetails,
  cdmDatabaseSchema = cdmDatabaseSchema,
  cohortDatabaseSchema = resultsDatabaseSchema,
  cohortTable = "cohorts_of_interest",
  cohortId = 1,
  covariateSettings = covariateSettings,
  aggregated = TRUE)

summary(covariateData2)
```

□□□□□□□□□□□□□

```
## CovariateData Object Summary
##
## Number of Covariates: 41330
## Number of Non-Zero Covariate Values: 41330
```

11.8.5

□□□□□covariateData□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□cov

```
covariateData2$covariates
covariateData2$covariatesContinuous
```

11.8.6

FeatureExtraction□□ //ohdsi.github.io/FeatureExtraction/

11.9 ATLAS

□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ Hripcsak7329 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□Cohort Pathways□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

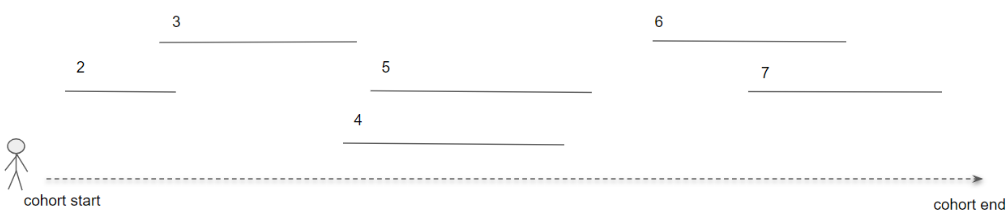


Figure 11.17: □□□□□□□□□□□□□□□□□□□

□ 11.17 □□□

□□□ATLAS□□□□□□□ Cohort Pathways □□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.9.1

□□□1□□□□3□□Appendix B.6, B.7□□□□□□□□□□2□□□□□□□□□□□□□□□

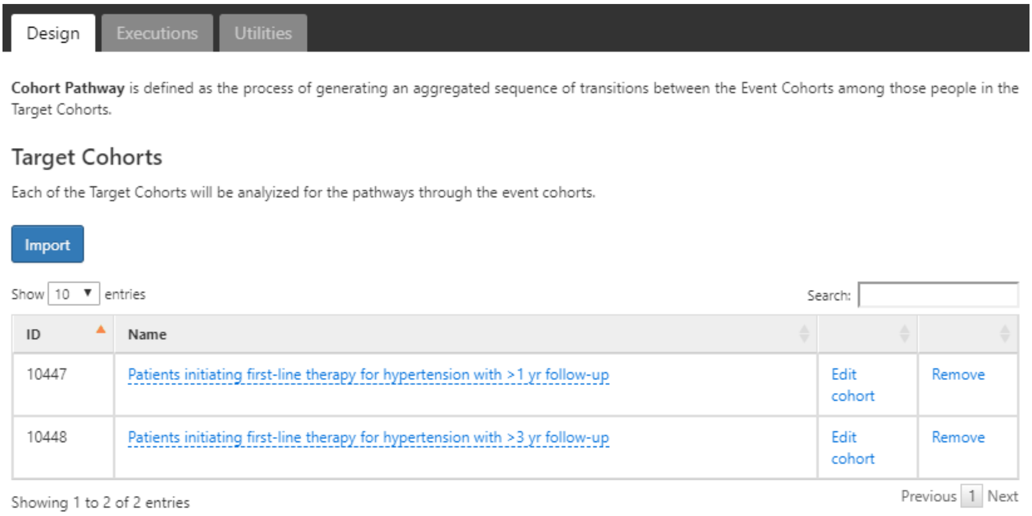


Figure 11.18: □□□□□□□□□□□□□□□□□□□□□

□□ACE□ B.8-B.16□□□□□□□□□□□□□□□□ Import □□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

- □□□□□□□□□□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ + □□□□□□□□2□□□□□□□□□□□□
- □□□□□: □□
- □□□□□: □□□□□□□□□□□□□□□□□□□□□□□□□

11.9.2

□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.9.3

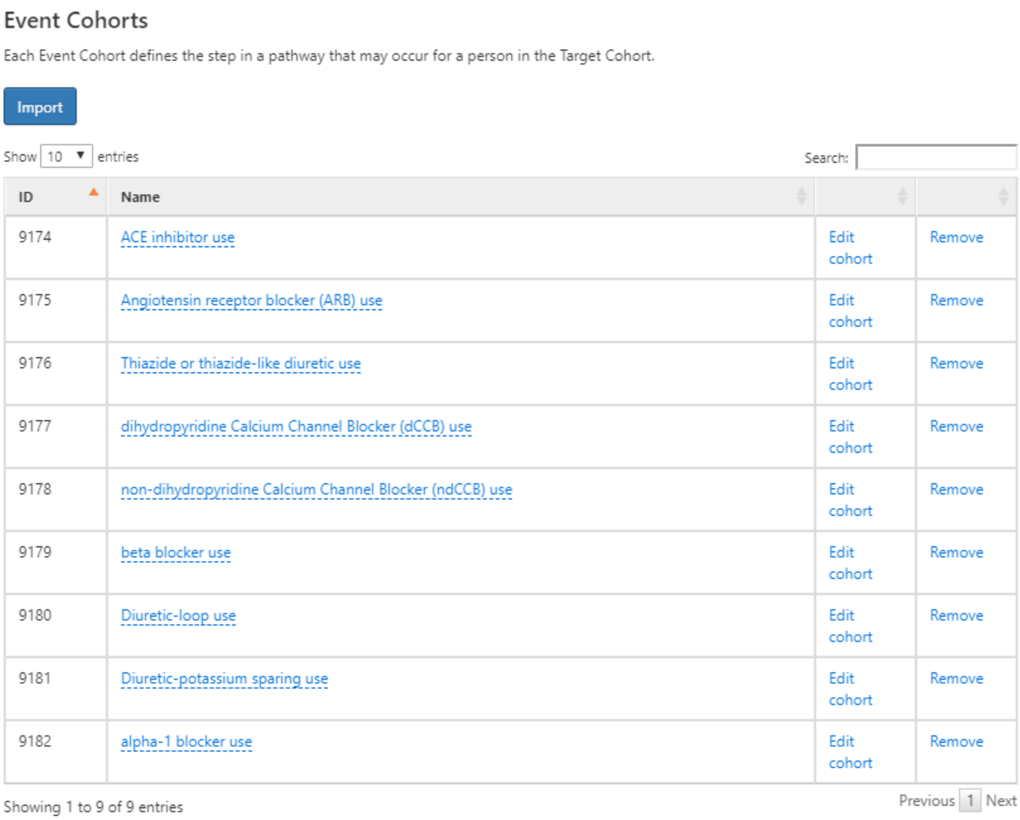
□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□

□□□

11.10 ATLAS

□□□ (ACEi) □□□□□□□□□□□□□□□□□□□ (THZ) □□□□□□□□□□□□□□ (Angioedema) □□□□□□□□□ (AMI) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ (TAR) □□□□□□□□□□□□□□□□□□□□□ (ACEi □□□ THZ) □□□□□□□□□□□□□□□□□□□□



Event Cohorts

Each Event Cohort defines the step in a pathway that may occur for a person in the Target Cohort.

Import

Show 10 entries   Search:

ID	Name		
9174	ACE inhibitor use	Edit cohort	Remove
9175	Angiotensin receptor blocker (ARB) use	Edit cohort	Remove
9176	Thiazide or thiazide-like diuretic use	Edit cohort	Remove
9177	dihydropyridine Calcium Channel Blocker (dCCB) use	Edit cohort	Remove
9178	non-dihydropyridine Calcium Channel Blocker (ndCCB) use	Edit cohort	Remove
9179	beta blocker use	Edit cohort	Remove
9180	Diuretic-loop use	Edit cohort	Remove
9181	Diuretic-potassium sparing use	Edit cohort	Remove
9182	alpha-1 blocker use	Edit cohort	Remove

Showing 1 to 9 of 9 entries   Previous 1 Next

Figure 11.19: □□□□□□□□□□□□□□□□□□□□□□□□□

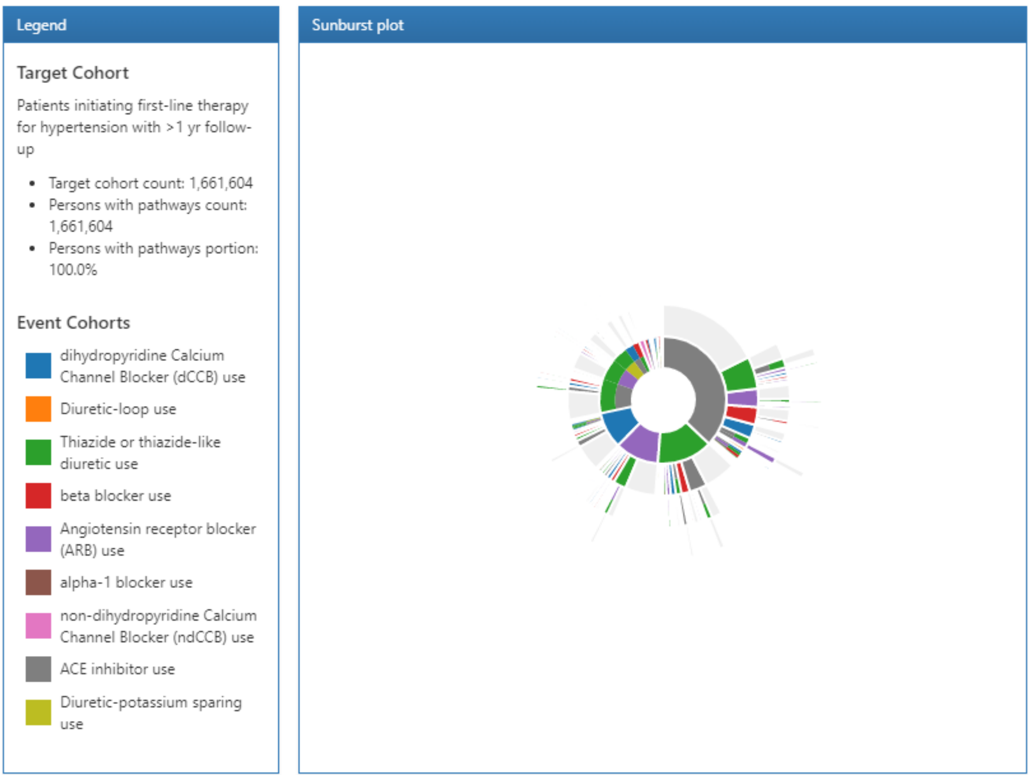

Legend
Sunburst plot
Target Cohort
Patients initiating first-line therapy for hypertension with >1 yr follow-up
Target cohort count: 1,661,604
Persons with pathways count: 1,661,604
Persons with pathways portion: 100.0%
Event Cohorts
dihydropyridine Calcium Channel Blocker (dCCB) use
Diuretic-loop use
Thiazide or thiazide-like diuretic use
beta blocker use
Angiotensin receptor blocker (ARB) use
alpha-1 blocker use
non-dihydropyridine Calcium Channel Blocker (ndCCB) use
ACE inhibitor use
Diuretic-potassium sparing use

Figure 11.20: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

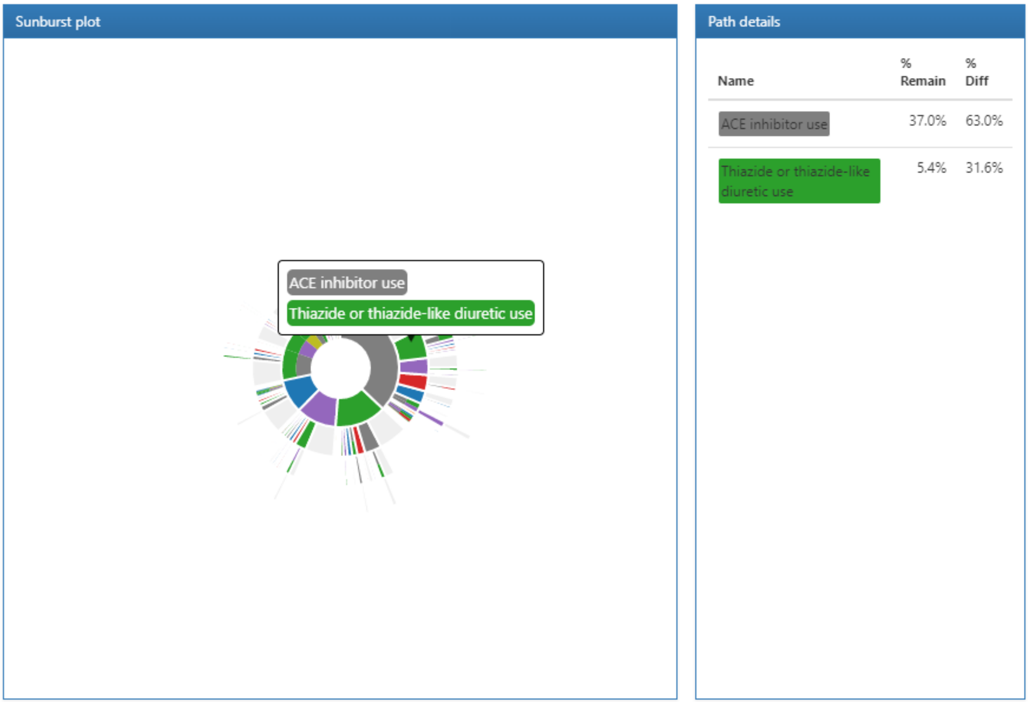


Figure 11.21: □□□□□□□□□□□□□□□□□□□□

(ARB) □□□□□□□□□□□□□□□ARB □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ ARB □□□□□□□□□□□□□□□□□□

□□□□□□ATLAS □□□□□□□ Incidence Rates □□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□

11.10.1

□□□□□□□□□□□□□□□□ ATLAS □□□□□□□□□□□□□□ (Chapter 10 □□□)□□□□□□□□□□□□□□□ (Appendix B.2, B.5) □□□□□□□□□□□□ (Appendix B.4, B.3, B.9) □□□□□□□□□□□□□□□□

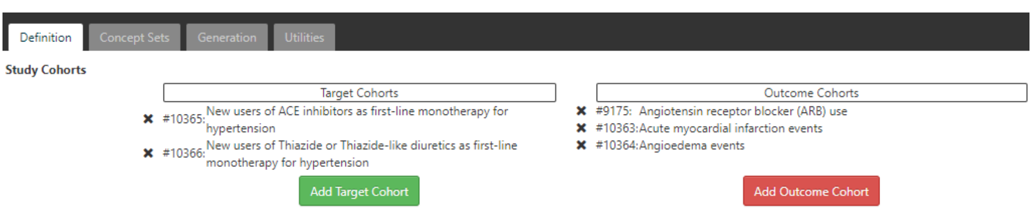


Figure 11.22: □□□□□□□□□□□□□□□□□□□□

□□□□□□ *New users of ACE inhibitors* □□□□□ *New users of Thiazide or Thiazide-like diuretics* □□□

acute myocardial infarction events □ *angioedema events* □□□□ *Angiotensin receptor blocker (ARB) use* □□

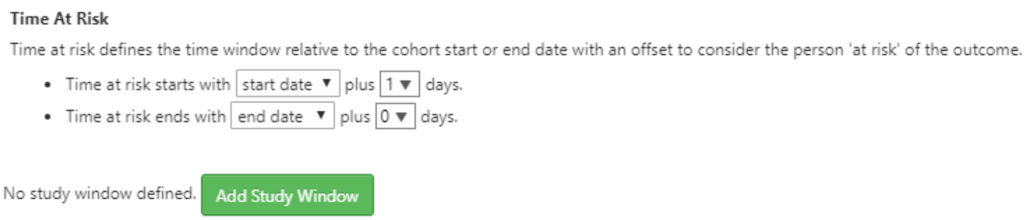


Figure 11.23: □□□□□□□□□□□□□□□□□□□□

□□□ □□□ THZ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

ATLAS □□□:

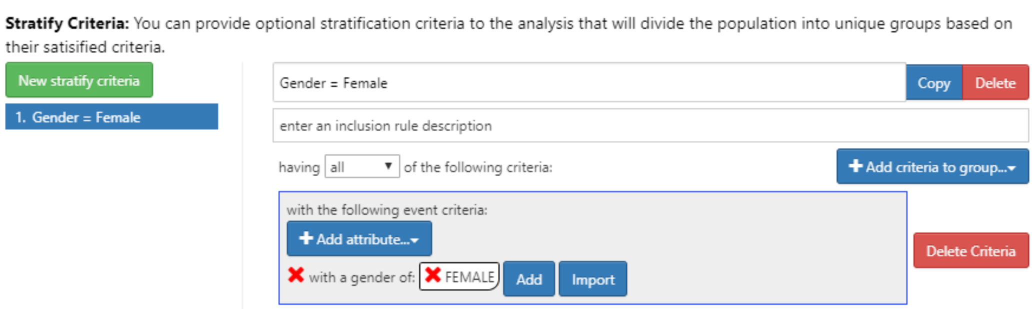


Figure 11.24: □□□□□□□□□□□□□

□□□□□□□□[□□□□□□□]□□□□□□□□□□Chapter 11 □□□□□□□□□□□□□□□□□□□□

11.10.2

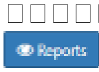
[□□] □□□□□□□□□ 

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□:

□□□□□□□□□□□□□□□□□[□□] □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.10.3

□□□

□□□□□□□□□□□□□□□□ ACEi □□□□□□□□□□□□□□□□□□□□ (AMI) □□□□□□□ 

□□□□□□□□□□□□□□□□□□□□□□□□□□□□:

□□□□□□□□□□□□□TAR □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 1000□□□ 1000□□□□□□□□□□□□□□□□□□

Figure 11.25: □□□□□□□□□

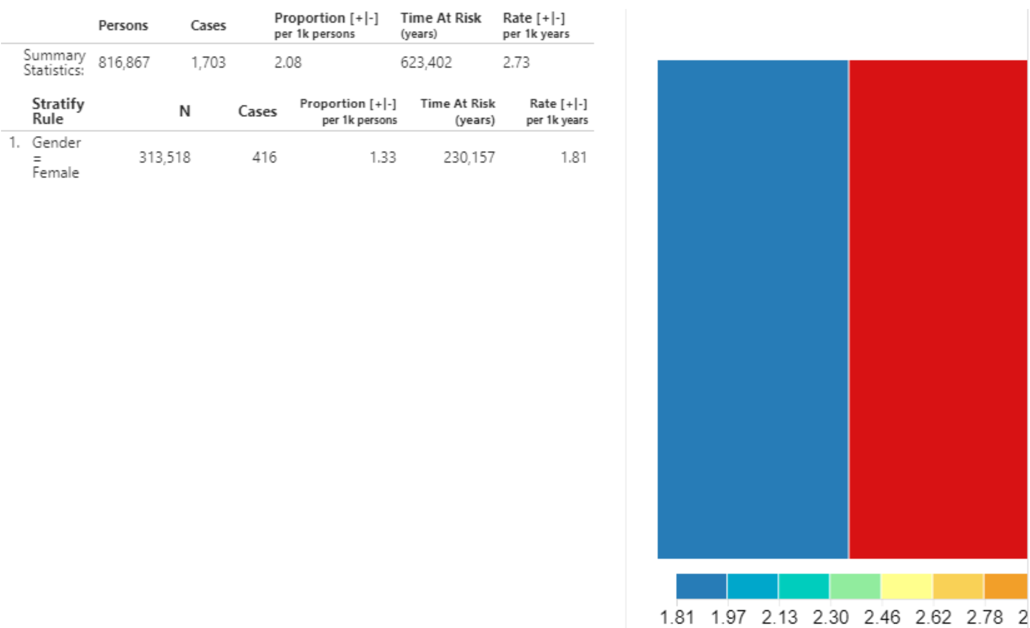
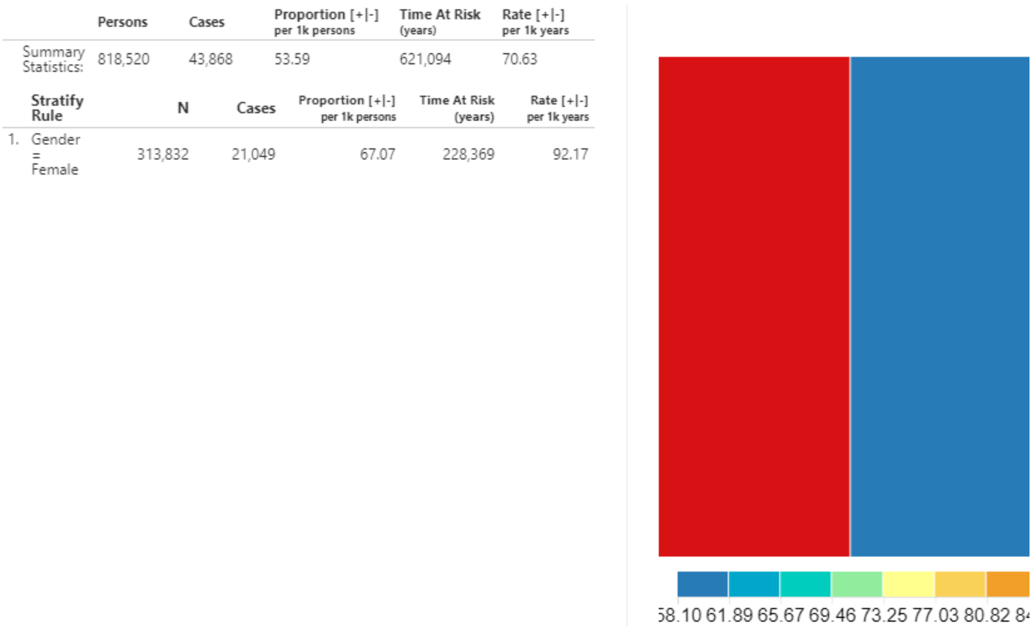


Figure 11.26: □□□□□□□□□ - AMI □□□□□□□□□□□□ ACEi □□□□

□□□

ACEi □□□□□ ARB □□
ARB □□□□□□□ Reports □□□□□□□□□□□□□□□□□□□□□



	Persons	Cases	Proportion [+|-] per 1k persons	Time At Risk (years)	Rate [+|-] per 1k years
Summary Statistics:	818,520	43,868	53.59	621,094	70.63

Stratify Rule	N	Cases	Proportion [+|-] per 1k persons	Time At Risk (years)	Rate [+|-] per 1k years
1. Gender = Female	313,832	21,049	67.07	228,369	92.17

Figure 11.27: □□□ - ACEi □□□□ ARB □□□□□□□□□□ ACEi □□□□

□□□ (ARB □□)
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

11.11

- OHDSI □□
- □□
- ATLAS □□□□□□□□□□□□□ OHDSI □□□□□□□□□□□□□□□□□□□□□□□□
- ATLAS □□

11.12

□□□□

□□□□□□□□□ATLAS □□□□□□□□□□□□□□□□□□http://atlas-demo.ohdsi.org
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ **11.1.** □□□
□□

□□ **11.2.** □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□365
□□□
(Chapter 10 □□□□□□□□□□□□)□

□□ **11.3.** □□□□□□□□□□□□□□□□□(GI □□) □□□□□□□□□□□□□□□□□□□□□□□□□□192
(“□□□□□□”) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ GI
□□
GI □□□□□□□□□□□□□□□□□□□□

□□□□□□□□ Appendix E.7 □□□□□□□□□□

Chapter 12

12 – –

□□□□□□□□: *Martijn Schuemie, David Madigan, Marc Suchard & Patrick Ryan*

□□

- □□□□□□: □□
- □□□□□□: □□

□□

□□

□□□□□□□OHDSI Methods Library□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.1

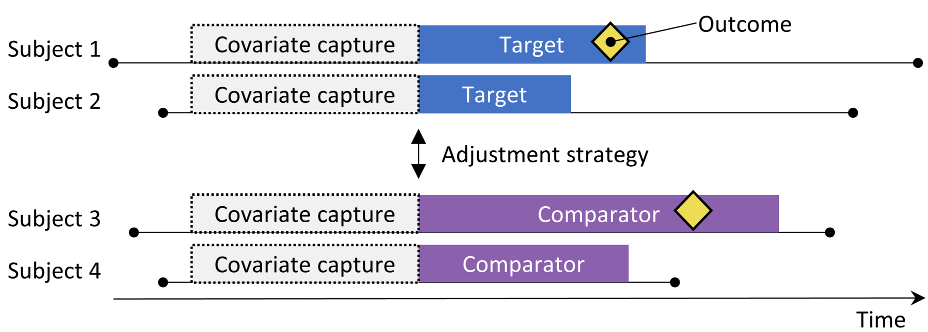


Figure 12.1: □□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ (Hernan and Robins, 2016,)□ □□□

□: □□□□□□□□□□□□□□□□□□□

□□	□□
□□□□□□□□□	□□□□□□□□□□□□□□□□
□□□□□□□□	□□□□□□□□□□□□□□□
□□□□□□□□□	□□□□□□□□□□□□□□□□□□□
□□□□□	□□
□□□	□□

□□

□□

□□

12.1.1

□□

□□□□□□□□□□□□□□*PS*□□□□□□□□□□□□□□□□□□□□□□ (Rosenbaum and Rubin, 1983,)□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0.5□□□□□ et al., 2012,)□

□□□□□□□□PS□□□□□□□□□□□□□□Jan□□□□□□□□□□□□□□□□□0.4□□□□□□□□□□ Jun□□□

□□ “□□□□□□□□”□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 18□□□□□□□

12.1.2

□□□□PS□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□ et al., 2018,)□□□ et al., 2013,)□□□□□□□□□□□Cyclops□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

□□□

12.1.3

□□□□□□□0□□1□□□ (2011) □□□□□□□□□□□□□□□□□□0.2□□□□□□□

12.1.4

□□□

$$\ln\left(\frac{F}{1-F}\right) = \ln\left(\frac{S}{1-S}\right) - \ln\left(\frac{P}{1-P}\right)$$

□□□ F □□□□□□□S □□□□□□□P □□□□□□□□□□□□□□□□□□□□

Walker et al. (2013) □□□

12.1.5

□□□□□□□PS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□12.11□□□□□□□ 2001,)□

12.2

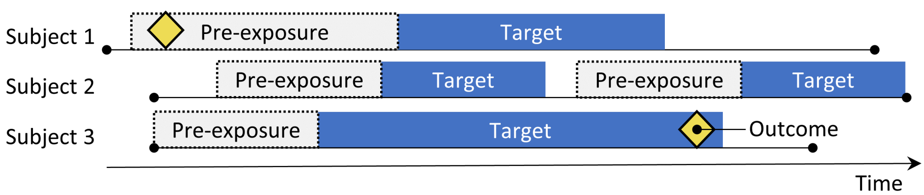


Figure 12.2: □□

□□□□□□□□□SCC□□□□□ (Ryan et al., 2013a,) □□□□□□□□□□□□□□□□□□□□□□□ 12.2 □□□4□□□□□□□□□□□□□□□□□□□□□□□□□

□□□

12.3

□□□□□□ (Vandenbroucke and Pearce, 2012,) □□□□□□□□□□□□□□□□□□□□□□□□□□ 12.3 □□□□□□□□□□□□□□□□□□□□

□□□1□□□□

Table 12.2: □□□□□□□□□□□□□□□□□□□□□□

Choice	Description
Target cohort	□□□□□□□□□
Outcome cohort	□□□□□□□□□□□□□□□□□
Time-at-risk	□□□
Control time	□□□□□□□□□□□□□□□□□□□

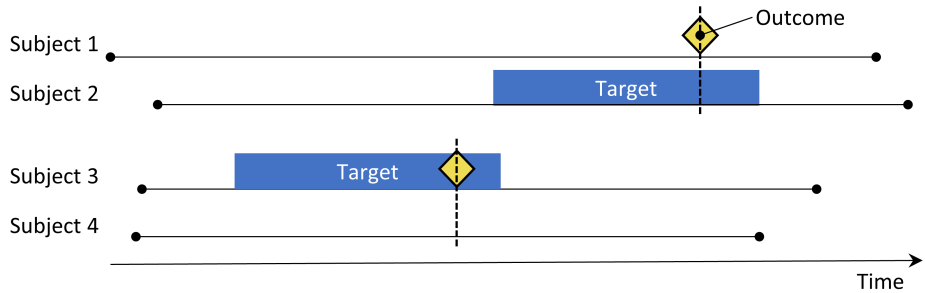


Figure 12.3: □□□

12.4

□□□□ (Maclure, 1991,) □□ 12.4 □□□□□□□□□□□□□□□□□□□□□□□

□□ □□-□□□□□□□□□□ (Suissa, 1995,) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.5

□□□□□□□□□□□SCCS□□□□□(Farrington, 1995; Whitaker et al., 2006,)□□□ SCCS □□□□□□□□□□

Table 12.3: □□□□□□□□□□□□□□□□□□

Choice	Description
Outcome cohort	□□□□□□□□□□□□□□□□□□□□□
Control cohort	□□□
Target cohort	□□□□□□□□□
Nesting cohort	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Time-at-risk	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

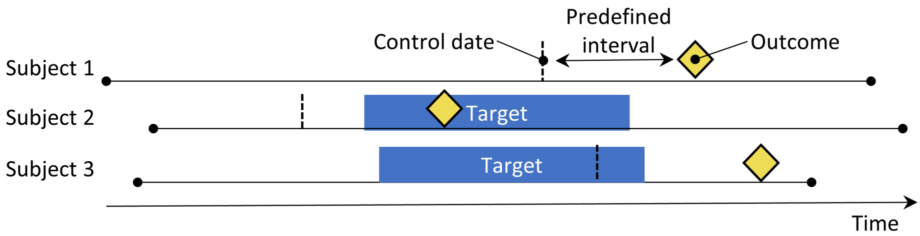


Figure 12.4: □□□

Table 12.4: □□□□□□□□□□□□□□□□□□

Choice	Description
Outcome cohort	□□□□□□□□□□□□□□□□□□□□□
Target cohort	□□□□□□□□□
Time-at-risk	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Control time	□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□SCCS □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ et al., 2013,)□□□

SCCS□□□ et al., 2011,)□

12.6

12.6.1

ACE □□□□ACEi□□ et al., 2002,)□□□ and Black, 1997)□□ et al., 2012,)□□□□□□□□□□□□ACEi□β□□□□□□□□□□□□□□□□□□□□□□□□□□□□(et al., 2010; Toh et al., 2012,)□□β□□□□□□□□□□□□□□□□□□□□□□□□□□□ (Whelto et al., 2018,)□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□THZ□□□□□□□□□□□□

Table 12.5: □□□□□□□□□□□□□□□□□□□□□□□

□□□	□□
□□□□□□□□□	□□□□□□□□□□□
□□□□□□□□□	□□□□□□□□□□□□□□□□□□□
□□□□□	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□	□□□□□□□□□□□□□□□□□□□□□□□□

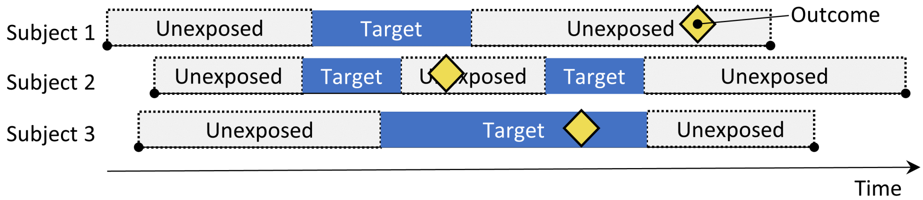


Figure 12.5: □□□

□□

ACE□□

ACE□□

□□□□□□□□□□□□□□□□□□□@ref(CohortMethod)□□□□□□□□□□□□□□□□□□

12.6.2

□□□□□□□□□□□□□□□□□□ACEi□□□THZ□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.6.3

□□□□□□□□□□□□□□□□□□ER□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□7□□

12.6.4

□□30□□□□□□□□□□

12.6.5

□□□□□□□□□□□□□□□□□PS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.6.6

□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□	□
□□□□□□□□□	□□□□□□□□□□□□□□□□ACE□□□□□□□□□□
□□□□□□	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□	□□□□□□□□□□□□□□□□□
□□□□□	□□□□□□□□□□□□□□□□□□□□□□□
□□□	□□□□□□□□□□□□□□□□□□□□□□□□□□

12.6.7

□□□
@ref□MethodValidity□□□□□□□□□□

12.7 ATLAS

□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□
Estimation □□□
□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.7.1

□□□□1□□□
□□2□□□□
10□□□□□□□□□□□Appendix B.2□□□□□Appendix @ref(ThiazidesMono□□□□□□□□□□
@ref(Angioedema□□□□Appendix @ref(Ami□□□□□□□□□□□□□□□□□□□□□□□□□□
12.6□□□□□□□□□□

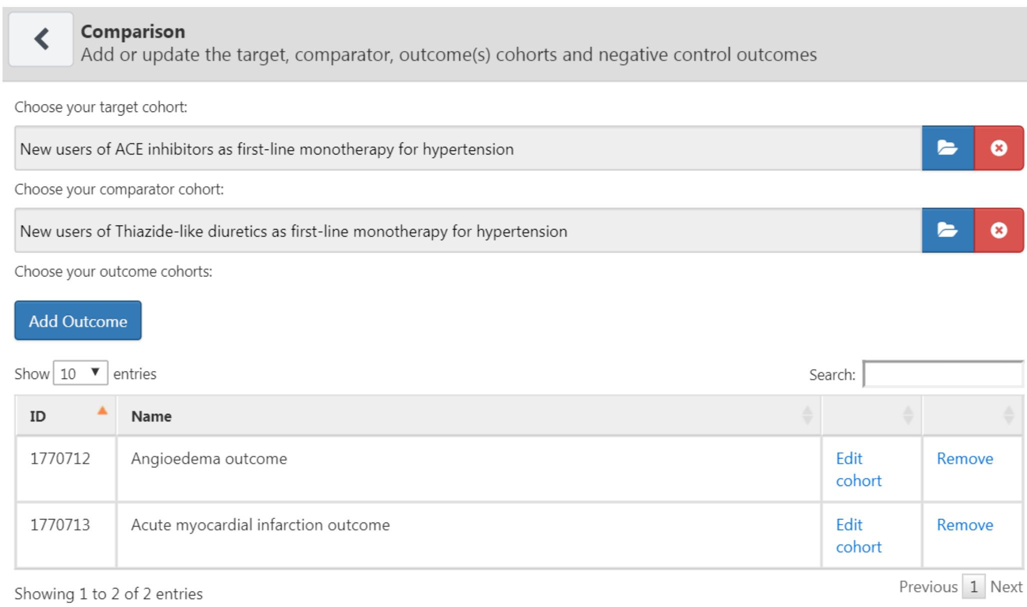


Figure 12.6: □□□□□□□

□□□

□□□□□□□□□

□□□
18□□□

12.7□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

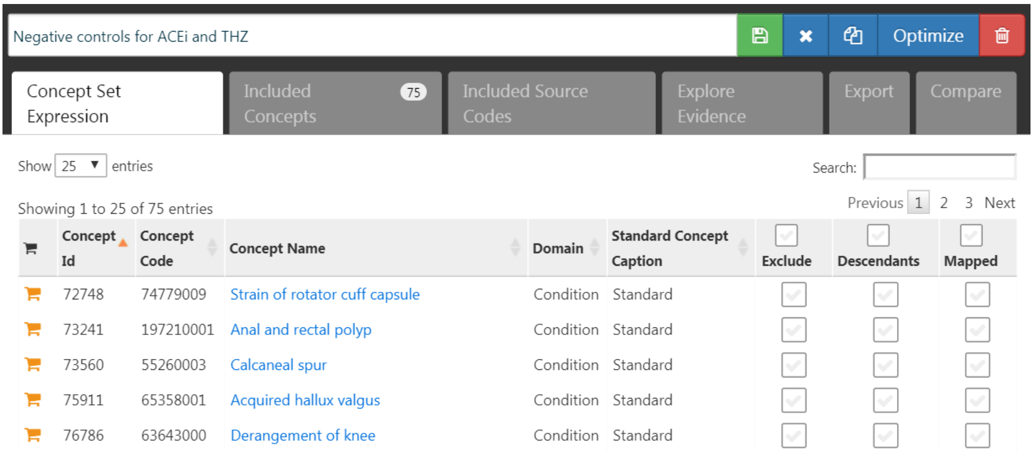


Figure 12.7: □□□□□□□□□□□□□

□□□□□□□□

□□□

□□□□□□□□□

□□□
12.8□□

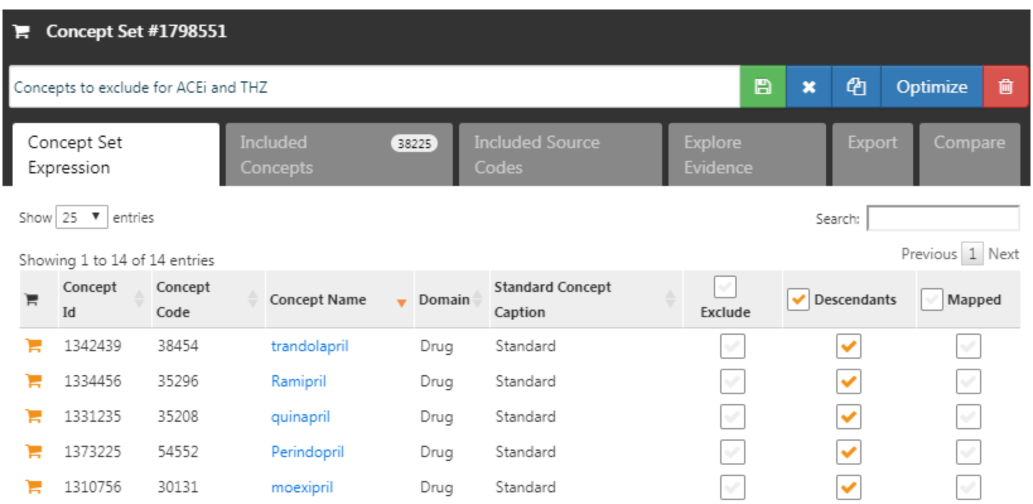


Figure 12.8: □□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Figure
12.9□□□□□□□□□

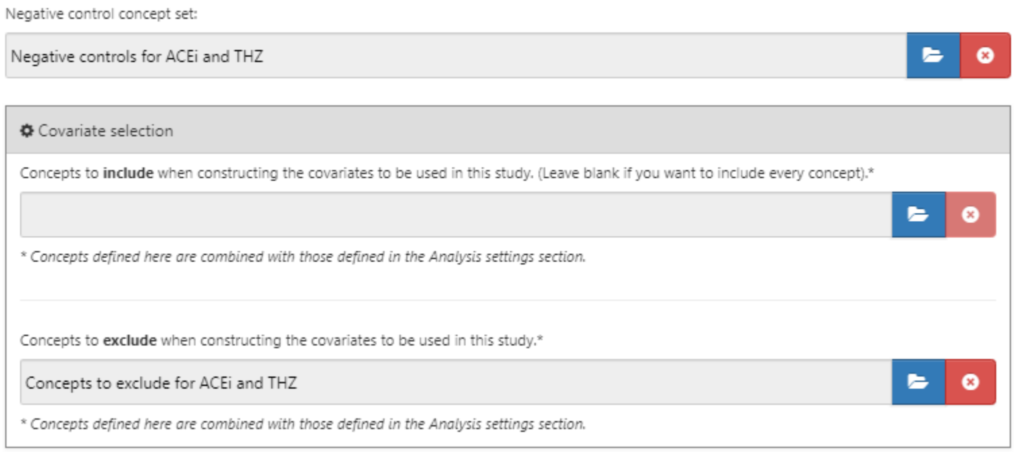


Figure 12.9: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.7.2

□□□

□□□□

□□□

□□□
□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.8 R

□□□□□□□□□□R□□□□□□□□□□□ATLAS□□□□□□□□□□R□□□□□□□□□□□□□□□
R□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□CohortMethod□□□□□□□□□□□□□□□□□□□□□CohortMethod□□CDM□□□□□□□
12.8.6□□□□□□□□□□□□AMI□□□□□□□□□□□□□□□□□□□□□□□□□□

12.8.1

□□□
(10)□□□□□□□□□□□□□□□□□□□□(B.2)□□□□□□□(B.5)□□□□□□□□□□□□□
(B.4)□□□□□□□□□□□□□□□□□□□□ACEi□THZ□□□□□□□□□□□□□□□□□□□□□
1□2□3□□□ `scratch.my_cohorts` □□□□□□□□□□□□□□□□□□□□□□□□

12.8.2

□□□□R□□□□□□□□□□□□□□□□□□□□□□□ CohortMethod□DatabaseConnector□□□□□

```
library(CohortMethod)
connDetails <- createConnectionDetails(dbms = "postgresql",
                                       server = "localhost/ohdsi",
                                       user = "joe",
                                       password = "supersecret")

cdmDbSchema <- "my_cdm_data"
cohortDbSchema <- "scratch"
cohortTable <- "my_cohorts"
cdmVersion <- "5"
```

□□□4□□cdmDbSchema□cohortDbSchema□□□□cohortTable□□□□CDM□□□□□□□□□□□□
SQL Server□□□cdm
<- "my_cdm_data.dbo"□□□□□□□□□□

□□□CohortMethod□□

```
#
aceI <- c(1335471,1340128,1341927,1363749,1308216,1310756,1373225,
          1331235,1334456,1342439)
thz <- c(1395058,974166,978555,907013)

#
cs <- createDefaultCovariateSettings(excludedCovariateConceptIds = c(aceI,
                                                                    thz),
                                     addDescendantsToExclude = TRUE)

#
cmData <- getDbCohortMethodData(connectionDetails = connectionDetails,
                                cdmDatabaseSchema = cdmDatabaseSchema,
                                oracleTempSchema = NULL,
                                targetId = 1,
                                comparatorId = 2,
                                outcomeIds = 3,
                                studyStartDate = "",
                                studyEndDate = "",
                                exposureDatabaseSchema = cohortDbSchema,
                                exposureTable = cohortTable,
                                outcomeDatabaseSchema = cohortDbSchema,
                                outcomeTable = cohortTable,
                                cdmVersion = cdmVersion,
                                firstExposureOnly = FALSE,
                                removeDuplicateSubjects = FALSE,
                                restrictToCommonPeriod = FALSE,
                                washoutPeriod = 0,
                                covariateSettings = cs)
cmData
```

```
## CohortMethodData
##
##     ID 1
##       ID 2
##      ID(s) 3
```

□□□□□□□□□□□□□□□□□□□CohortMethod□□□□□□□□□□□□□□□□`createDefault`
(12.1)□□□2

□□`cohortMethodData`
(8.4.2)□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□`summary()`□□□□□□□□□□

```r
summary(cmData)
```

```
## CohortMethodData
##
##     ID 1
##       ID 2
##      ID(s) 3
##
##      67166
##      35333
##
##
##
## 3              980          891
##
##
##    58349
##       24484665
```

`cohortMethodData`□□

```r
saveCohortMethodData(cmData, "AceiVsThzForAngioedema")
```

□□□□□□□□□□□□□□□□□□□□□`loadCohortMethodData()`□□□□□□□□□□

□□□□□□□□□□

□□□

1. □□□□□□□□□
2. □□□□□□`getDbCohortMethodData`□□□□□□□□□□□□□□`firstExposureOnly`□`rem`
3. `createStudyPopulation`□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□1□□□□□□□□□□□□□□□□CohortMethod□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.8.3

□□□

```
studyPop <- createStudyPopulation(cohortMethodData = cmData,
                                  outcomeId = 3,
                                  firstExposureOnly = FALSE,
                                  restrictToCommonPeriod = FALSE,
                                  washoutPeriod = 0,
                                  removeDuplicateSubjects = "remove all",
                                  removeSubjectsWithPriorOutcome = TRUE,
                                  minDaysAtRisk = 1,
                                  riskWindowStart = 1,
                                  startAnchor = "cohort start",
                                  riskWindowEnd = 0,
                                  endAnchor = "cohort end")
```

firstExposureOnly□removeDuplicateSubjects□FALSE□□□□□washoutPeriod□0□□□□□
= 1□□□startAnchor = "cohort start"□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
= 0□□□endAnchor = "cohort end"□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
getAttritionTable(studyPop)
```

```
##                                 ...
## 1                          67212              35379 ...
## 2                67166              35333 ...
## 3                            67061              35238 ...
## 4    1              66780              35086 ...
```

12.8.4

getDbcohortMethodData()□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
ps <- createPs(cohortMethodData = cmData, population = studyPop)
```

createPs□□□Cyclops□□

□□□□□□□□□□□□□□□□□□□□PS□□□□□□□

```
matchedPop <- matchOnPs(population = ps, caliper = 0.2,
                        caliperScale = "standardized logit", maxRatio = 100)
```

□□□□□PS□trimByPs□trimByPsToEquipoise□□□□stratifyByPs□□□□□□□□□□□□□□□

12.8.5

```
outcomeModel <- fitOutcomeModel(population = matchedPop,
                                modelType = "cox",
                                stratified = TRUE)
outcomeModel
```

```
##      cox
##    TRUE
##      FALSE
##       FALSE
##     OK
##
##                 95%   95% logRr seLogRr
##        4.3203   2.4531   8.0771 1.4633   0.304
```

12.8.6

C

```
#
ois <- c(3, 4) #    AMI

#
ncs <- c(434165,436409,199192,4088290,4092879,44783954,75911,137951,77965,
         376707,4103640,73241,133655,73560,434327,4213540,140842,81378,
         432303,4201390,46269889,134438,78619,201606,76786,4115402,
         45757370,433111,433527,4170770,4092896,259995,40481632,4166231,
         433577,4231770,440329,4012570,4012934,441788,4201717,374375,
         4344500,139099,444132,196168,432593,434203,438329,195873,4083487,
         4103703,4209423,377572,40480893,136368,140648,438130,4091513,
         4202045,373478,46286594,439790,81634,380706,141932,36713918,
         443172,81151,72748,378427,437264,194083,140641,440193,4115367)

tcos <- create

##     {#studyOutputs}

                                       ATLAS     R                 R

###

                                     (ROC)    AUC       AUC 1
```

??Coho

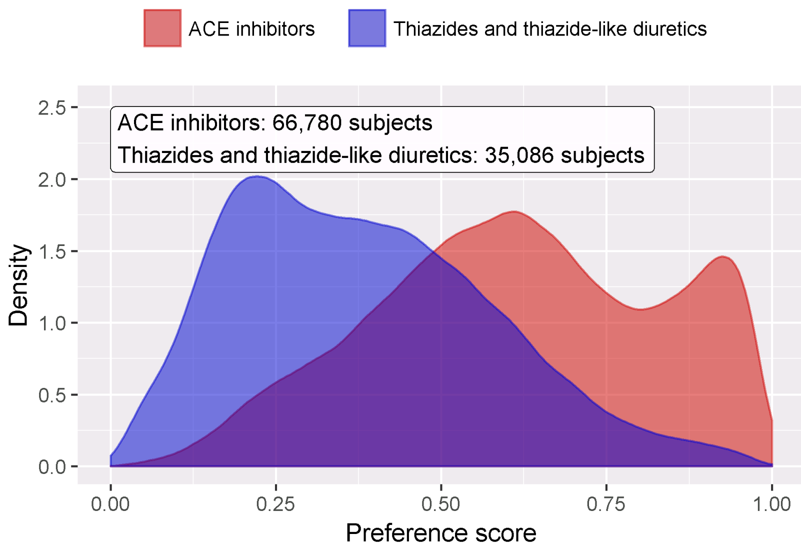


Figure 12.10: □□□□□□□□

□□□	□□□
-1.42	□□□□□-30□□□0□□□□□□□□□□□: □□
-1.11	□□□□□0□□□0□□□□□□□□□□□: □□□□□□
0.68	□□□□□□: 05-09
0.64	□□□□□-365□□□0□□□□□□□□□: □□□
0.63	□□□□□-30□□□0□□□□□□□□□□□: □□□
0.57	□□□□□-30□□□0□□□□□□□□□□□: □□□□□
0.55	□□□□□-365□□□0□□□□□□□□□□□: □□□□□□□□□□
-0.54	□□: □□□□□□□□□□□□□□□
0.52	(□□)
0.50	□□: □□

□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

12.8.7

PS□□□□□□□□□□2□□□
12.11□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□0.1□□□□□□□□□□□□□□□□□

12.8.8

□□□
12.12□□□□□□□`drawAttritionDiagram`□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

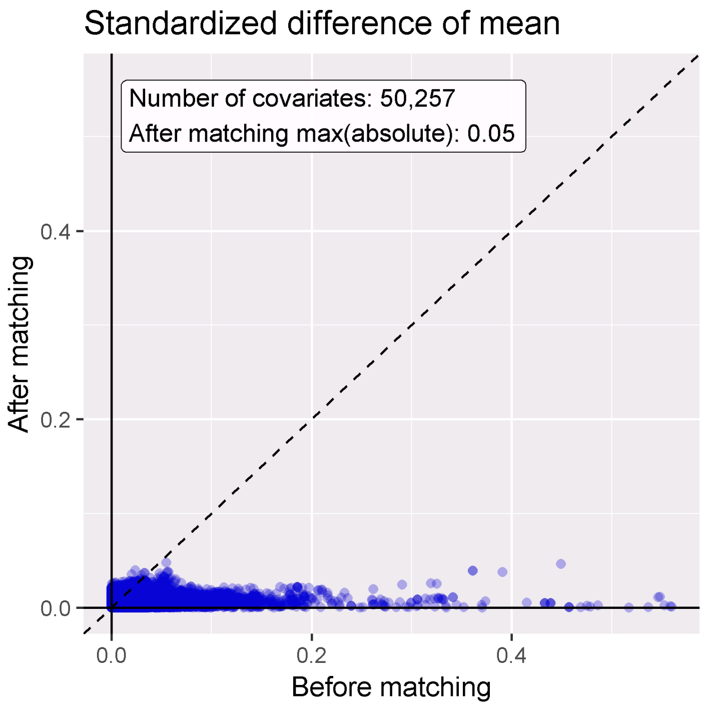


Figure 12.11: [illegible]

14[illegible]

12.9

- [illegible]
- [illegible]
- [illegible]
- OHDSI[illegible]

12.10

[illegible]

[illegible]R[illegible]R-Studio[illegible]Java[illegible] 8.4.5 [illegible]

```r
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
```

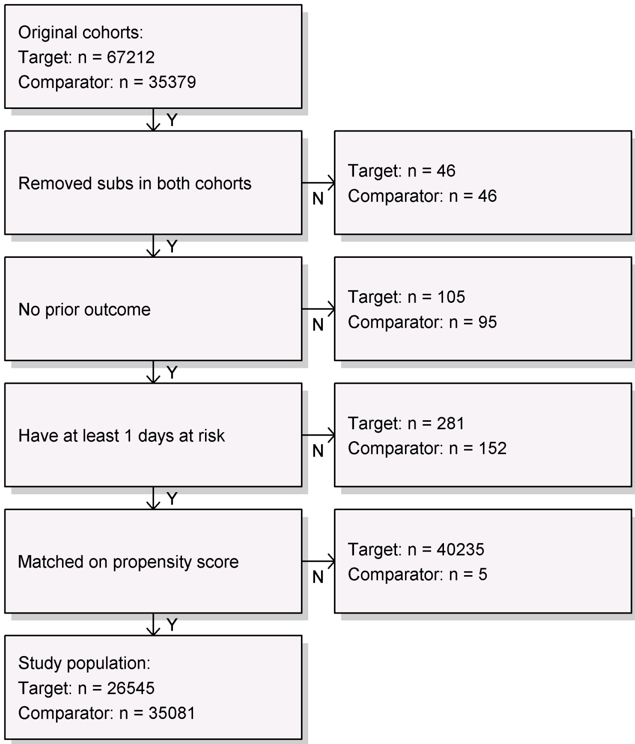


Figure 12.12: □□

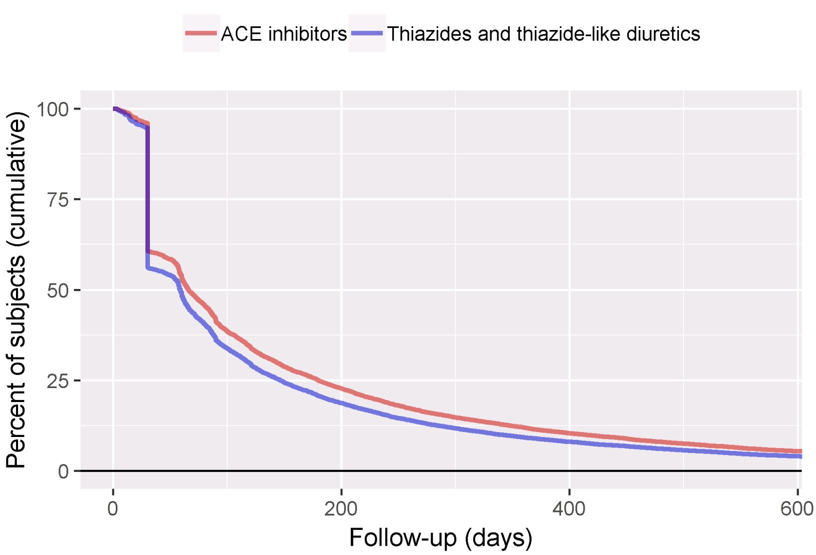


Figure 12.13: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

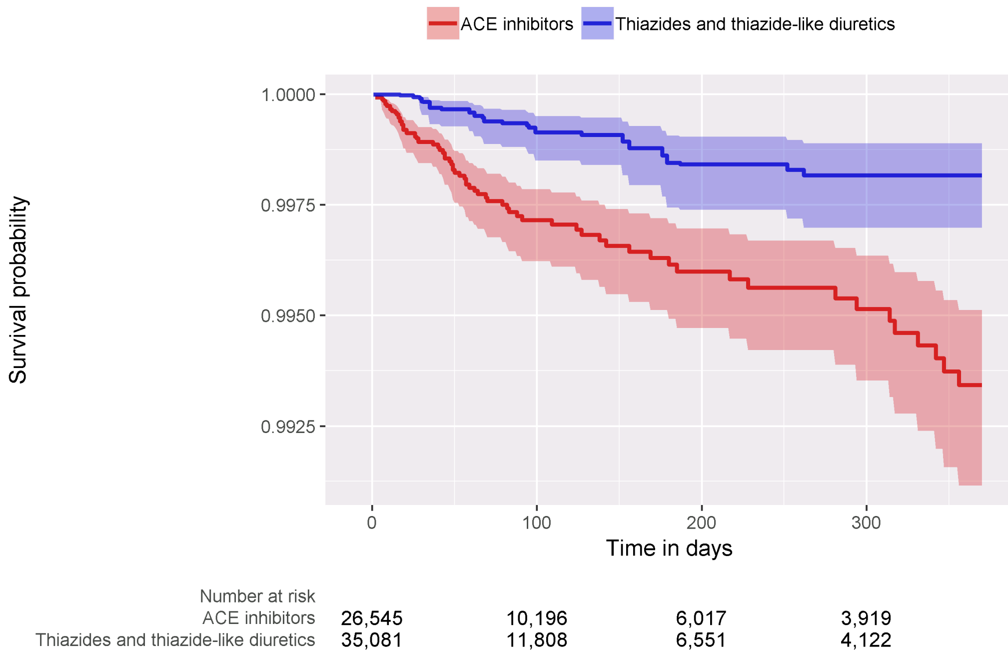


Figure 12.14: □□□□−□□□□□□□□□

```r
remotes::install_github("ohdsi/CohortMethod")
```

Eunomia□□□□□□□□□□□□R□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□

```r
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM□□□□□□□□□□□□□main□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Eu

```r
Eunomia::createCohorts(connectionDetails)
```

□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□GI□□□□□□□□□

□□□□□□□□□□□□□□□□COHORT_DEFINITION_ID□1□□□□□□□□□□□□□□□□□□

□□ **12.1.** CohortMethod R□□□□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□CohortM

□□ **12.2.** `createStudyPopulation`□□□□□□□□180□□□□□□□□□□□□□□□□□□□□□

□□ **12.3.** □□

□□ **12.4.** □□□□□□□□□□□□□□□□2□□□□□□□□□□□□

□□ **12.5.** 5□□□□□□□□PS□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ **12.6.** PS□□

□□□□□□□□□□ E.8 □□□□□□□□□□□□□

Chapter 13

13 – –

Lead: Peter Rijnbeek & Jenna Reps

10

OHDSI CDM

OHDSI (Reps et al., 2018,) PatientLevelPrediction R

13.1

13.1 t = 0

13.1 t=0

[ ] [ ]

:

[1] https://www.equator-network.org/reporting-guidelines/tripod-statement/

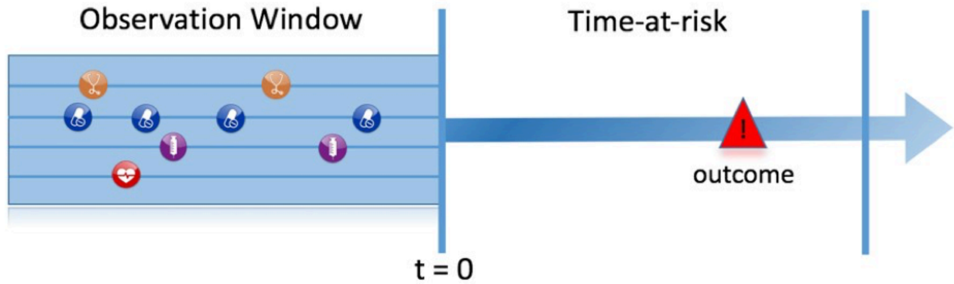


Figure 13.1: □□□□□

□□□	□□
□□□□□□□□□	□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□	□□□□□□□□□□□□□□□□□□□□□□□
□□□□□	t=0□□□□□□□□□□□□□□□□□□□□□□□□
□□□	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□
  - □□: *[□□]*□□□□□□□□□□□□□□□*[*□□□□□□□□□*]*□*[*□□□□□□□□□□□*]*□□□□□
  - □: □□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□□□□
- □□□□
  - □□: *[□□□□□□□□]*□□□□[□□1]□□□[□□2]□□□□□□□□□□□□[□□1]*□□□□□□□
  - □: □□□
- □□□□
  - □□: *[□□]*□□□□□□□□□□*[*□□□□*]*□*[*□□□□*]*□□□□□□□□□□□□
  - □: □□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□
- □□□□□
  - □□: *[□□]*□□□□□□□□□□*[*□□□□*]*□*[*□□□*]*□□□□□□□□□□□□
  - □: □□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□
- □□□□
  - □□: *[□□]*□□□□□□□□□□*[*□□□*]*□*[*□□□□*]*□□□□□□□□□□□□
  - □: □□□□□□□□□□□□□□□□□□□□1□□□80%□□□□□□□□□□□□□□□□□□□□

13.2

□□□

□□□□“□□□□”□“□□”□□□□”□□□□”□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
11□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□“□□□”□□□”□□□”□□□□□□□□□□□□□□□□□□□□□□

13.2.1

13.2.1 2 COHORT ID 1

COHORT COHORT_END_DATE

COHORT_DEFINITION_ID	SUBJECT_ID	COHORT_START_DATE
1	1	2000-06-01
1	2	2001-06-01
2	2	2001-07-01

13.2.1 CONDITION_OCCURRENCE Concept ID 320128

CONDITION_OCCURRENCE 3

PERSON_ID	CONDITION_CONCEPT_ID	CONDITION_START_DATE
1	320128	2000-10-01
2	320128	2001-05-01

1
1 0 ID 2 1
1 0 ID 2 1 1

13.2.2 vs

ID
1

13.3

13.2

2

13.2

Free Lunch 1

PatientLevelPrediction

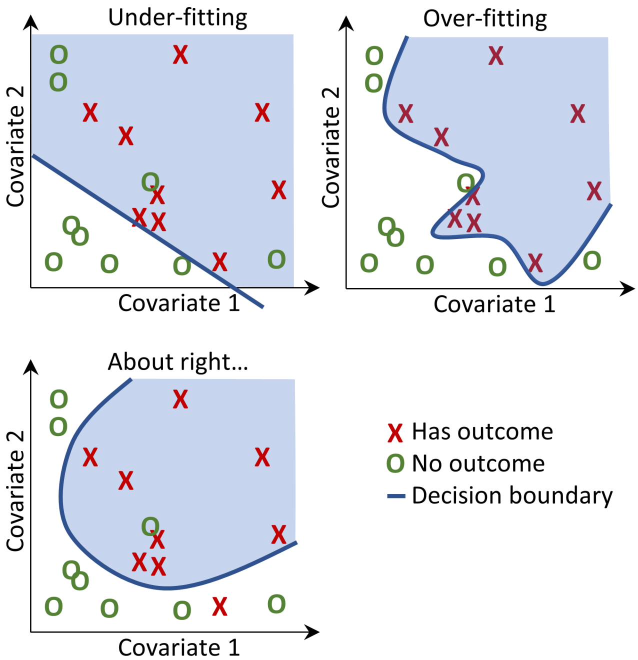


Figure 13.2: Decision boundary.

13.3.1

LASSO□□
Cyclops □□□

Table 13.4: □□□□□□□□□□□□□□□□□□□□□□□

□□□□□□	□□	□□□□□□
□□□□	□□□□□□□□□	0.1

□□□

13.3.2

□□□
R□□□□□□□□□□□□□□

Table 13.5: □□□□□□□□□□□□□□□□□□□□□□□

□□□□□□	□□	□□□□□□
earlyStopRound	□□□□□□□□□□□□□□□	25

□□□□□	□□	□□□□□
learningRate	□□□□□□□□□□□	0.005,0.01,0.1
maxDepth	□□□□□□	4,6,17
minRows	□□□□□□□□□□□□□□□□	2
ntrees	□□□	100,1000

13.3.3

□□□

Table 13.6: □□□□□□□□□□□□□□□□□□□□

□□□□□	□□	□□□□□
maxDepth	□□□□□□	4,10,17
mtries	□□□□□□	-1 = □□□□□□□□□,5,20
ntrees	□□□	500

13.3.4 K-

K-nearest neighbors (KNN) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Table 13.7: K-□□□□□□□□□□□□□□

□□□□□	□□	□□□□□
k	□□□	1000

13.3.5

□□□

13.3.6 AdaBoost

AdaBoost□□□
AdaboostClassifier□□□□□□□□□□□

Table 13.8: AdaBoost□□□□□□□□□□□

□□□□□	□□	□□□□□
nEstimators	□□□□□□□□□□□□□□□□□□□□	4
learningRate	□□□□□□□□□□□□□learning_rate□□□□□□□□□□□□learningRate□nE	1

□□□□□	□□	□□□□□

13.3.7

□□
DecisionTreeClassifier□□□□□□□□□□□

Table 13.9: □□□□□□□□□□□□□□

□□□□□	□□	□□□□□
classWeight	"Balance"□□□"None"	None
maxDepth	□□□□□□	10
minImpuritySplit	□□□□□□□□□□□□□□□□	10^-7□□□□□□□□□□□□□□□□□□□□□□
minSamplesLeaf	□□□□□□□□□□□□	10
minSamplesSplit	□□□□□□□□□□□	2

13.3.8

□□

Table 13.10: □□□□□□□□□□□□□□□□□□□□

□□□□□	□□	□□□□□
alpha	l2□□□	0.00001
size	□□□□□□□	4

13.3.9

□□
□□

13.3.10

□□□Patien
□□□□□□"Adding Custom Patient-Level Prediction Algorithms"□□□□□□□□□□□□

13.4

13.4.1

□□

> □□
> 13.3 □□□

□□□□□□□□□□□□□□□□□□□□

- □□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□2□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□
  - □□
  - □□□□□□□□□□□□□□□□□□□□□
- □□□n□□□□□□□□□□□□ 10□□

□□□

13.4.2

□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□0□□1□□□□□□□□□□□□□□□□□□ 13.11 □□□□□□□□□□0.5□□□□□□□□□□1□3□7□□□□□10□□□□0.5□□□□□□□□□□□□□□□□□□□□

Table 13.11: □□□□□□□□□□□□□□□□□□□□

□□ID	□□□□□	0.5□□□□□□□□□□	□□□□□□□□□□□□□□□	□□□
1	0.8	1	1	TP
2	0.1	0	0	TN
3	0.7	1	0	FP
4	0	0	0	TN
5	0.05	0	0	TN
6	0.1	0	0	TN
7	0.9	1	1	TP
8	0.2	0	1	FN
9	0.3	0	0	TN
10	0.5	1	0	FP

□□□TP□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□

- □□: $(TP + TN)/(TP + TN + FP + FN)$

- □□: $TP/(TP+FN)$
- □□□: $TN/(TN+FP)$
- □□□□□: $TP/(TP+FP)$

□□
FN□□□

□□□

□□
- □□□□y□□□□□□□□□□□□□□□□□□□□□□□□□ROC□□□□□□□□□□□□□□□□□□□□□□□
13.11 □□□□□□□□□□□□□□□□□□□□□□□□□□□□AUC□□□□□□□□□□□□□□□□□□□□□□□□□□0.5□□□□

AUC□□
13.3 □□□□□□□□□□□□□□

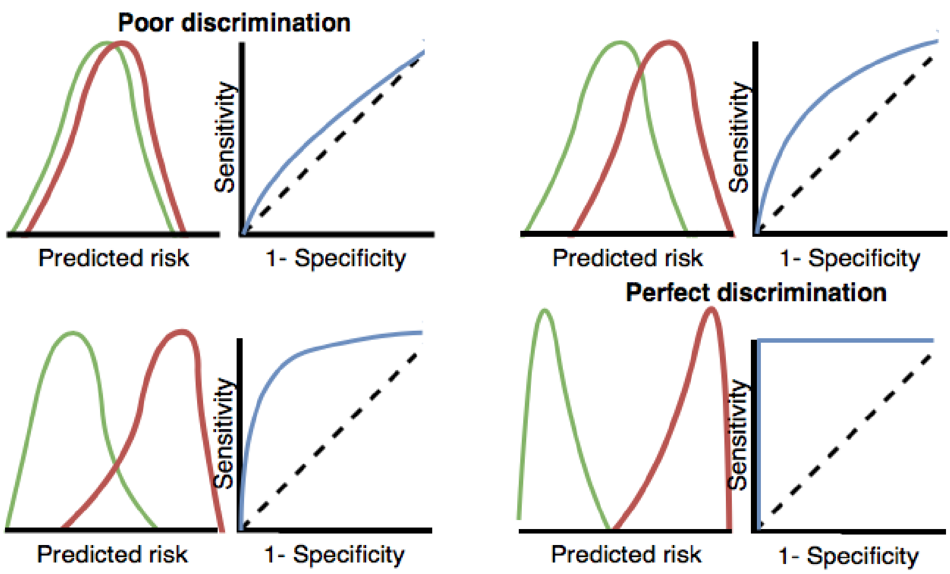


Figure 13.3: □□□□□□□□ROC□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□ROC□□□□□

□□□□□□□□□□□□□□□□AUC□□□
□□□□□□□□□□AUPRC□□□□□□□□□□□□□□□□□□□□□AUPRC□□□□□x□□□□□□□□□□□□

□□□□□□□□□

□□100□□□□□10□□□□□□
=y□□
13.12 □□1□
Calibration Curves□□□□□□□□□

13.5

13.1

13.5.1

ACE ACE
(Byrd et al., 2006,)
et al., 2013,)
et al., 2013; Thompson and Frable, 1993,) ACE
et al., 2006,)
et al., 2004,) (O'Mara
and O'Mara, 1996,)

ACE

13.5.2

-
-
-
- 2
-

13.5.3

13.3
Boosting Machines GBM

13.5.4

-25

13.5.5

□□□□□ 13.5.5□□□□□□□□□□□□□□□□□□□□□□

□: □□□□□□□□□□□□□□□□□□□□□

□□	□
□□□□□□□□□□□□□□□□□□□□	□□□ACE□□□□□□□□□□□□□□□□□□□□365□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□
□□□□□	□□□□□□□□1□□□365□□□□□□□364□□□□□□□□□□□
□□□	Gradient Boosting Machine with hyper-parameters ntree: 5000, max depth: 4 or 7 or 10 and learning rate: 0.001 or 0.01 or 0.1 or 0.9. Covariates will include gender, age, conditions, drugs, drug groups, and visit count. □□□□□: 75□□□□□□□□ - 25□□□□□□□□□□□□□□□□□□□□□□□□□□

13.6 ATLAS

□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□ 

□□□□□□□□□□□□□□□□□□□

□□4□□□□□□□□□□□□□□□□□□□□

13.6.1

□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 10□□□□□□□□□□□□□□□□□□□□□□□□□□□Appendix B.1□□□□□□□□□□□□□□□□Appendix B.4□□Add Target Cohort□□□Add Outcome Cohort□□□ 13.4□□□□□□□□□

13.6.2

□□

□□□□□

□□□□□□□□□□□□□1□□□ Model Settings□□□□□□□□□□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□13.5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

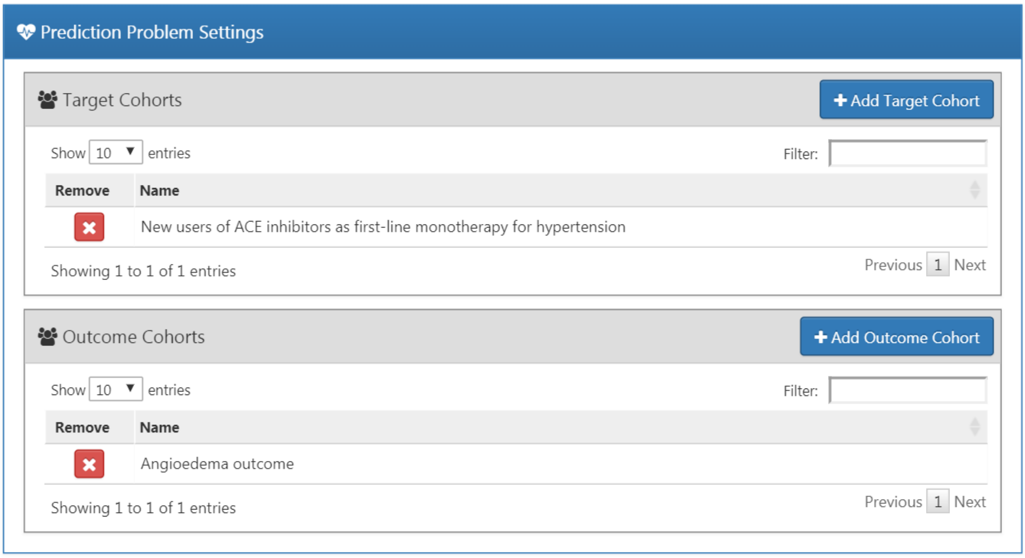


Figure 13.4: □□□□□□□.

□□□□□

CDM□□

□□□□□□□□□□□□□□□□□Add Covariate Settings□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ **concepts do you want to include in baseline covariates in the patient-level prediction model? (Leave blank if you want to include everything)**□□□□ □□□□□□□□□□□□□□□□ **descendant concepts be added to the list of included concepts?**□□□□□□yes□□□□□□□□□□□ **concepts do you want to exclude in baseline covariates in the patient-level prediction model? (Leave blank if you want to include everything)**□□□□□□□□□□□□□□□□□□□□□ **comma delimited list of covariate IDs that should be restricted to** □□□□□□□□ID□□□ID□□□ 13.6□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□: □□□□□□□□□□□□□□□□□
- □□: □□□□□□□□
- □□□□□□: 5□□□□□□□□□□□0-4□5-9□…□95+□
- □□: □□□□□□□□□□
- □□: □□□□□□□□□□
- □□□□□□□: □□□□□□□□□□□□□□□□
- □□□□□□□: □□□□□□□□□□□□□□□□
- □□□□□□□: [□□□□□□□□□□□□] □□□□□□□□□□
- □□□□□□□: [□□□□□□□□□□□□] □□□□□□□□□□□
- □□□□□□: □□□□□□□□□□□□
- □□□□□□□□□: [□□□□□□□□□□□□] □□□□□□□□□□□□□□□□

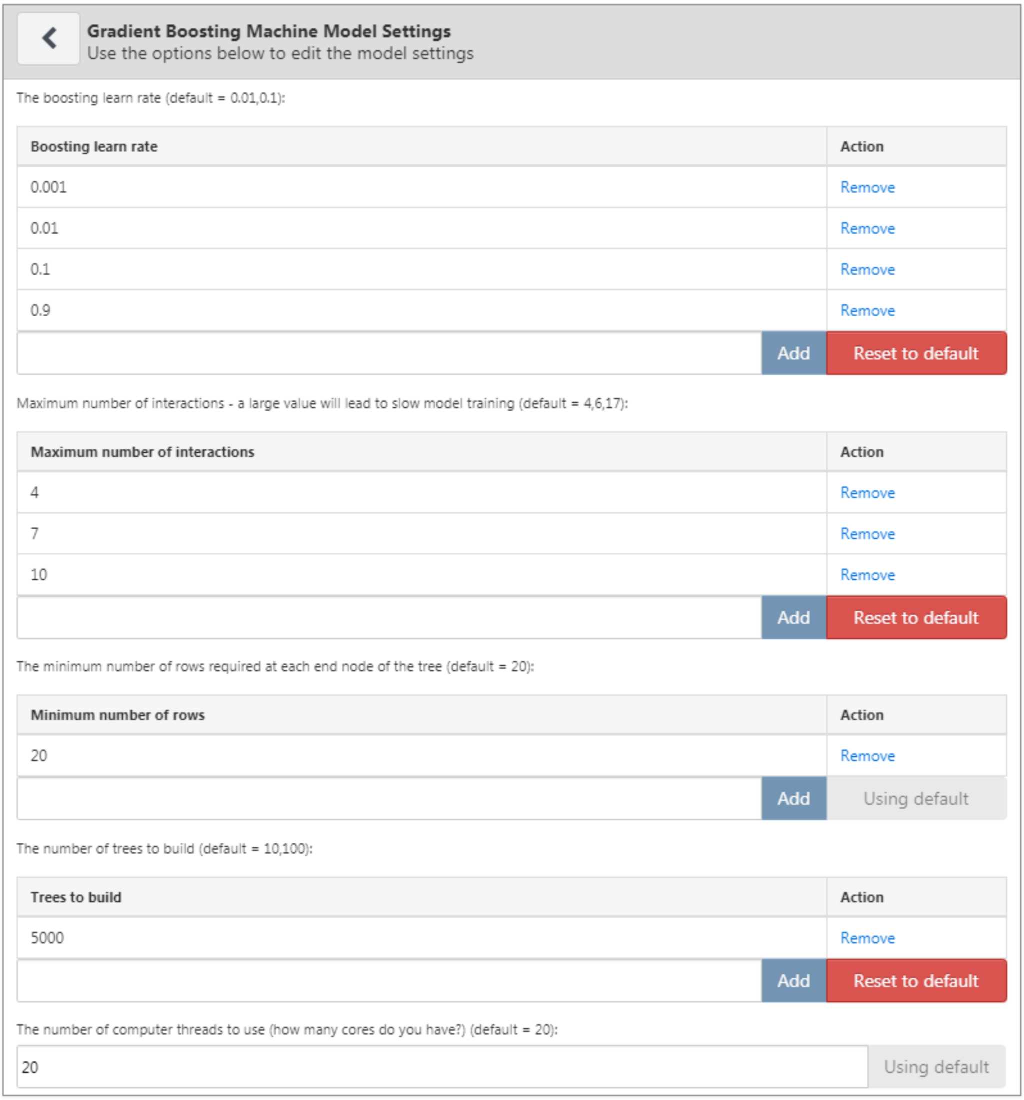

Gradient Boosting Machine Model Settings
Use the options below to edit the model settings
The boosting learn rate (default = 0.01,0.1):
Boosting learn rate
Action
0.001
Remove
0.01
Remove
0.1
Remove
0.9
Remove
Add
Reset to default
Maximum number of interactions - a large value will lead to slow model training (default = 4,6,17):
Maximum number of interactions
Action
4
Remove
7
Remove
10
Remove
Add
Reset to default
The minimum number of rows required at each end node of the tree (default = 20):
Minimum number of rows
Action
20
Remove
Add
Using default
The number of trees to build (default = 10,100):
Trees to build
Action
5000
Remove
Add
Reset to default
The number of computer threads to use (how many cores do you have?) (default = 20):
20
Using default

Figure 13.5: □□□□□□□□□□□□□□

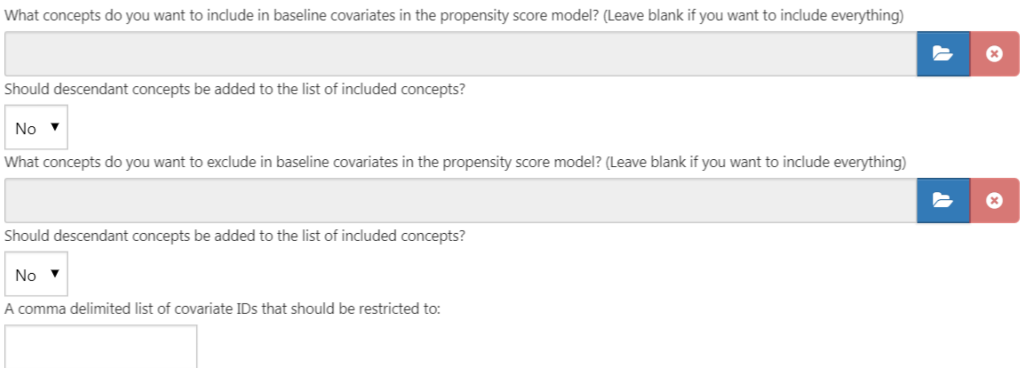


Figure 13.6: □□□□□□□□□□□□.

□□□□□□□□□□□□□□□□□□ 13.7□□□□□□□□□□□

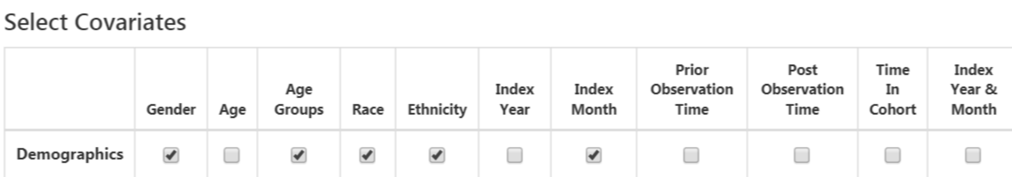


Figure 13.7: □□□□□□.

□□□□□□□□□□□□□3□□□□□□□□□□□□□□

- □□□: □□□□□□□□□□□□□[□□□□□□0□]
- □□ [□□□□□-365□□□□□□□□□□□□]
- □□ [□□□□□-180□□□□□□□□□□□□]
- □□ [□□□□□-30□□□□□□□□□□□□]

□□□□□□□□□□□□□□□□□□ 13.8□□□□□□□□□□□

Figure 13.8: □□□□□□□□□□.

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- Condition: □□□□□□□□□□□□□□□□□□□□□□ID□□□□□□□□□□□□□□□□□
- Condition group: □□□□□□□□□□□□□□□□□□ID□□□□□□□□□ID□□□□□□□
- Drug: □□□□□□□□□□□□□□□□□ID□□□□□□□□□□□□□□□ID□□□□□1□□
- Drug group: □□□□□□□□□□□ID□□□□□□□

□□
□□□□□□□□□□□□□□□□□□ 13.9□□□□□□□□□□□


Set the time bound era covariates

Domain	Any Time Prior	Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)	Overlapping	Era Start Long Term (-365 days)	Era Start Medium Term (-180 days)	Era Start Short Term (-30 days)
Condition	☐	☐	☐	☐	☐	☐	☐	☐
Condition Group	☐	☑	☐	☑	☐	☐	☐	☐
Drug	☐	☐	☐	☐	☐	☐	☐	☐
Drug Group	☐	☑	☐	☑	☐	☐	☐	☐

Figure 13.9: □□□□□□□.

□□□□□□□□□□□□□□□□□□□□□□ID□□□□□□□□□□□□□□

- Condition: □□□□□□□ID□□□□□□□□□
- Condition Primary Inpatient: □□□□□□□□□□□
- Drug: □□□□□
- Procedure: □□□□□□
- Measurement: □□□□□

13.7 R

□□□□□□□ATLAS□□□□□□□□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□PatientI
CDM□□□

13.7.1

□□
10□□□□□□□□□□Appendix□□□□□□□□□□□(Appendix B.1)□□□□□□□□□□(Appendix
B.4)□□□□□□□□□□□□□□□□□□ACE□□□□□□□□ID□1□□□□□□□□□□ID□2□□□□□□□

13.7.2

□□□R□□□□□□□□□□□□□□□□□□□□□PatientLevelPrediction□□DatabaseConnector

```
library(PatientLevelPrediction)
connDetails <- createConnectionDetails(dbms = "postgresql",
                                       server = "localhost/ohdsi",
                                       user = "joe",
                                       password = "supersecret")

cdmDbSchema <- "my_cdm_data"
cohortsDbSchema <- "scratch"
```

```
cohortsDbTable <- "my_cohorts"
cdmVersion <- "5"
```

□□□4□□cdmDbSchema□cohortsDbSchema□cohortsDbTable□□□□□□□CDM□□□□□□
SQL Server□□
<- "my_cdm_data.dbo"□□□□□□□□□□

□□

```
sql <- paste("SELECT cohort_definition_id, COUNT(*) AS count",
"FROM @cohortsDbSchema.cohortsDbTable",
"GROUP BY cohort_definition_id")
conn <- connect(connDetails)
renderTranslateQuerySql(connection = conn,
                        sql = sql,
                        cohortsDbSchema = cohortsDbSchema,
                        cohortsDbTable = cohortsDbTable)
```

```
##   cohort_definition_id  count
## 1                    1 527616
## 2                    2   3201
```

PatientLevelPrediction□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Featur

```
covariateSettings <- createCovariateSettings(
  useDemographicsGender = TRUE,
  useDemographicsAge = TRUE,
  useConditionGroupEraLongTerm = TRUE,
  useConditionGroupEraAnyTimePrior = TRUE,
  useDrugGroupEraLongTerm = TRUE,
  useDrugGroupEraAnyTimePrior = TRUE,
  useVisitConceptCountLongTerm = TRUE,
  longTermStartDays = -365,
  endDays = -1)
```

□□□□□□□□□□□□□□getPlpData□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
plpData <- getPlpData(connectionDetails = connDetails,
                      cdmDatabaseSchema = cdmDbSchema,
                      cohortDatabaseSchema = cohortsDbSchema,
                      cohortTable = cohortsDbSchema,
                      cohortId = 1,
                      covariateSettings = covariateSettings,
                      outcomeDatabaseSchema = cohortsDbSchema,
                      outcomeTable = cohortsDbSchema,
                      outcomeIds = 2,
```

```
                    sampleSize = 10000
)
```

getPlpData□□□□□□□□□□□□□□□□□□□□□□□□□□□PatientLevelPrediction□□□□□□□□

plpData□□

```
savePlpData(plpData, "angio_in_ace_data")
```

□□□□□□□□□□□□□□□□□□□□□□□□□loadPlpData()□□□□□□□□□□

13.7.3

□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
= 30□riskWindowEnd = 365□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
= TRUE□□□

□□□

```
population <- createStudyPopulation(plpData = plpData,
                                    outcomeId = 2,
                                    washoutPeriod = 364,
                                    firstExposureOnly = FALSE,
                                    removeSubjectsWithPriorOutcome = TRUE,
                                    priorOutcomeLookback = 9999,
                                    riskWindowStart = 1,
                                    riskWindowEnd = 365,
                                    addExposureDaysToStart = FALSE,
                                    addExposureDaysToEnd = FALSE,
                                    minTimeAtRisk = 364,
                                    requireTimeAtRisk = TRUE,
                                    includeAllOutcomes = TRUE,
                                    verbosity = "DEBUG"
)
```

13.7.4

□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Gradient Boosting Machine□□□□□□□□□□□□ntr
= c(100,200), maxDepth = 4□□□□□□□□□□□□□ntrees = 100□□□maxDepth
= 4□□□□ntrees = 200□□□maxDepth = 4□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
gbmModel <- setGradientBoostingMachine(ntrees = 5000,
                                       maxDepth = c(4,7,10),
                                       learnRate = c(0.001,0.01,0.1,0.9))
```

`runPlP`□□□□□□`plpData`□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□75
25%□□□□□□□□□□□□□□□□□□□□□□□□□□□`testSplit`□□/□□□□□□`testFractio`

```r
gbmResults <- runPlp(population = population,
                     plpData = plpData,
                     modelSettings = gbmModel,
                     testSplit = 'person',
                     testFraction = 0.25,
                     nfold = 2,
                     splitSeed = 1234)
```

□□□□□□□□□□□□R□xgboost□□□□□□□□□□□75%□□□□□□□□□□□□□□□□□□□□

`runPlp`□□□□□`plpData`□`plpResults`□`plpPlots`□`evaluation`□□□□□□□□□□□□□□□

□□□□□□□□□□□

```r
savePlpModel(gbmResults$model, dirPath = "model")
```

□□□□□□□□□□□□

```r
plpModel <- loadPlpModel("model")
```

□□□□□□□□□□□□□□□□□□□□□□□

```r
savePlpResult(gbmResults, location = "gbmResults")
```

□□□□□□□□□□□□□□□□□□□

```r
gbmResults <- loadPlpResult("gbmResults")
```

13.7.5

□□□□□□□□□`runPlp`□□□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□
`= gbmResults)`□□□□□□□□□□□□Shiny□□□□□□□□□□□□□□□□□□□□□□□□□□□
13.10□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```r
plotPlp(gbmResults, "plots")
```

□□□□□□□□□□□□□Section 13.4.2□□□□□□□□□□

13.7.6

```
#
plpModel <- loadPlpModel("model")

#   plpData
plpData <- loadPlpData("newData")

population <- createStudyPopulation(plpData = plpData,
                                    outcomeId = 2,
                                    washoutPeriod = 364,
                                    firstExposureOnly = FALSE,
                                    removeSubjectsWithPriorOutcome = TRUE,
                                    priorOutcomeLookback = 9999,
                                    riskWindowStart = 1,
                                    risk

##

###

               `viewPlp`                     R       `runPLp`           ATLAS
```

```
plpResult <- loadPlpResult(file.path(outputFolder,
                                     'Analysis_1',
                                     'plpResult'))
```

Analysis_1

Shiny

```
viewPlp(plpResult)
```

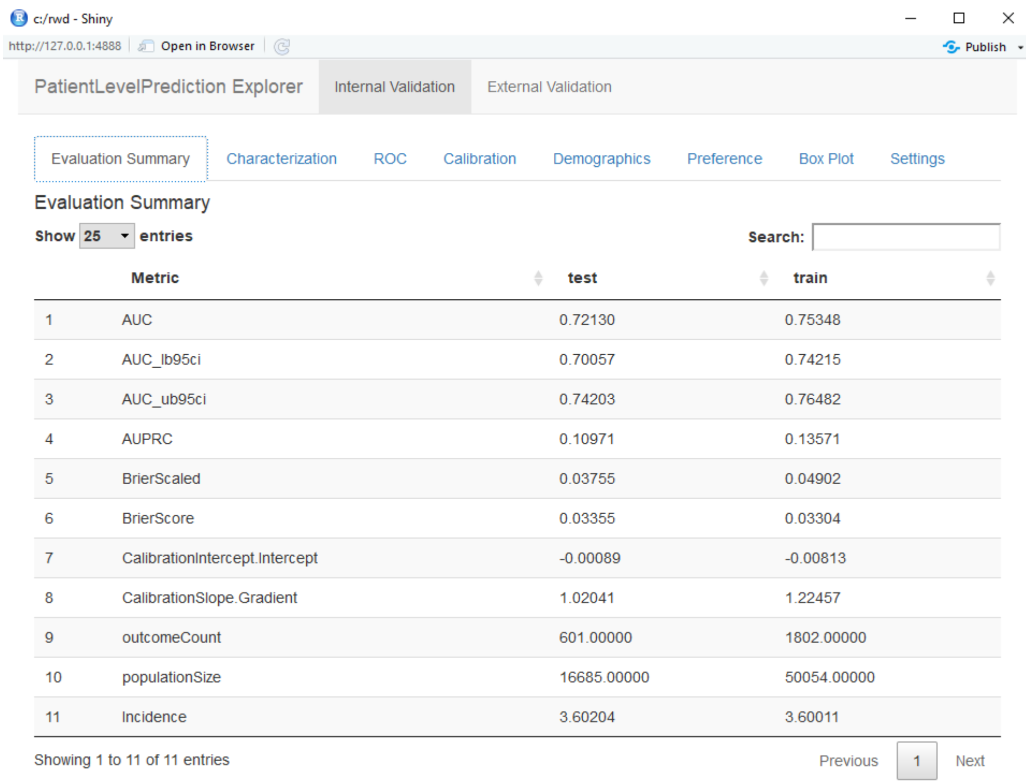
Shiny
13.10 AUC 0.78 0.74
13.11

13.12
13.13 40

attrition
13.14

13.7.7

ATLAS



	Metric	test	train
1	AUC	0.72130	0.75348
2	AUC_lb95ci	0.70057	0.74215
3	AUC_ub95ci	0.74203	0.76482
4	AUPRC	0.10971	0.13571
5	BrierScaled	0.03755	0.04902
6	BrierScore	0.03355	0.03304
7	CalibrationIntercept.Intercept	-0.00089	-0.00813
8	CalibrationSlope.Gradient	1.02041	1.22457
9	outcomeCount	601.00000	1802.00000
10	populationSize	16685.00000	50054.00000
11	Incidence	3.60204	3.60011

Figure 13.10: Shiny□□□□□□□□□□□□□□□□□□□□

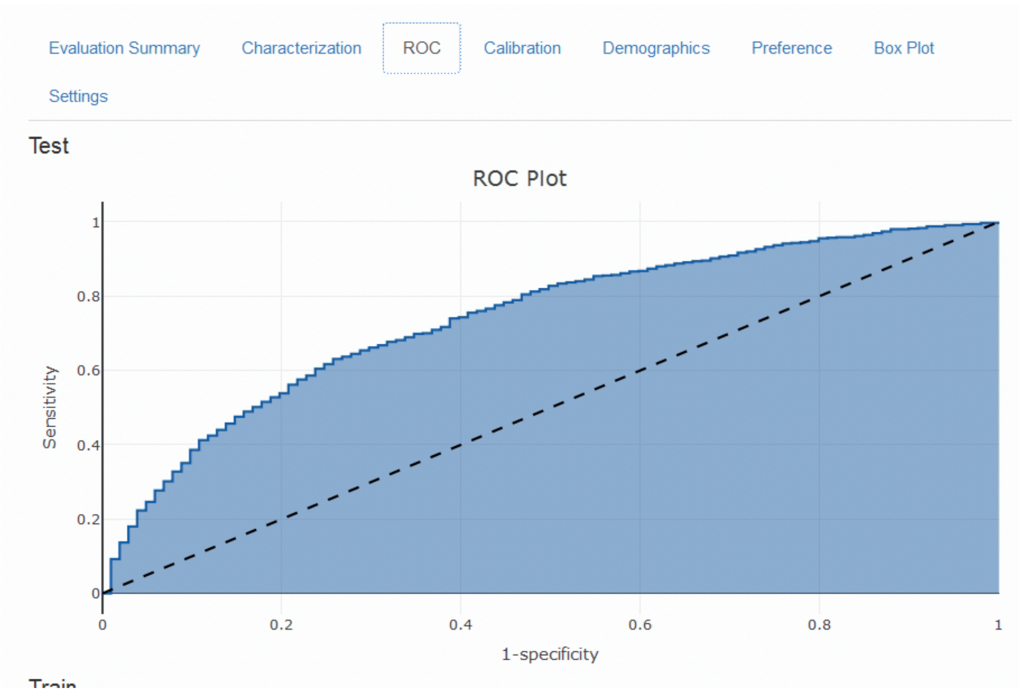


Figure 13.11: ROC□□□□□

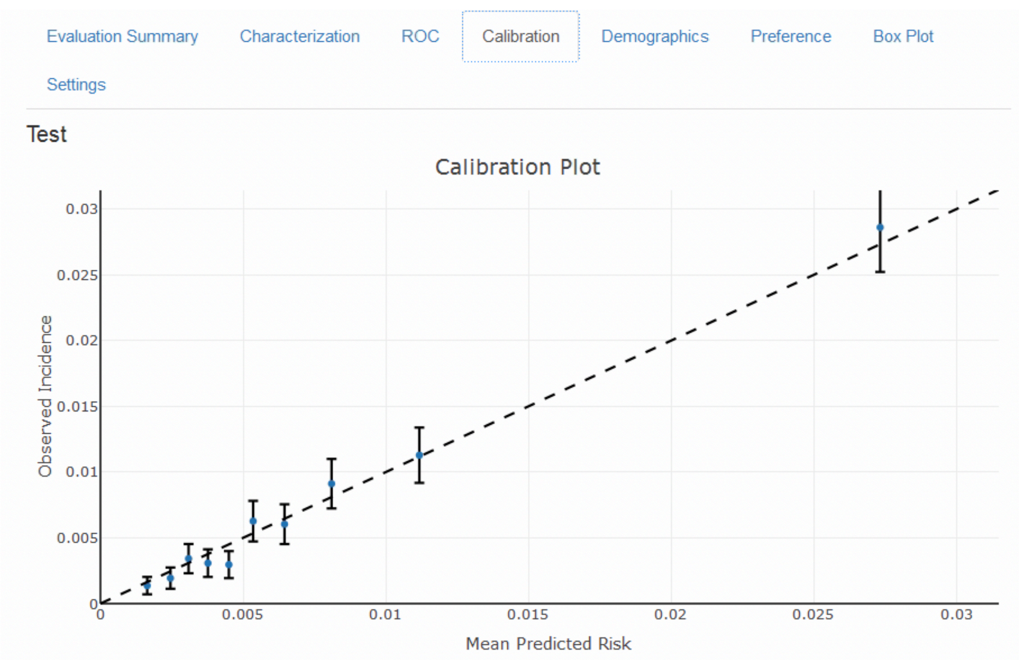


Figure 13.12: □□□□□□□□□□□□□

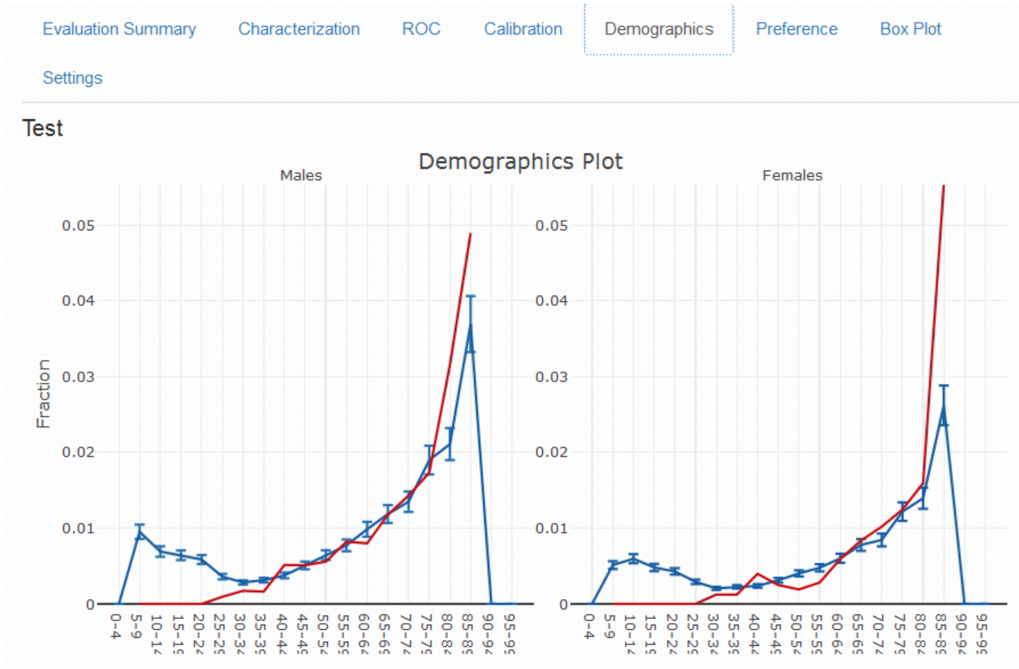
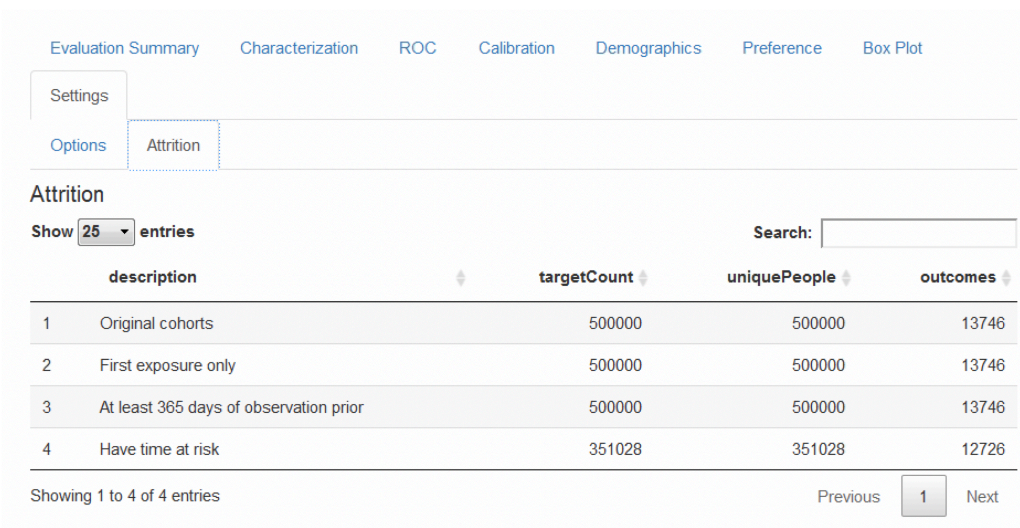


Figure 13.13: □□□□□□□□□□□□□□□□□□□□

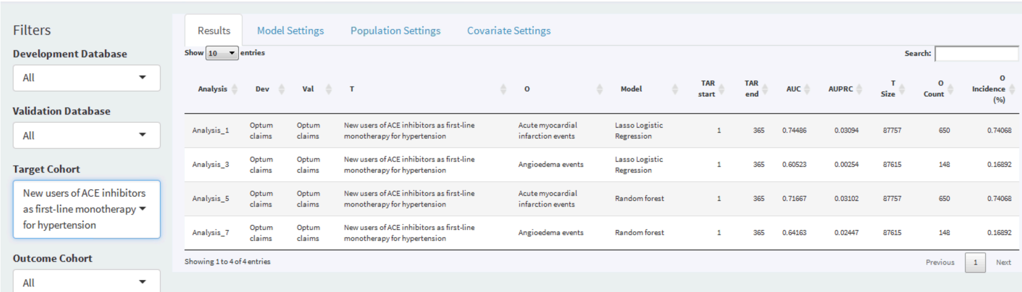


	description	targetCount	uniquePeople	outcomes
1	Original cohorts	500000	500000	13746
2	First exposure only	500000	500000	13746
3	At least 365 days of observation prior	500000	500000	13746
4	Have time at risk	351028	351028	12726

Figure 13.14: □□□□□□□□attrition□□□□

□□□□□□□□□□□□

□□□□□□□□□Shiny□□□□□□ 13.15□□□□□□□□□□□□□□□



Analysis	Dev	Val	T	O	Model	TAR start	TAR end	AUC	AUPRC	T Size	O Count	O Incidence (%)
Analysis_1	Optum claims	Optum claims	New users of ACE inhibitors as first-line monotherapy for hypertension	Acute myocardial infarction events	Lasso Logistic Regression	1	365	0.74486	0.03094	87757	650	0.74068
Analysis_3	Optum claims	Optum claims	New users of ACE inhibitors as first-line monotherapy for hypertension	Angioedema events	Lasso Logistic Regression	1	365	0.60523	0.00254	87615	148	0.16892
Analysis_5	Optum claims	Optum claims	New users of ACE inhibitors as first-line monotherapy for hypertension	Acute myocardial infarction events	Random forest	1	365	0.71667	0.03102	87757	650	0.74068
Analysis_7	Optum claims	Optum claims	New users of ACE inhibitors as first-line monotherapy for hypertension	Angioedema events	Random forest	1	365	0.64163	0.02447	87615	148	0.16892

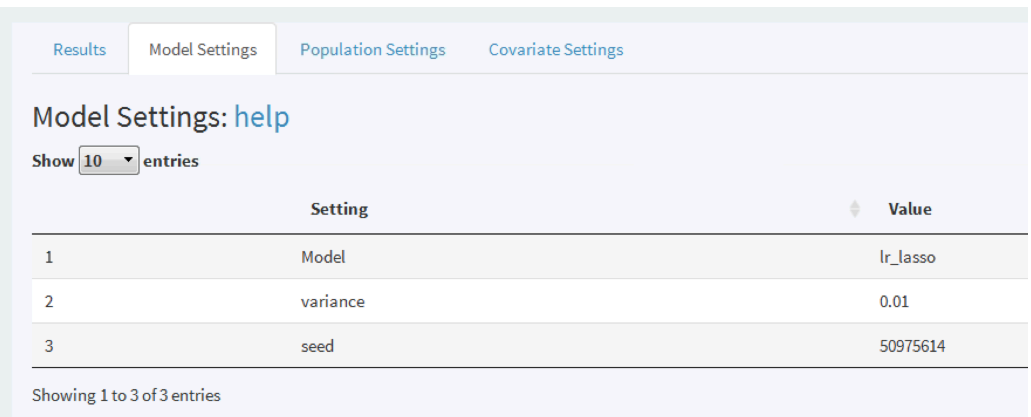
Figure 13.15: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Shiny□□□□□□

□□□□□□□□□□□□□□□□□□□□□□

- □□
- □□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□AUC□AUPRC

□□□□□□□□□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□Settings□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□



	Setting	Value
1	Model	lr_lasso
2	variance	0.01
3	seed	50975614

Figure 13.16: □□□□□□□□□□□□□□□□□□□□□□

□□□

□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 

□□□□□□□□□□□□□□□□13.17□□□□□

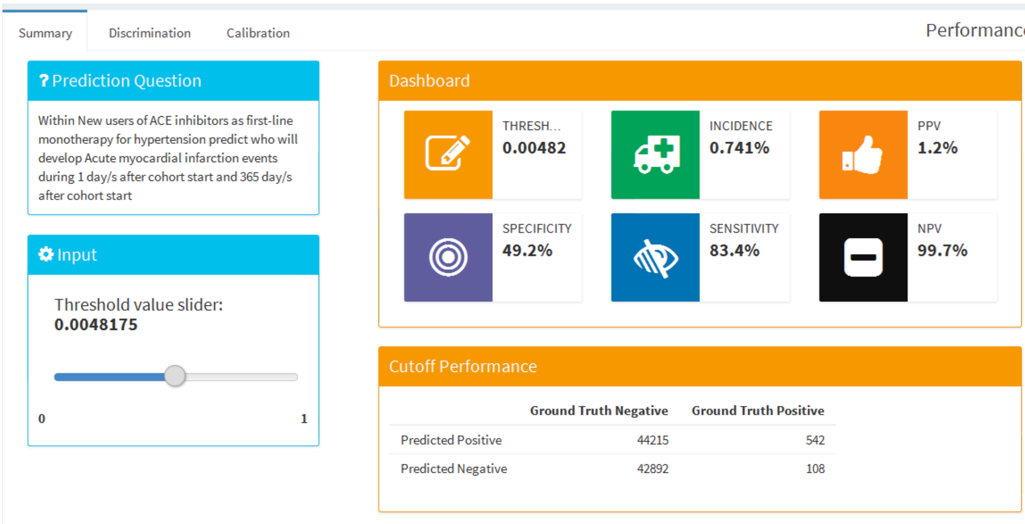


Figure 13.17: □□□□□□□□□□□□□□□□□□□□□□□□

□□PPV□□□□□□□ 13.4.2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 13.17□□□□□□□

□□□□□□□□□□□□□□□□□Discrimination□□□□□□□□□ROC□□□□□□□-□□ 13.18□ROC□□□□□-□□□□□□□□□□□□□□ROC□□□

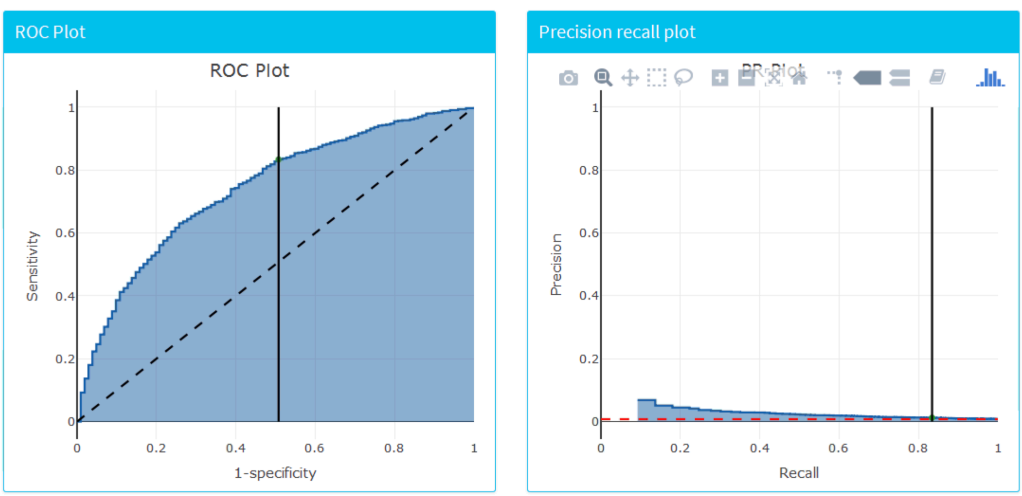


Figure 13.18: ROC□□□□□-□□□□□□□□□□□□□□□□□□□□□□□□□□

□ 13.19□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□Calibration□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

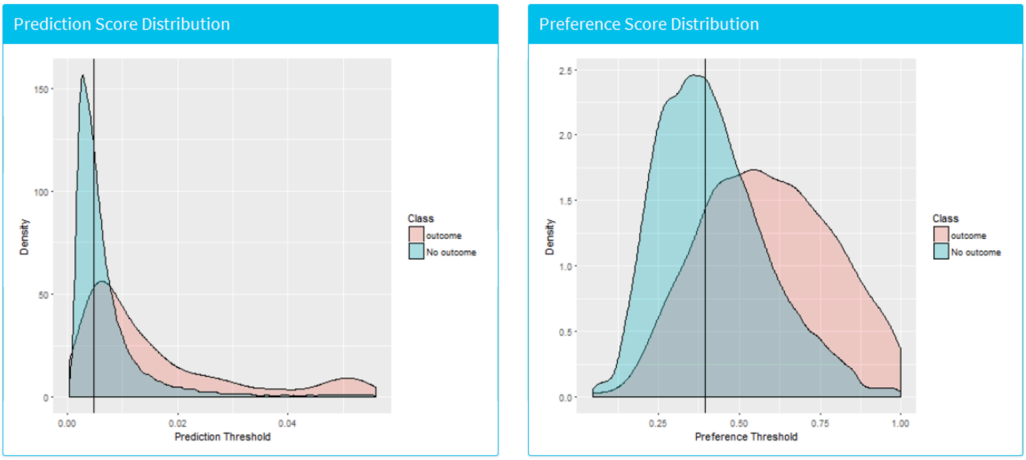


Figure 13.19: □□□□□□□□□□outing□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

13.20□□

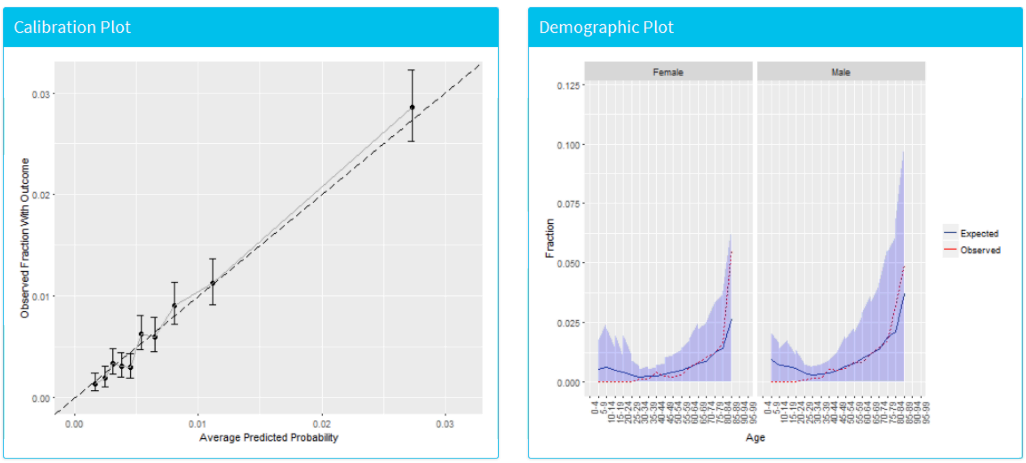
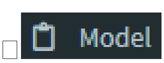


Figure 13.20: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

1□□

□□□□□□

□□□□□□□□□□□□□□□□□□□□□□  □□□□□□□□□□□□□□□□□□□
13.21□□□□□□□□□□□□□□□□□□ 13.22□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□ 13.22□□

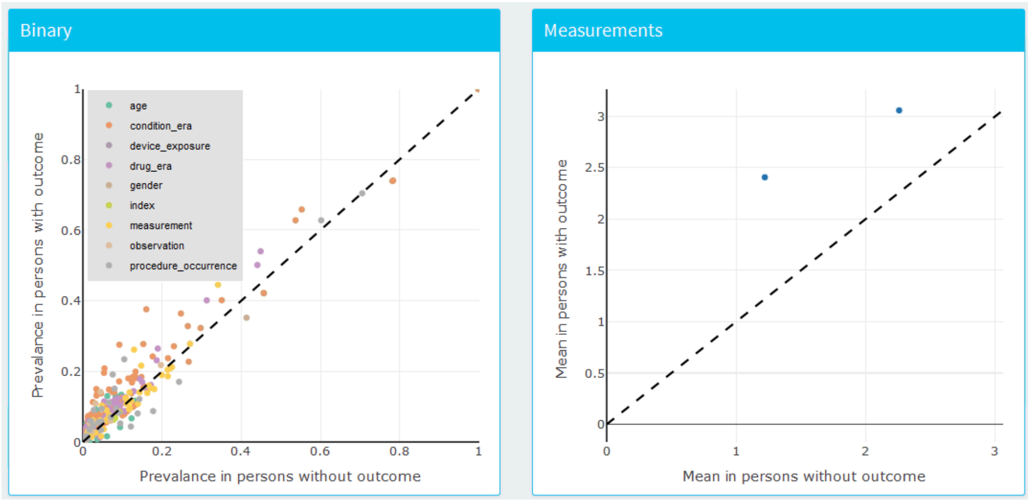


Figure 13.21: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Model Table

Download Model

Show 10 entries

Search:

	Covariate Name	Value	Outcome Mean	Non-outcome Mean
1	age group: 00-04	0	0.0004	0.0001
2	age group: 05-09	0	0	0.0003
3	index month: 1	0	0.1307	0.1096
4	observation during day -365 through 0 days relative to index: Domain	0	0.1188	0.0514
5	Charlson index - Romano adaptation	0	2.4783	1.3817
6	Diabetes Comorbidity Severity Index (DCSI)	0.1478	2.4056	1.2207
7	CHADS2VASc	0.9279	3.0573	2.2576
8	visit_occurrence concept count during day -365 through 0 concept_count relative to index	0	19.5263	13.8837
9	age group: 10-14	0	0	0.001
10	index month: 2	0	0.0934	0.0909

Showing 1 to 10 of 67,897 entries   Previous 1 2 3 4 5 … 6790 Next

Figure 13.22: □□□□□□□□□□

13.22

13.8

13.8.1

1

```r
createPlpJournalDocument(plpResult = <your plp results>,
            plpValidation = <your validation results>,
            plpData = <your plp data>,
            targetName = "<target population>",
            outcomeName = "<outcome>",
            table1 = F,
            connectionDetails = NULL,
            includeTrain = FALSE,
            includeTest = TRUE,
            includePredictionPicture = TRUE,
            includeAttritionPlot = TRUE,
            outputLocation = "<your location>")
```

13.9

-
-
- PatientLevelPrediction OMOP-CDM
-
- OHDSI

13.10

8.4.5 R R-Studio Java

```
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
remotes::install_github("ohdsi/PatientLevelPrediction")
```

Eunomia□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□R□□□□□□

```
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM□□□□□□□□□□□□□main□□□□□□□□□□□□□□□□□□□□□□□□□□□Eunomia

```
Eunomia::createCohorts(connectionDetails)
```

□□□□

□□□NSAIDs□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□GI□□□□
NSAID□□□□□□□□□□□□COHORT_DEFINITION_ID□4□□□GI□□□□□□□COHORT_D

□□ **13.1.** PatientLevelPrediction R□□□□□□□□□□□□□□□□□□□□□□□□□CDM□□PLP

□□ **13.2.** □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□createStudyPopulati

□□ **13.3.** LASSO□□□□□□□□□□□□□□Shiny□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□ E.9□□□□□□□□□□□□

Part II

— —

Chapter 14

: Patrick Ryan & Jon Duke

14.1

Desired attribute	Question	Researcher	Data	Analysis		Result
Repeatable	Identical	Identical	Identical	Identical	=	Identical
Reproducible	Identical	Different	Identical	Identical	=	Identical
Replicable	Identical	Same or different	Similar	Identical	=	Similar
Generalizable	Identical	Same or different	Different	Identical	=	Similar
Robust	Identical	Same or different	Same or different	Different	=	Similar
Calibrated	Similar (controls)	Identical	Identical	Identical	=	Statistically consistent

Figure 14.1:

Austin Bradford Hill
1965)

et al., 2013a)

et al., 2016, 2018a,b)

14.2

□□□

□□□(Bots
et al., 2010; Hersh et al., 2013; Sherman et al., 2016) □□□
et al., 2012; Liaw et al., 2013; Weiskopf and Weng, 2013) OHDSI□□□
CDM□□□OHDSI□□□(Huser et al., 2016; Kahn et al., 2015; Callahan et al., 2017; Yoon et al., 2016)

□□10□□□
McDonald□□□

□□□

□□□-□□□

□□□

□□□ 14.2 □□□4□□□

□: □□□4□□□

□□□	□□□
□□□	□□□
□□□	□□□
□□□	□□□
□□□	□□□

14.3

□□□OHD
□□□

- □□□ □ □□□ □ □□□ □ □□□ □ □□ □□□
- □□□
  * □□□
  * □□□
  * □□□
  * □□□
- □□□

Chapter 15

— —

□□□□□: *Martijn Schuemie, Vojtech Huser & Clair Blacketer*

□□□
Lei (1991) □□

> □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□DQ□□□□□□□□□□□□□□□□□ (Roebuck, 2012):

> □□

□□

DQ□□□2□□□DQ□□□□□□□□
(Weiskopf and Weng, 2013): □□□□DQ□□□□□□□□□□□□□□□□□□□□□□DQ□□□□□□□□□

□□□□□□□□□DQ□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DQ□□□□□□□□□□

15.1

Chapter 14 □□
and Johnson (2003) □□DQ□□□□□□

1. □□: □□□□□□□□□□□□□□□□□□□□
2. □□
3. □□
4. □□□

□□□1□□□□□□□□□□□□□□□□□

□□□2□□□□□□□□□□□□□□□□□□
15.1.2 □□□□□□□□□□□DQ□□□
(Defalco et al., 2013; Makadia and Ryan, 2014; Matcho et al., 2014; Voss et al., 2015a,b; Hripcsak et al., 2018) □□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□3□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□DQ□□□□16□Chapter 17□Chapter 18 □□□□□□□□□□□□□ ## □□□□□□□□□□□

□□□Kahn et al. (2016) □□□□□□□□□□□□□□□□□□□□□□□DQ□□□□3□□□□□□□□□□□□□□□□□□□□□□□□□□

1. □□□: □□□□□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□
    - □: □□□
    - □□: □□DRUG_EXPOSURE □□□□□PROVIDER_ID □ PROVIDER □□□□□□□□□□□□□□□□□□□□□□□□□□
    - □□: □□BMI□□□□
2. □□□: □□
3. □□□: □□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□□□□□
    - □□□: □□□□□PERSON □□□□□□□□ PERSON_ID □□□□□□□□□□□□□□□□□
    - □□□□: □□
    - □□□: □□□

□□□□□□□□□□2□□□□□□□□□□□□□□□□□

- □□ □□
- □□□□□□□□□□□□□ □□

15.1.1

□□□ □□□□□□□□□□□
□□ACHILLES □□ (Huser et al., 2018) ACHILLES □□CDM □□ (Huser et al., 2016) ACHILLES □□□ ATLAS □□□□□□□□□□□□□

ACHILLES □ 170 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ ID □□□□□□□□□□□□□□□□□□□□□□□□□□□□715: DRUG_CONCEPT_ID □□ DAYS_SUPPLY □□□□□□□506: □□□ ATLAS □□□□□□□□□□□□□□

□□ □□□□□□□□□□□□□□□□□□□ (DQD) □□□ACHILLES □ CDM □□CDM □□□□□□□□□□□□□□□□□□□□□□□□□ 1,500 □□□ FAIL □□□□□□□□□□□ 15.1.1 □□□□□□□□□□□□□□□□□□□□□□□□□□□□

□: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□	□□□□□□□□□□	□□	□□
0.34	VISIT_OCCURRENCE □ provider_id □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ yes □□□ no □□□	0.05	FAIL

□□□□□□□	□□□□□□□□	□□	□□
0.99	MEASUREMENT □□□□□ measurement_source_value □□□□□□□□□□□□□□□□□□□□□□□□□□ 0 □□□□□□□□□□□□	0.30	FAIL
0.09	DRUG_ERA □□□□□ drug_concept_id □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□	0.10	PASS
0.02	DRUG_EXPOSURE □□□□□ DRUG_EXPOSURE_END_DATE □□□□□□□□ DRUG_EXPOSURE_START_DATE □□□□□□□□□□□□□□□□□□□□□□□□□□□□□	0.05	PASS
0.00	PROCEDURE_OCCURRENCE □□□□□ procedure_occurrence_id □□□□□□□□□□□□□□□□□□□□□□□□□□□	0.00	PASS

□□
CDM □□□
CDM □□□
ID □□□
ID □□□

ACHILLES □ DQD □ CDM □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CDM □□

15.1.2 ETL

□□
CDM □□□□□□ ETL (Extract-Transform-Load) □□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□ETL □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ET

ETL □□
6 □□□□□□ Rabbit-in-a-Hat □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□ CDM □□□□□□□□□ R □□□□□□□□□□□□□□□□□□□□
CDM □□□

```
source("Framework.R")
declareTest(101, "Person gender mappings")
add_enrollment(member_id = "M000000102", gender_of_member = "male")
add_enrollment(member_id = "M000000103", gender_of_member = "female")
expect_person(PERSON_ID = 102, GENDER_CONCEPT_ID = 8507
expect_person(PERSON_ID = 103, GENDER_CONCEPT_ID = 8532)
```

□□□□□□Rabbit-in-a-Hat □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ gender mappings □□□□□□□□□□□□□□□□□□□□□□□□□□□□ ENROLLMENT □□□□□□□□Rabbit-in-a-Hat □□□□□□□□□ add_enrollment □□□□□□□□MEMBER_ID □□□ GENDER_OF_MEMBER □□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□ETL □□□□□□□□□□□□□□□□□□ PERSON □□□□□2□□□□□□□□□□□□□□□□□□□□□

ENROLLMENT □□White Rabbit □□□□□□□□□□□□□□□□□□□□

ETL □□ SQL □□□□□□□□□□□□□ETL □□□□□□□□□□□□□□□□□□□□□□□ SQL □□□□□□□□□□□□□□□□□□□□□□

```
insertSql <- generateInsertSql(databaseSchema = "source_schema")
testSql <- generateTestSql(databaseSchema = "cdm_test_schema")
```

□□□□□□□□□ 15.1 □□□□□□□□□

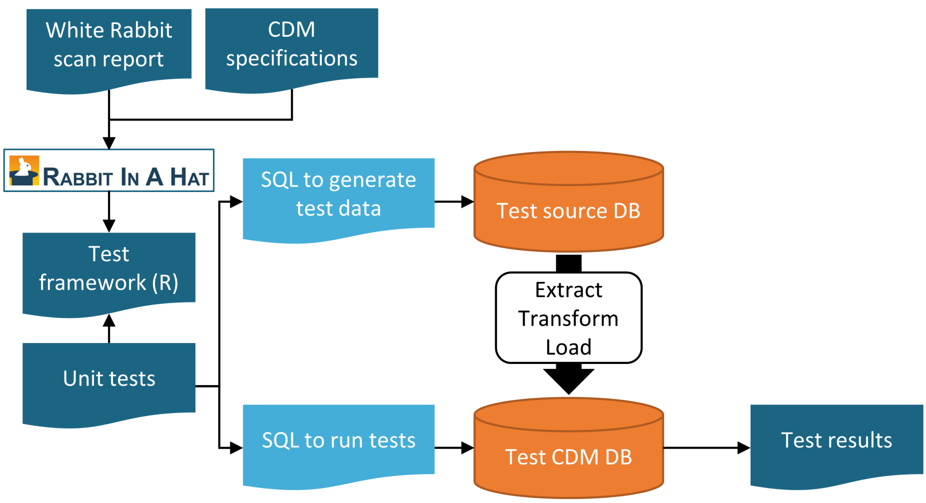


Figure 15.1: Rabbit-in-a-Hat □□□□□□□□□□□□□□□ ETL (Extract-Transform-Load) □□□□□□□□□□□

□□□ SQL □□ 15.1.2 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□

□: ETL □□□□□□□□□□

ID	□□	□□
101	Person gender mappings	PASS
101	Person gender mappings	PASS

□□□□□□□□□□□□□ETL □□□□□□□□□□□□□□□□□□□□□□□□□□□

15.2

□□□□□□□□□□□□□□□□□□□□DQ□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□DQ□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□ACHILLES□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□ 16 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□DQ□□□□□□□□□

15.2.1

□□Vocabulary □□ Vocabulary □□□□□□□□□[1] □□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□MethodEvaluation R □□□□□□ checkCohortSourceCodes □□□□□□□□□□□□□□□□□□□□□□□□ATLAS □□□□□□□□□□□□□□□□□□□□□□ 15.2 □□ 440383 (“□□□□□□”) □□□□□□□□□□□□□□□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 9□□□3.11□□□ICD-10□□□F32.8□F32.89□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 9□□□ICD-10□□□□□□□□□□□□□□□□□□□□□□□□ 10□□□□□□□□□□□ICD-9□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ICD-10□□□F32.9□□

‰ per month	Max monthly ‰	Person count	Description
	26.81	92,019,885	**Depressive Disorder**
	6.64	15,969,198	**Depressive disorder** 440383
	6.64	15,686,275	311 (ICD9CM) Depressive disorder, not elsewhere classified
	0.46	188,230	F328 (ICD10CM) Other depressive episodes
	0.38	94,693	F3289 (ICD10CM) Other specified depressive episodes
	3.10	12,010,783	**Adjustment disorder with mixed emotional features** 433454
	3.07	9,839,712	30928 (ICD9CM) Adjustment disorder with mixed anxiety and depressed mood
	3.03	2,049,618	F4323 (ICD10CM) Adjustment disorder with mixed anxiety and depressed mood
	0.04	121,453	3091 (ICD9CM) Prolonged depressive reaction
	3.17	9,237,192	**Dysthymia** 433440

Figure 15.2: checkCohortSourceCodes□□□□□□□□□□□

□□□ R □□□□□□ `findOrphanSourceCodes` □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 15.3 □□□□□□□□ICD-10□□□J85.0□□□□□□□□□□□□□□□□□□4324261□□□□□□□□□□

[1] https://github.com/OHDSI/Vocabulary-v5.0/issues

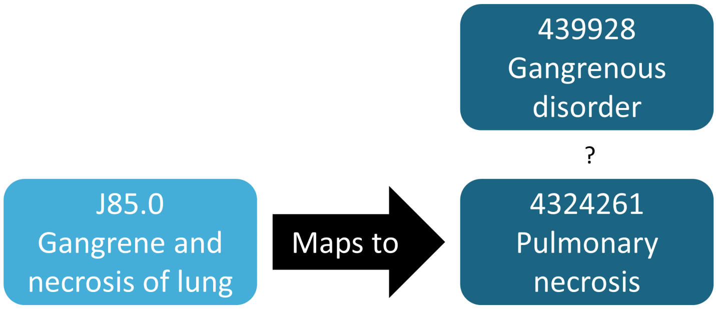


Figure 15.3: □□□□□□□□□□□□□□□□

15.3 ACHILLES

□□□□□CDM□□□□□□□□□□□□□ACHILLES□□□□□□□□□□□□□

□□□R□□□□□□□□□□□□□□□□□□□□□□□ACHILLES□ DatabaseConnector □□□□□□□□□□□□□□□□□□ `createConnectionDetails` □□□□□□□□□□□□□□□□□□□□

```
library(Achilles)
connDetails <- createConnectionDetails(dbms = "postgresql",
                                       server = "localhost/ohdsi",
                                       user = "joe",
                                       password = "supersecret")

cdmDbSchema <- "my_cdm_data"
cdmVersion <- "5.3.0"
```

□□□2□□ `cdmDbSchema` □□□CDM□□□□□□□□□□□□□□□□□□□□CDM□□□□□□□□□□□ SQL Server□□ `<- "my_cdm_data.dbo"` □□□

□□□ACHILLES□□□□□□□

```
result <- achilles(connectionDetails,
                   c
## Data Quality Dashboard  {#dqdInPractice}

  CDM       Data Quality Dashboard            \@ref(achillesInPractice)   CDM
```

```
cdmDbSchema <- "my_cdm_data.dbo"
```

□□□Dashboard□□□□□□…

```r
DataQualityDashboard::executeDqChecks(connectionDetails = connectionDetails,
                                      cdmDatabaseSchema = cdmDbSchema,
                                      resultsDatabaseSchema = cdmDbSchema,
                                      cdmSourceName = "My database",
                                      outputFolder = "My output")
```

□□□

```r
viewDqDashboard(jsonPath)
```

□□jsonPath□□Dashboard□□□□□□JSON□□□□□□□□□□□□□□□□□□□□□□□□□□□

Dashboard□□□□□□□□□15.4□□□□□□□□□□□□□□□□□□□□□□□□□□□□Kahn□□□□

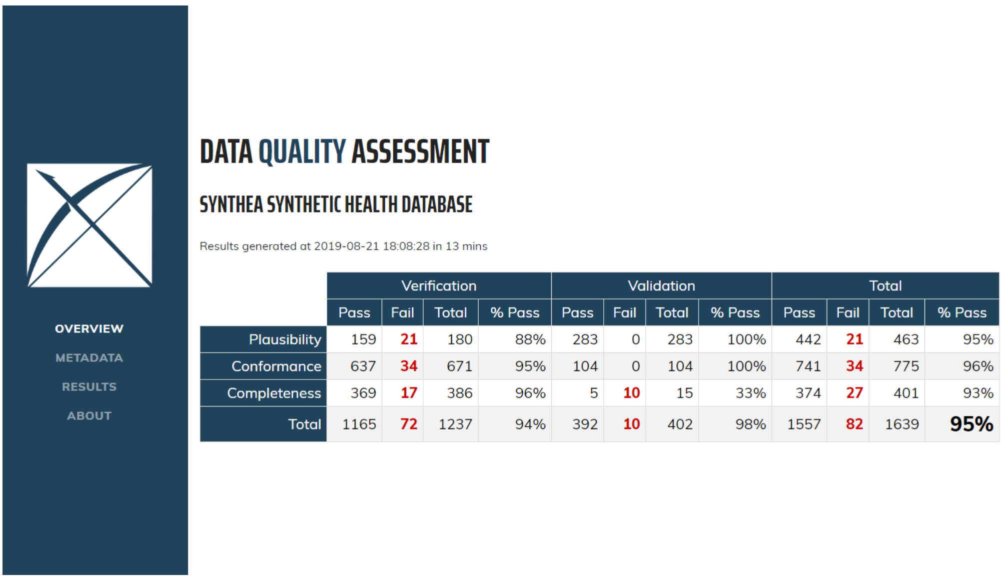


Figure 15.4: Data Quality Dashboard□□□□□□□□□□□□□□□□□

□□□□□□□□*Results*□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□15.5□

□□

15.4

□□□□□□B.4□□

```r
library(MethodEvaluation)
json <- readChar("cohort.json", file.info("cohort.json")$size)
sql <- readChar("cohort.sql", file.info("cohort.sql")$size)
checkCohortSourceCodes(connectionDetails,
```

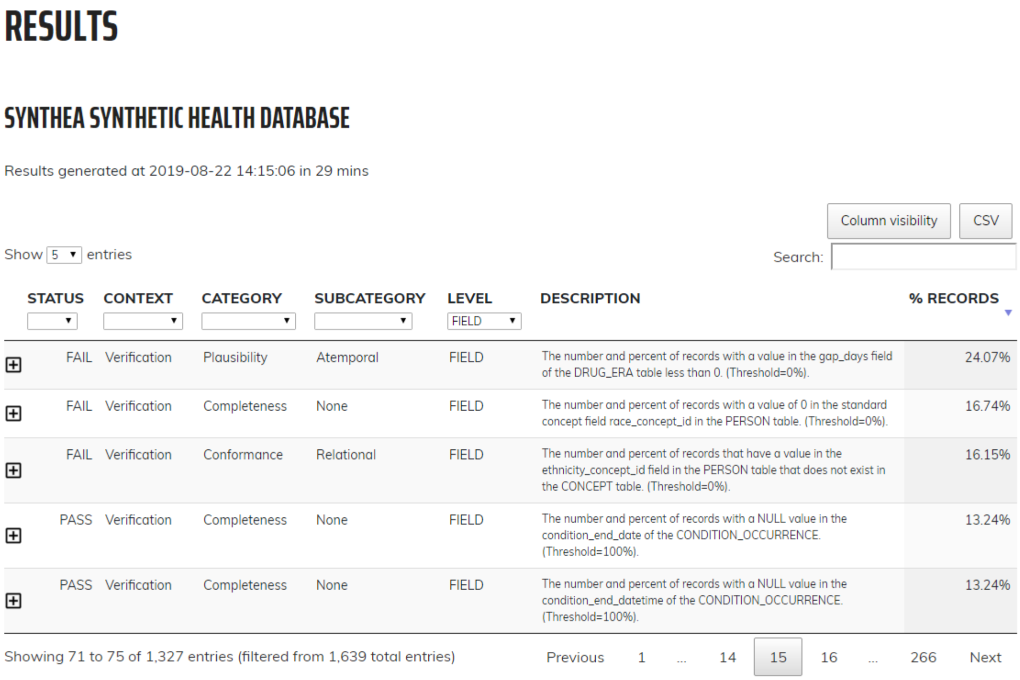


Figure 15.5: Data Quality Dashboard□□□□□□□□□□□□□□□□□

```
                  cdmDatabaseSchema = cdmDbSchema,
                  cohortJson = json,
                  cohortSql = sql,
                  outputFile = "output.html")
```

□□□□□□□□□□□□□□□□□□□□□□□□□□15.6□□□□□□□□□□□□□□□□□□□□□□□”□□
9□□□□□□□ICD-10□□

□□Angioe

```
orphans <- findOrphanSourceCodes(connectionDetails,
                                 cdmDatabaseSchema = cdmDbSchema,
                                 conceptName = "Angioedema",
                                 conceptSynonyms = c("Angioneurotic edema",
                                                     "Giant hives",
                                                     "Giant urticaria",
                                                     "Periodic edema"))
View(orphans)
```

□□□	□□	□□ID	□□□□□□□
T78.3XXS	Angioneurotic edema, sequela	ICD10CM	508
10002425	Angioedemas	MedDRA	0

□□□	□□	□□ID	□□□□□□□
148774	Angioneurotic Edema of Larynx	CIEL	0
402383003	Idiopathic urticaria and/or angioedema	SNOMED	0
232437009	Angioneurotic edema of larynx	SNOMED	0
10002472	Angioneurotic edema, not elsewhere classified	MedDRA	0

□□□□□□□□□□□□□□□□□□□□□□□□□□□”Angioneurotic edema, sequela”□□□

15.5

- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□
- □□□
- □□□□□□□□□□□□□Data Quality Dashboard□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

15.6

□□□□

□□□□□□□□□□□□□□8.4.5□□□□□□□□R□R-Studio□□□□Java□□□□□□□□□□□□□□

```
install.packages(c("SqlRender", "DatabaseConnector", "remotes"))
remotes::install_github("ohdsi/Achilles")
remotes::install_github("ohdsi/DataQualityDashboard")
remotes::install_github("ohdsi/Eunomia", ref = "v1.0.0")
```

Eunomia□□□□□□□□□□□□R□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□

```
connectionDetails <- Eunomia::getEunomiaConnectionDetails()
```

CDM□□□□□□□□□□□□main□□□□

□□ **15.1.** Eunomia□□□□□□□□□□ACHILLES□□□□□□□□□□

□□ **15.2.** Eunomia□□□□□□□□□□Data Quality Dashboard□□□□□□□□□□

□□ **15.3.** DQD□□□□□□□□□□□□□□□□□□□□□□□

‰ per month	Max monthly ‰	Person count	Description
	60.60	24,189,656	**Inpatient or ER visit**
	39.50	15,003,249	**Emergency Room Visit** 9203
	39.50	15,003,249	ER (None) No matching concept
	23.90	9,186,407	**Inpatient Visit** 9201
	23.90	9,186,407	IP (None) No matching concept
	0.27	76,711	**Angioedema**
	0.27	76,711	**Angioedema** 432791
	0.26	64,726	9951 (ICD9CM) Angioneurotic edema, not elsewhere classified
	0.20	8,822	T783XXA (ICD10CM) Angioneurotic edema, initial encounter
	0.09	3,163	T783XXD (ICD10CM) Angioneurotic edema, subsequent encounter

Figure 15.6: □□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□E.10□□□□□□

Chapter 16

— —

□…: *Joel Swerdel, Seng Chan You, Ray Chen & Patrick Ryan*

> □…
> □…*, 1935*

OHDSI□…

16.1

□…ACE□…ACE□…

□…
□… {#CohortValidation}

Hripcsak and Albers (2017) □…

□…1)
□…2)
□…3) □…

OHDSI□…

□…

□…Chapter 12□ "ACE□…

□…
□…2) □…Phe-
Valuator□

16.1.1

□…
16.1 □…

		Gold Standard True	Gold Standard False
Cohort Definition	True	True Positive	False Positive
	False	False Negative	True Negative

Figure 16.1: □□□□□□□□□□□□□□

□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□incorrect□ □□□□□□□□□□□□□□□

□□

1. □□□□□□□□□ – □□

   □□ = □□□ / (□□□ + □□□)

2. □□□□□□□□□□ – □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

   □□□ = □□□ / (□□□ + □□□)

3. □□□□□□□□□□□□□**PPV**□ – □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

   PPV = □□□ / (□□□ + □□□)

4. □□□□□□□□□□□□□**NPV**□ – □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

   NPV = □□□ / (□□□ + □□□)

□□□□□□□□□□□□□□100%□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Rubbo et al. (2015) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□33□□□□□□□1□□□□□□

□□
□□
□□□□□□□□□

□□

1. □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□IRB□□□□□□□□□□□□□□□□
2. □□
3. □□
4. □□□
5. □□
6. □□

□□□□□□□□□□□□□□□□□□□□PPV□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□
6□□

□□

16.1.2

□□

1. □□□□□□□□OHDSI□□□□□□□□□□□□□□□□□IRB□□□□□□□□
2. □□□□□□□□□□□□□□□□□□□□□ATHENA□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□ 4163261□□□ID 4314337□□□□□□□□□□□□□□□□□□□□□ID 4048666□□□□□□
3. ATLAS□□□□□□□□□□□□□□□□□□□□□□□□100□□□□□□□□□□□□□□□□PERS
4. □□□□□□□□□□□□□□□□EHR□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
5. □□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
6. □□□
7. □□□1□□□□□□□2□□□□□□□□□□□□3□□□□□□□□□□□□□□□□□□
8. □□□□□□□□PPV□□□□□□□□□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□□□□□□□
9. □□□CUIMC□□
10. □□□□□□PPV□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□

□□□□□□MI□□□□□□□□□□□□□□□□□□□□□□□□□□□□Rubbo et al. (2015) □□□□□□□□□□□□□□□□□□□□□□
PheValuator

OHDSI□□ et al., 2019) □□ PheValuator□□Patient Level Prediction□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□

1. □□□□□□□□**“xSpec”**□□□
2. □□□□□□□□□□□**“xSens”**□□
3. xSpec□□□xSens□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 13□□□□□□□□□□□□□□□□□□□
4. □□
5. □□

□□

□□□Patient-Level Prediction□□Chapter 13□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ et al., 2013) □□

16.1.3 PheValuator

□□

MI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

[1]https://github.com/OHDSI/PheValuator

□□□□ 1: xSpec □□□□□□□

□□□□MI□□
16.2 □ATLAS□□□□MI□xSpec□□□□□□□□□□□□□□

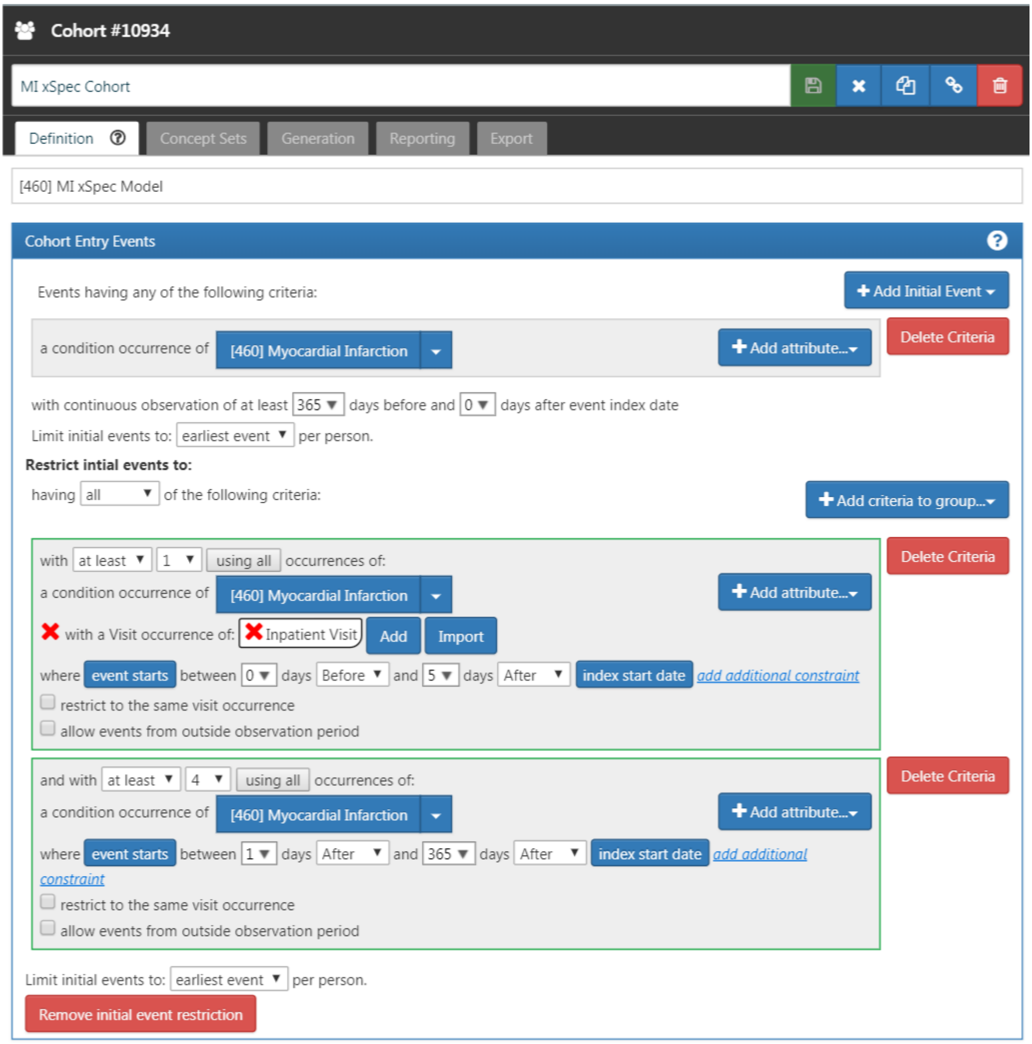


Figure 16.2: □□□□□□□□□□□□□□□□□□□□□xSpec□□

□□□□ 2: xSens □□□□□□□

□□□□□□□□□□□□□□□□□xSens□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
16.3 □ATLAS□□□□MI□xSens□□□□□□□□□□□□□□

□□□□ 3: □□□□□□□□

`createPhenoModel` □□

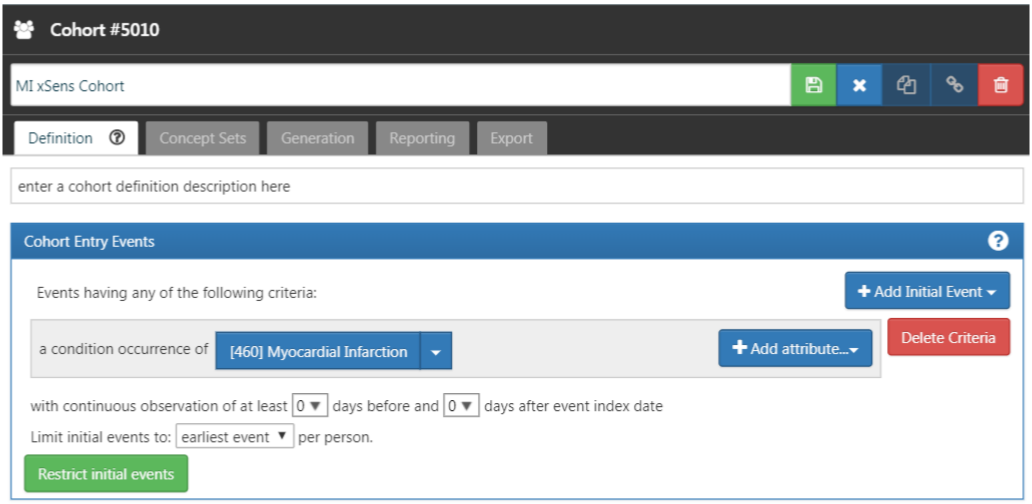


Figure 16.3: □□□□□□□□□□□□□□□□□□□□xSens□□

xSpec□□

□□□ □□ pcd-f□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□:

```
setwd("c:/temp")
library(PheValuator)
connectionDetails <- createConnectionDetails(
  dbms = "postgresql",
  server = "localhost/ohdsi",
  user = "joe",
  password = "supersecret")

phenoTest <- createPhenoModel(
  connectionDetails = connectionDetails,
  xSpecCohort = 10934,
  cdmDatabaseSchema = "my_cdm_data",
  cohortDatabaseSchema = "my_results",
  cohortDatabaseTable = "cohort",
  outDatabaseSchema = "scratch.dbo", #
  trainOutFile = "5XMI_train",
  exclCohort = 1770120, #xSens
  prevCohort = 1770119, #
  modelAnalysisId = "20181206V1",
  excludedConcepts = c(312327, 314666),
  addDescendantsToExclude = TRUE,
  cdmShortName = "myCDM",
  mainPopnCohort = 0, #
  lowerAgeLimit = 18,
```

```
  upperAgeLimit = 90,
  gender = c(8507, 8532),
  startDate = "20100101",
  endDate = "20171231")
```

□□□□□□“my_results” □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□cohortDatabase
- “my_results.cohort”□□□□□□□□□□□□□□□□□□□□□□□□□cdmDatabaseSchema
- “my_cdm_data”□□□□□□□□□□□□□□□□□□□□□□□□□□□□2010□1□1□□□2017□12□31□□□□CDM
□□□□□□□□□□□□□□□□ID312327□314666□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□ **4:** □□□□□□□□□

`createEvalCohort`□□□□PatientLevelPrediction□□□□□□`applyModel`□□□□□□□□□□□□□□□

□□□□

```
setwd("c:/temp")
connectionDetails <- createConnectionDetails(
  dbms = "postgresql",
  server = "localhost/ohdsi",
  user = "joe",
  password = "supersecret")

evalCohort <- createEvalCohort(
  connectionDetails = connectionDetails,
  xSpecCohort = 10934,
  cdmDatabaseSchema = "my_cdm_data",
  cohortDatabaseSchema = "my_results",
  cohortDatabaseTable = "cohort",
  outDatabaseSchema = "scratch.dbo",
  testOutFile = "5XMI_eval",
  trainOutFile = "5XMI_train",
  modelAnalysisId = "20181206V1",
  evalAnalysisId = "20181206V1",
  cdmShortName = "myCDM",
  mainPopnCohort = 0,
  lowerAgeLimit = 18,
  upperAgeLimit = 90,
  gender = c(8507, 8532),
  startDate = "20100101",
  endDate = "20171231")
```

□□□□□□□□□□□□□□□□□c:/temp/lr_results_5XMI_train_myCDM_ePPV0.75_20181206V1.rds□

□□□□ **5:** □□□□□□□□□□□□□

□□
□□□□□□□□□ {#GeneralizabilityOfEvidence}

□□□
□□□□□2□□□

OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□vs.□□□□□□□□□□□□
vs.□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□vs.□□□□□□□□□□□□
vs.□□/□□□

16.2

- □□□
- □□□
- □□□
- □□□□□□□□□□□□□□□□□□□PheValuator□□□□□□□□□□□□□□□□□□2□
- OHDSI□□□

Chapter 17

—　—

□□□□□□□□: *Martijn Schuemie*

□□□□□□□□□□□□□□□□□□□□□

　　□□□□□□□□□□□□□□□□□□□□

□□□ 17.1.1 □□CDM□□□ 8.1 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI Methods Library□□□□□□□□□□□□□□□□□□□□Methods Library□□□□□□□□□□□□

□□Methods Library□□

17.1

17.1.1

□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□R□□□□□□□□□DatabaseCo R□□□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□ OHDSI Methods Library □□□R□□□□□□□□□□□□□□□□

17.1.2

□□ 2008) □□□4□□□□□

- □□□: □□

- :
- : <- spl(y, 100) sampledPatients <- takeSample(patients, sampleSize = 100)
- : 1

17.1.3

- : 1 1
- : 2 2

CDM Library

17.1.4 Methods Library

OHDSI Methods Library Library Library Methods Library
##

OHDSI OHDSI Suchard Martijn Schuemie OHDSI Peter Rijnbeek Jenna Reps

GitHub drat

OHDSI License V2 R C++ SQL Java OHDSI

17.1.5

GitHub git

OHDSI

- : 4.3.2 4.3.3
- : 4.3.3 4.4.0
- : 4.4.0 5.0.0

[1]https://github.com/OHDSI/CohortMethod/issues
[2]http://forums.ohdsi.org/

17.1.6

[illegible]R[illegible]
[illegible]1[illegible]

[illegible]

17.1.7

[illegible]GitHub[illegible]
drat[illegible]

17.1.8

[illegible]OHDSI[illegible]

17.1.9

OHDSI[illegible]

OHDSI[illegible]OHDSI[illegible]

17.1.10

OHDSI[illegible]GitHub[4][illegible]GitHub[illegible]https://github.com/security[illegible]OHDSI[illegible]

17.1.11

OHDSI[illegible]GitHub[illegible]GitHub[illegible]https://github.com/security[illegible]

17.2

[illegible]

17.2.1

[illegible]O[illegible]

[illegible]

17.2.2

[illegible]
[illegible]

[3]https://ohdsi.github.io/MethodsLibrary/
[4]https://github.com/

17.3

- [illegible]CDM[illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]

Chapter 18

— —

: Martijn Schuemie

18.1

OHDSI
#ref(studyOutputs) CohortMethod

- Cox

18.1
18.1 2

18.2

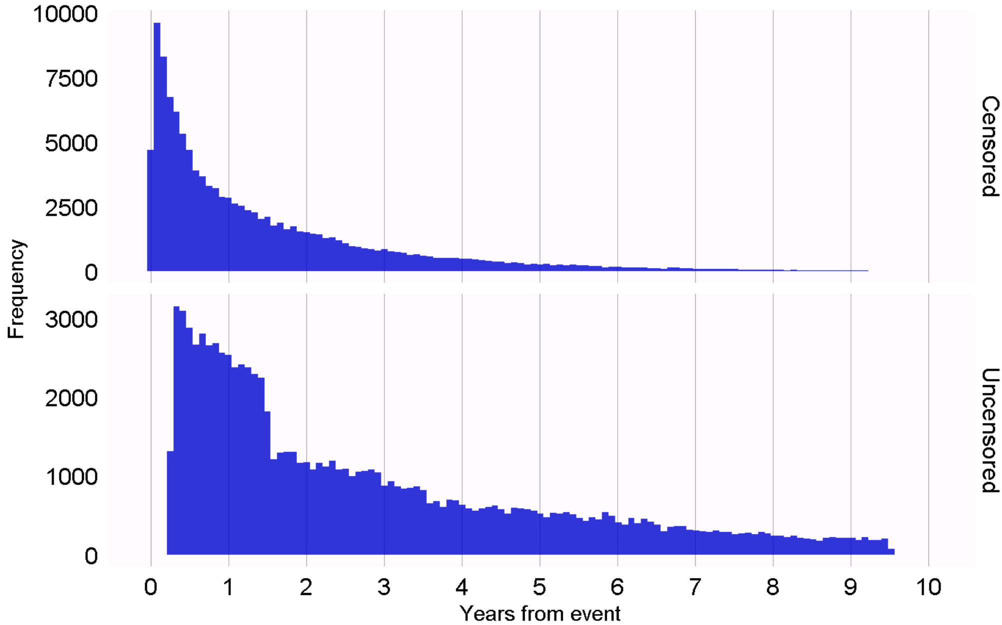


Figure 18.1: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□

18.2.1

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□-□□□□□□□□□□□□□□(Lipsitch et al., 2010) □□□ et al., 2016) □□□□(Zaadstra et al., 2008) □□□□□□□□□□□□□□□□□□□□□□□□MS□□□□□□□

□□

- □□
- □□□1□□□
- □□
- □□
- □□

□□□□□□□□□□□□□□□□□□□-□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ et al., 2016) □□□

18.2.2

□□□□□□□□1□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ et al., 2014)

□□□□□□OHDSI□□

□□□ et al., 2013)□□□□□□□□□□□□□□□□□□□□□□□□□□10□□□□□□□□□□□□□□□□□

□18.2□□

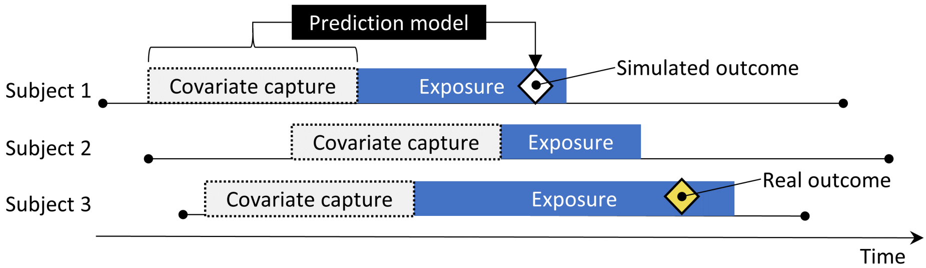


Figure 18.2: □□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□

18.2.3

□□□

- □□□□□□□□□□□□□**AUC**□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□95□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□$1/(\quad)^2$□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□**MSE**□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□**1**□□□□□□□□□□□□□□□□□□□□□□□□□□□□□$\alpha = 0.05$□□□□□□□□□□ □□□□□□□
- □□□**2**□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□$\alpha = 0.05$□□□□□□□□□□□□□□□$1 -$ □□□□□□□
- □□

□□□□□□□□□□□□□□

□□□

18.2.4 P

□□□□□□□□□1□□□□□$\alpha = 0.05$□□5□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□5 et al., 2014) □□□

□□□□□□□□□□□□□□□□□□□□□□□□□□-□□□□□□□□□□□□□□□□□□□□□□□□ $\hat{\theta}_i$ □□□□□□□□□□□□□□□□□ $\hat{\tau}_i$ □□□□□□$i = 1, ..., n$□□□□□□□□□□□□□□□□□□□□□□□ θ_i □0□□□□□□□□□ i □□□□□□□□□□□□□□□□□□□□□ β_i □□□□□□□□□□□θ_i □□□□□□□□□□□□□□□□ β_i □□□□□□□□□□p□□□□□□□□□$\hat{\theta}_i$ □□□□ $\theta_i + \beta_i$ □□□□□$\hat{\tau}_i^2$ □□□□□□□□□□□□□□□□□□□□□□□□□□p□□□□□□ β_i □□□□□□□□□□□□□□□□□□□□□ β_i □□□□□□ μ □□□□ σ^2 □□□μ □ σ^2 □□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□:

$$\beta_i \sim N(\mu, \sigma^2) \text{ □□□ } \hat{\theta}_i \sim N(\theta_i + \beta_i, \tau_i^2)$$

□□□ $N(a,b)$ □□□□ a □□□ b □□□□□□□□□□□□□□:

$$L(\mu, \sigma | \theta, \tau) \propto \prod_{i=1}^{n} \int p(\hat{\theta}_i | \beta_i, \theta_i, \hat{\tau}_i) p(\beta_i | \mu, \sigma) \mathrm{d}\beta_i$$

□□□□□□□□□□ $\hat{\mu}$ □ $\hat{\sigma}$ □□□□□□□□□□□□□□□p□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□ $\hat{\theta}_{n+1}$ □□□□□□□□□□□□□□ $\hat{\tau}_{n+1}$ □□□□□□□□□□□□□□
$□_{n ## □□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□ACEi□□□□□□□□□□□□□□□□□□□□□□THZ□□□□□□□□□□

18.2.5

□□□
100□□

□□□

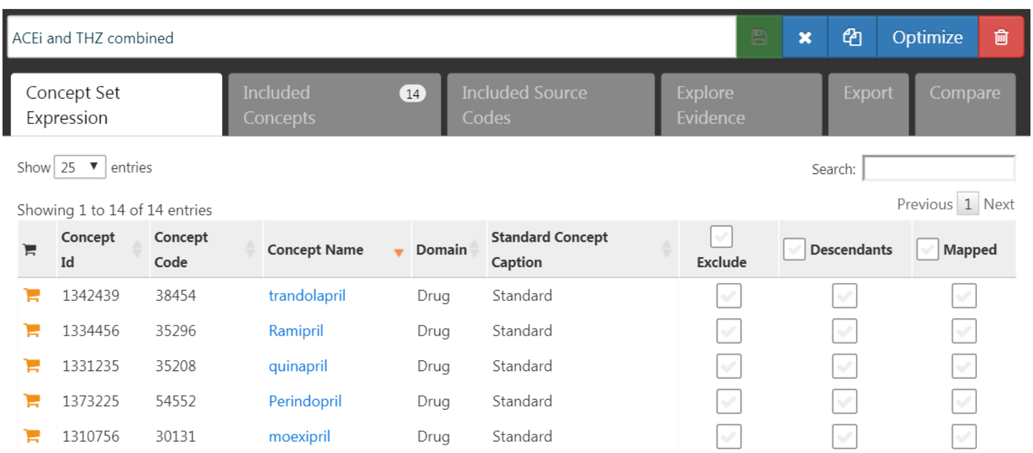


Figure 18.3: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□

□□□

□□□

□□□□□□□□□□□□□C.1□□□□□□□76□□□□□□□□□□□□□□□□□□□

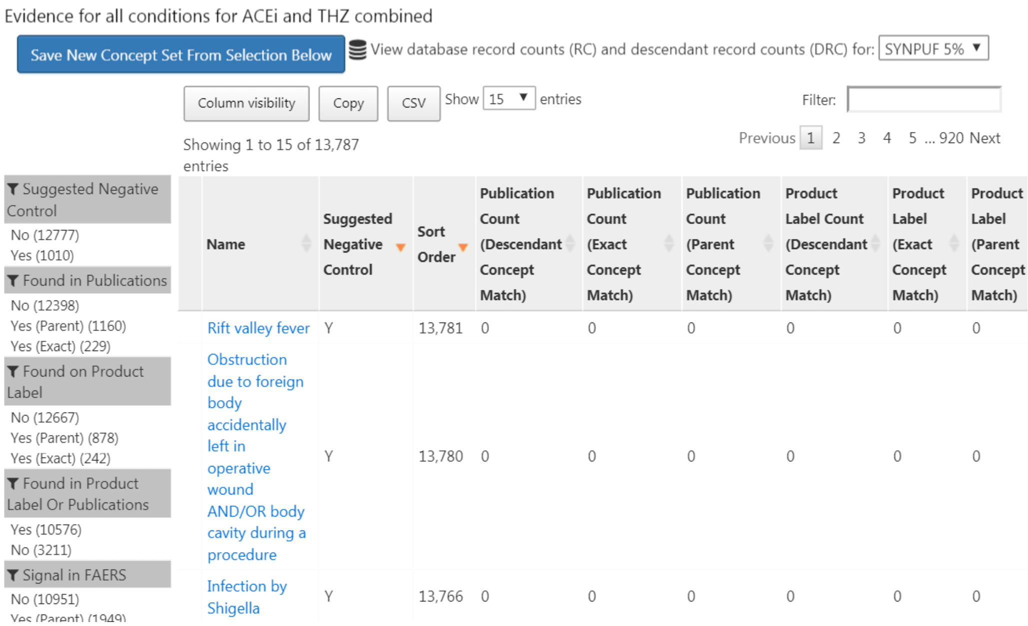


Figure 18.4: □□

18.2.6

□□ Query Language□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 9□□□SQL□□□R□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□SQL□□□R□□□□

OHDSI□□□

```
library(MethodEvaluation)
#     ACEi = 1          -
eoPairs <- data.frame(exposureId = 1,
                      outcomeId = ncs)

pcs <- synthesizePositiveControls(
  connectionDetails = connectionDetails,
  cdmDatabaseSchema = cdmDbSchema,
  exposureDatabaseSchema = cohortDbSchema,
  exposureTable = cohortTable,
  outcomeDatabaseSchema = cohortDbSchema,
  outcomeTable = cohortTable,
  outputDatabaseSchema = cohortDbSchema,
  outputTable = cohortTable,
  createOutputTable = FALSE,
  modelType = "survival",
  firstExposureOnly = TRUE,
  firstOutcomeOnly = TRUE,
```

```
  removePeopleWithPriorOutcomes = TRUE,
  washoutPeriod = 365,
  riskWindowStart = 1,
  riskWindowEnd = 0,
  endAnchor = "cohort end",
  exposureOutcomePairs = eoPairs,
  effectSizes = c(1.5, 2, 4),
  cdmVersion = cdmVersion,
  workFolder = file.path(outputFolder, "pcSynthesis"))
```

□□□synthe
□□□cohortDbSc

□□□
Methods Library□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

18.2.7

□18.5□□

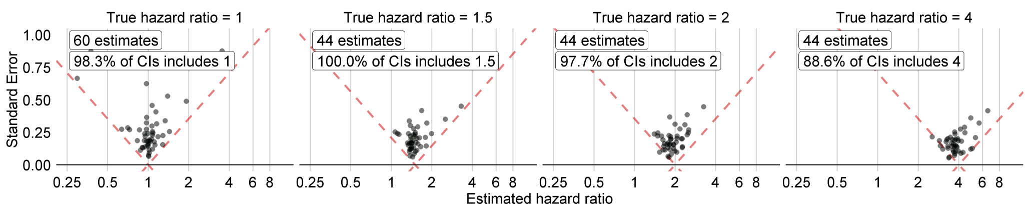


Figure 18.5: □□□□□□□□□□□□□□=1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□>
1□□

□□□□□□□□□□□□□[MethodEvaluation](https://ohd ## OHDSI□□□□□□□□□□□□

□□□

□□□□□OHDSI□□□

□□□

□□□1□□□□□□□□□□□□□□□□

18.3

- □□□
- □□□

[1] http://data.ohdsi.org/MethodEvalViewer/

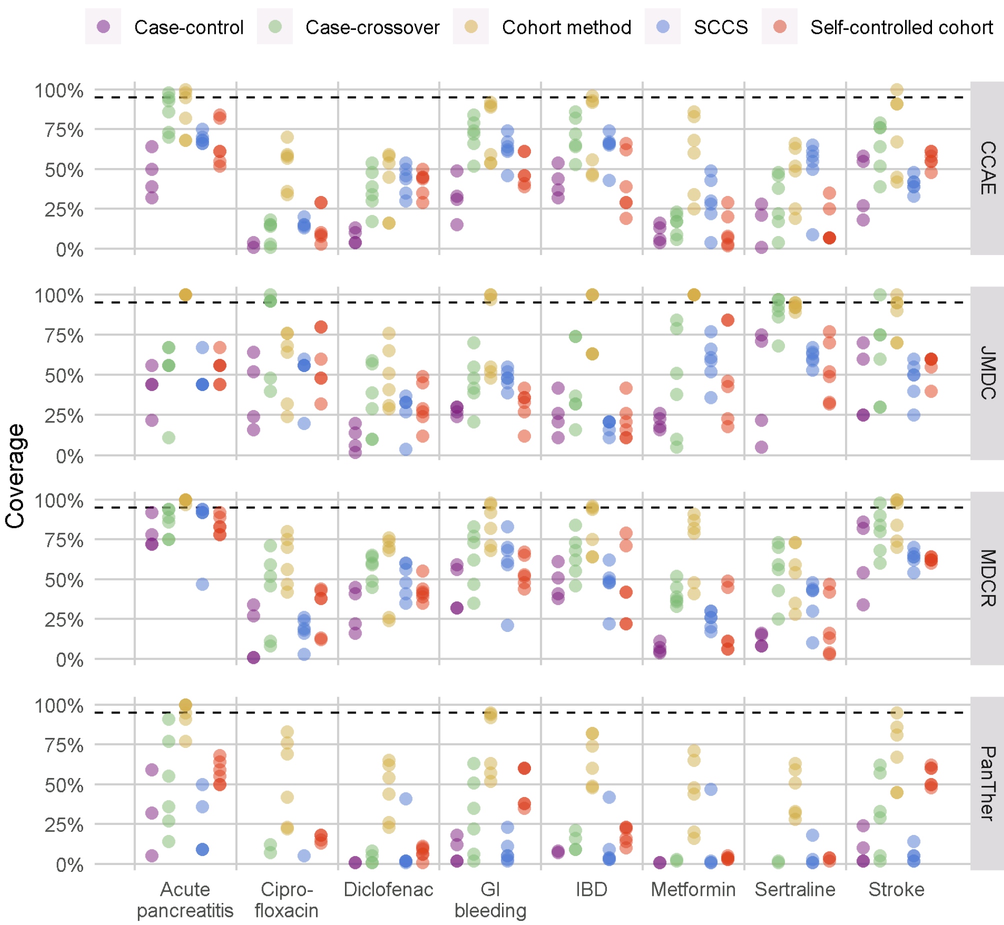


Figure 18.6: □□□□□□□□□□□□□□□□□95%□□□□□□□□□□□□□□□□□□□□□□□□ = □□□□□□□□□□□GI = □□□IBD = □□□□□□□

- □□□
- □□□□□□p□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□
- □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□p□□□□□

Part III

OHDSI

Chapter 19

[illegible]

[illegible]: *Sara Dempster & Martijn Schuemie*

[illegible]OHDSI[illegible]

[illegible]OHDSI[illegible]

[illegible]OHDSI[illegible]R[illegible]SQL[illegible]
CDM[illegible]OMOP
ETL[illegible]6[illegible]
[illegible]

19.0.1

[illegible]
Level Estimation[illegible]
12 [illegible] 18 [illegible]

19.0.2

[illegible]
[illegible]EHR[illegible]
Level Prediction[illegible]PLE[illegible]

19.0.3

[illegible]
[illegible]

19.0.4

OHDSI[illegible]
2, 7, 11, 12, 13 [illegible]

□□

OHDSI□□□
?? □□□□□□□□□□□

19.0.5

□□

19.0.6 CDM

OHDSI□□□□□□□□□□□□□□OMOP□□□□□□□□□CDM□□□□□□□□□□□□□□□□□□□□
4 □□□□□□□□□□□□□□□□□□□□□□□ETL□□□□□□□□□□□□□□□□□□□□□□□□□□
CDM□□

CDM□□□□□OMOP□□□□□□□□□□5□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□
CDM□□□□□□□□□□□□OMOP□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□

19.0.7

□□
(T2DM) □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 7 □□□□□□□□3□□□□□

□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□

□□□T2DM□□□□□□□□□□□□T2DM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□T2DM□□□□T2DM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□OHDSI□□□□□□□□□□
7 □□□

19.0.8

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□Chapter 15 □□□□□□□□□□□□□□□□□□

□□□□□□T2DM□□□□T2DM□□□□□□□□□□□□□□□□□□□□□□□□□□□T2DM□□□□□
7 □□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□PLE□□□□□□□□□□□□□□□□

19.0.9

□□

□□

□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□ATLAS□□□□□□□□□□□□□□□□□□□□□
@ref(Ohd ## □□□

- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]
- [illegible]

Chapter 20

– – OHDSI

□□□□: *Kristin Kostka, Greg Klebanov & Sara Dempster*

OHDSI□□□
19□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
Research Network□□□□□□□□□□□□□□□□□□

20.1 OHDSI

OHDSI□□□
CDM□□

□□□□□□□**OHDSI**□□□□□□□□□□□□□

- □□□□□□□□□□□**: OHDSI**□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□□□□**:** □□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□□□**:** □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

20.2 OHDSI

□□□Chapter **?**(StudySteps)□□□□CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

20.2.1 OHDSI

□□□

□□□

□□□

20.2.2 OHDSI

> OHDSI CDM

OHDSI OMOP CDM

OHDSI OHDSI OHDSI
CDM OHDSI

OHDSI OHDSI

- OHDSI GitHub
- CDM R SQL
- OHDSI OHDSI
- OHDSI GitHub
- R Shiny data.ohdsi.org

20.2.3 OHDSI

OHDSI

OHDSI

20.2.4 OHDSI

OHDSI OHDSI
CDM OHDSI OHDSI

- CDM
- RStudio
- R

> Chapter **?**(ExtractTransformLoad) ETL
> **?**(DataQuality)

Level Effect Estimation PLE Patient Level Prediction PLP OHDSI

□□□
□□□
OHDSI□□□□□□□□□□□

OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□
- □□□□
- □□□□□□□□

20.2.5

□□□

□□□

□□□
CDM□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□CMS Synthetic Public Use Files□Mitre□SyntheticMass□□□□Synthea□□□□□□□□□□□□□□□□□□□OHDSI□□□
19□□□□□□□□□□□□□□□□□□IRB□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□

20.2.6

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□

□□□□□□□□□□OHDSI□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□
GitHub□□□□□□□□OHDSI□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□

□□□

□□□□□□□□□□□□□□□□□□□SFTP□□□□□□Amazon S3□□□□□□□□□□□□□□□□□

□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□/□□□□□□□□□□□□□□□□□□□□□

□□□

OHDSI□□□

□□□

20.2.7

□□□
OHDSI□□□□□□□□□□□□□□□□□□□□GitHub□□□□□□□□□□□□□□□Shiny□□□□□

> OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□JANE□Journal/Author Name Estimator□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□[1]

□□
□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□Wiki□GitHub□□□□□□□□□□

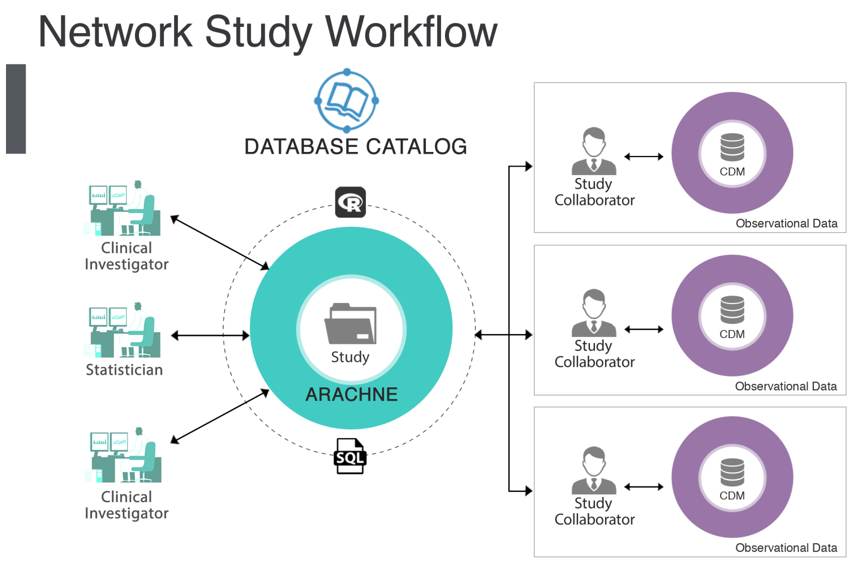


Figure 20.1: ARACHNE□□□□□□□□□□□□

ARACHNE□□□ARACHNE□OH

ARACHNE□□ACHILLES□□□□□□□□□□□□ATLAS□□□□□□□□□□□□□□□□□□□□□□□□□□□□

20.3 OHDSI

□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□: □□
CDM□□□

□□□□: □□□□□□□□□□□□□ATLAS□OHDSI Methods Library□□□OHDSI Study Skeletons□□

□□□□□: □□
GitHub□□□□□□□□data.ohdsi.org□R Shiny□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
CDM□□□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□□□□□□□□□□□□□
4□□
CDM□□

[2]http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html ## □□: □□□□□□□□□□□□□□□□□□□

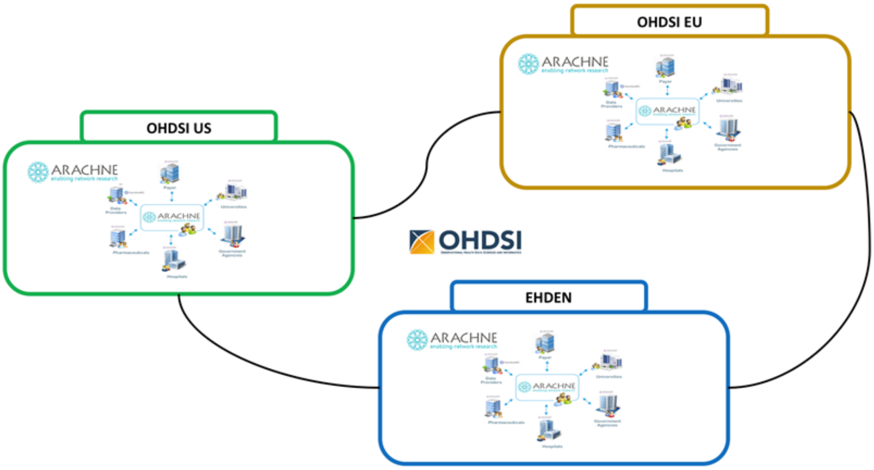


Figure 20.2: ARACHNE□□□□□□□□□□□□□□□□

20.4

- OHDSI□□□□□□□□□□□□□CDM□□□□□□□OHDSI□□□□□□□□□□□□□□
- OHDSI□□O
- □□□□□□□□□□□□□□□□□□□□□OHDSI Study Nurture Committee□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
- □□□□□□□□□□ □□□□□□□□□□□□□□□□□□□OHDSI GitHub□□□R Shiny□□□□□□□□□□□□□□□□□□□□□□OHDSI□□□□□□□□□□□□□□□

Chapter A

ACHILLES

ARACHNE OH

ATLAS

Bias

Boolean 2

Care site

Case control

Causal effect

Characterization 11

Claims data

Clinical trial

Cohort 1 10

Concept SNOMED CT 5

Concept set

Common Data Model, CDM

Comparative Effectiveness 2

Condition

Confounding

Covariate

Data quality

Device

Drug

Domain

Electronic Health Record, EHR

Epidemiology

Evidence-based medicine

ETL **Extract-Transform-Load**

Matching

Measurement

Measurement error

Metadata

Methods Library OHDSI R

Model misspecification OHDSI

Negative control -

Observation

Observation period «««< HEAD :

======= :

»»»> d84f15c84472a96d3c9d50b58c4246cdb83ccaa1

Observational study

OHDSI SQL R SqlRender SQL SQL SQL SQL Server SQL 9

Open science

Outcome

Patient-level prediction

Phenotype

Population-level estimation

Positive control -

Procedure

Propensity score, PS 2

Protocol

Rabbit-in-a-Hat ETL Rabbit 7

Selection bias

Self-controlled designs

Sensitivity analysis

SNOMED

Study diagnostics

Study package 17

Source code ICD-10

Standard Concept

THEMIS OHDSI

Visit 1

Vocabulary

White Rabbit ETL

Chapter B

B.1 ACE

- *ACE* B.1

365 0

ACE B.1

- 30
- 0

30

Table B.1: ACE□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1308216	□□□□□□	□□□	□□	□□□
1310756	□□□□□□□	□□□	□□	□□□
1331235	□□□□□	□□□	□□	□□□
1334456	□□□□□	□□□	□□	□□□
1335471	□□□□□□	□□□	□□	□□□
1340128	□□□□□□	□□□	□□	□□□
1341927	□□□□□□	□□□	□□	□□□
1342439	□□□□□□□□	□□□	□□	□□□
1363749	□□□□□□□	□□□	□□	□□□
1373225	□□□□□□□	□□□	□□	□□□

B.2 ACE

□□□□□□□□□□

□□□□□□□□□□□□

- □□□□□□□□□□□□□*ACE*□□□□□B.2□□□□□

□□□□□□□□□□□□□□□□□□□□365□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□

□□□□□

□□□□#1□□□□□□1□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□365□□□□□□□□□□□□□0□□□□□□□□□□1□□□□□□□□□□ B.3□□□□□□□□□□□□□□□□

□□□□#2□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□

- □□□□□□□□□1□□□□□□□□□□□□□□□□□□□□B.4□□□□□□□□□□0□□□□

□□□□#3□ACE□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□

- □□□□□□□□□0□□□□7□□□□□□□□□□□□□□B.4□□□□□□□□□□□□□□□□□□□1□

□□

□□□□□□□

□□□

ACE□□□□□□□□□□□□ (□ B.2)

- □□□□□□□30□□□□□□□□
- □□□□□□□0□□□□□□□

□□□□□□□□□□

0□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.2: ACE□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1308216	□□□□□□	□□□	□□	□□□
1310756	□□□□□□□	□□□	□□	□□□
1331235	□□□□□	□□□	□□	□□□
1334456	□□□□□	□□□	□□	□□□
1335471	□□□□□□	□□□	□□	□□□
1340128	□□□□□□	□□□	□□	□□□
1341927	□□□□□□	□□□	□□	□□□
1342439	□□□□□□□□	□□□	□□	□□□
1363749	□□□□□□□	□□□	□□	□□□
1373225	□□□□□□□	□□□	□□	□□□

Table B.3: □□□□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
316866	□□□□□□	□□□	□□	□□□

Table B.4: □□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
904542	□□□□□□□	□□□	□□	□□□
907013	□□□□□	□□□	□□	□□□
932745	□□□□□	□□□	□□	□□□
942350	□□□□□	□□□	□□	□□□
956874	□□□□□	□□□	□□	□□□
970250	□□□□□□□□	□□□	□□	□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
974166	□□□□□□□□□□	□□□	□□	□□□
978555	□□□□□□	□□□	□□	□□□
991382	□□□□□	□□□	□□	□□□
1305447	□□□□□	□□□	□□	□□□
1307046	□□□□□□□	□□□	□□	□□□
1307863	□□□□□	□□□	□□	□□□
1308216	□□□□□□	□□□	□□	□□□
1308842	□□□□□□	□□□	□□	□□□
1309068	□□□□□□	□□□	□□	□□□
1309799	□□□□□□	□□□	□□	□□□
1310756	□□□□□□□	□□□	□□	□□□
1313200	□□□□□	□□□	□□	□□□
1314002	□□□□□□	□□□	□□	□□□
1314577	□□□□□□	□□□	□□	□□□
1317640	□□□□□□□	□□□	□□	□□□
1317967	□□□□□□	□□□	□□	□□□
1318137	□□□□□□	□□□	□□	□□□
1318853	□□□□□□	□□□	□□	□□□
1319880	□□□□□□	□□□	□□	□□□
1319998	□□□□□□□	□□□	□□	□□□
1322081	□□□□□□□	□□□	□□	□□□
1326012	□□□□□□	□□□	□□	□□□
1327978	□□□□□□□	□□□	□□	□□□
1328165	□□□□□□	□□□	□□	□□□
1331235	□□□□□	□□□	□□	□□□
1332418	□□□□□□	□□□	□□	□□□
1334456	□□□□□	□□□	□□	□□□
1335471	□□□□□□	□□□	□□	□□□
1338005	□□□□□□□	□□□	□□	□□□
1340128	□□□□□□	□□□	□□	□□□
1341238	□□□□□	□□□	□□	□□□
1341927	□□□□□□	□□□	□□	□□□
1342439	□□□□□□□□	□□□	□□	□□□
1344965	□□□□□□□	□□□	□□	□□□
1345858	□□□□□□	□□□	□□	□□□
1346686	□□□□□□□	□□□	□□	□□□
1346823	□□□□□□□	□□□	□□	□□□
1347384	□□□□□□□	□□□	□□	□□□
1350489	□□□□□	□□□	□□	□□□
1351557	□□□□□□□	□□□	□□	□□□
1353766	□□□□□□□□	□□□	□□	□□□
1353776	□□□□□□	□□□	□□	□□□
1363053	□□□□□□	□□□	□□	□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1363749	□□□□□□□	□□□	□□	□□□
1367500	□□□□□	□□□	□□	□□□
1373225	□□□□□□□	□□□	□□	□□□
1373928	□□□□□□	□□□	□□	□□□
1386957	□□□□□□	□□□	□□	□□□
1395058	□□□□□□□	□□□	□□	□□□
1398937	□□□□□	□□□	□□	□□□
40226742	□□□□□□□	□□□	□□	□□□
40235485	□□□□□□□	□□□	□□	□□□

B.3 AMI

□□□□□□□□□□

□□□□□□□□□□□□

- □□□□□□□□ B.5□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□0□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□0□□□ (□ B.6)□□□□□□□□□□□□□1□□

□□

□□□□□□□

□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□7□□□□□□□

□□□□□□□□□□

180□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.5: □□□□□□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
314666	□□□□□□□	□□	□□	□□□
4329847	□□□□	□□□	□□	□□□

Table B.6: □□□□□□□□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
262	□□□□□□□□□□□□	□□□	□□	□□□
9201	□□□□□□	□□□	□□	□□□
9203	□□□□□□□	□□□	□□	□□□

B.4

□□□□□□□□□□

□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□ B.7□

□□□□□□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

- □□□□□□□□□□□□ □□□□□□□□□□□□□□0□□□□□□□□□□□□□□□□0□□□□□□□ B.8)□□□□□□□□□□□□□□□□1□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□7□□□□□□

□□□□□□□□□□

30□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.7: □□□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
432791	□□□□□	□□□	□□	□□□

Table B.8: □□□□□□□□□□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
262	□□□□□□□□□□□□	□□□	□□	□□□

[illegible]ID	[illegible]	[illegible]	[illegible]	[illegible]
9201	[illegible]	[illegible]	[illegible]	[illegible]
9203	[illegible]	[illegible]	[illegible]	[illegible]

B.5

[illegible]

[illegible]

- [illegible] B.9[illegible]

[illegible]365[illegible]0[illegible]

[illegible]

[illegible]1[illegible]1[illegible]

[illegible]

- [illegible]365[illegible]0[illegible]1[illegible] ([illegible] B.10)[illegible]

[illegible]#2[illegible]

[illegible]

- [illegible]1[illegible]([illegible] B.11)[illegible]0[illegible]

[illegible]#3[illegible]ACE[illegible]

[illegible]

- [illegible]0[illegible]7[illegible]([illegible] B.11)[illegible]

[illegible]

[illegible]

[illegible]

[illegible] B.9[illegible]

- [illegible]30[illegible]
- [illegible]0[illegible]

[illegible]

0[illegible]

□□□□□□□□□□

Table B.9: □□□□□□□□□□□□□□□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
907013	□□□□□	□□□	□□	□□□
974166	□□□□□□□□□□	□□□	□□	□□□
978555	□□□□□□	□□□	□□	□□□
1395058	□□□□□□□	□□□	□□	□□□

Table B.10: □□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
316866	□□□□□□	□□□	□□	□□□

Table B.11: □□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
904542	□□□□□□□	□□□	□□	□□□
907013	□□□□□	□□□	□□	□□□
932745	□□□□□	□□□	□□	□□□
942350	□□□□□	□□□	□□	□□□
956874	□□□□□	□□□	□□	□□□
970250	□□□□□□□□	□□□	□□	□□□
974166	□□□□□□□□□□	□□□	□□	□□□
978555	□□□□□□	□□□	□□	□□□
991382	□□□□□□	□□□	□□	□□□
1305447	□□□□□	□□□	□□	□□□
1307046	□□□□□□□	□□□	□□	□□□
1307863	□□□□□	□□□	□□	□□□
1308216	□□□□□□	□□□	□□	□□□
1308842	□□□□□□	□□□	□□	□□□
1309068	□□□□□□	□□□	□□	□□□
1309799	□□□□□□	□□□	□□	□□□
1310756	□□□□□□□	□□□	□□	□□□
1313200	□□□□□	□□□	□□	□□□
1314002	□□□□□□	□□□	□□	□□□
1314577	□□□□□□	□□□	□□	□□□
1317640	□□□□□□□	□□□	□□	□□□
1317967	□□□□□□	□□□	□□	□□□
1318137	□□□□□□□	□□□	□□	□□□
1318853	□□□□□□	□□□	□□	□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1319880	□□□□□□	□□□	□□	□□□
1319998	□□□□□□□	□□□	□□	□□□
1322081	□□□□□□	□□□	□□	□□□
1326012	□□□□□□	□□□	□□	□□□
1327978	□□□□□□□	□□□	□□	□□□
1328165	□□□□□□	□□□	□□	□□□
1331235	□□□□□	□□□	□□	□□□
1332418	□□□□□□	□□□	□□	□□□
1334456	□□□□□	□□□	□□	□□□
1335471	□□□□□□	□□□	□□	□□□
1338005	□□□□□□□	□□□	□□	□□□
1340128	□□□□□□	□□□	□□	□□□
1341238	□□□□□	□□□	□□	□□□
1341927	□□□□□□	□□□	□□	□□□
1342439	□□□□□□□□	□□□	□□	□□□
1344965	□□□□□□□	□□□	□□	□□□
1345858	□□□□□□	□□□	□□	□□□
1346686	□□□□□□□	□□□	□□	□□□
1346823	□□□□□□□	□□□	□□	□□□
1347384	□□□□□□□	□□□	□□	□□□
1350489	□□□□□	□□□	□□	□□□
1351557	□□□□□□□	□□□	□□	□□□
1353766	□□□□□□□□	□□□	□□	□□□
1353776	□□□□□□	□□□	□□	□□□
1363053	□□□□□□	□□□	□□	□□□
1363749	□□□□□□□	□□□	□□	□□□
1367500	□□□□□	□□□	□□	□□□
1373225	□□□□□□□	□□□	□□	□□□
1373928	□□□□□□	□□□	□□	□□□
1386957	□□□□□□	□□□	□□	□□□
1395058	□□□□□□□	□□□	□□	□□□
1398937	□□□□□	□□□	□□	□□□
40226742	□□□□□□□	□□□	□□	□□□
40235485	□□□□□□□	□□□	□□	□□□

B.6

□□□□□□□□□□

□□□□□□□□□□□□

- □□□□□□□□□□□□□□□□□□□□□□□□□B.12□□□□□

□□□□□□□□□□□□□□□□□□□□365□□□□□□□□□□□365□□□□□□□□□□□□□□□□□□□

□□□□□

□□□□□□□□□□□□□□□

- □□□□□□□□□1□□□□□□□□□□□□□□□□1□□□□□□□□□□B.13□□□□□□□□□□□□
- □□□□□□□□□□□□□□□□□□□□365□□□0□□□□□□□□□□□□□□
B.14□□□□□□□□□□□□□□□□1□□□□□□□

□□ □□□□□□□□□□□□□□

□□□□□□□

□□

□□□□□□□□□□

0□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.12: □□□□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
907013	□□□□□	NO	YES	NO
974166	□□□□□□□□□□	NO	YES	NO
978555	□□□□□□	NO	YES	NO
1307863	□□□□□	NO	YES	NO
1308216	□□□□□□	NO	YES	NO
1308842	□□□□□□	NO	YES	NO
1310756	□□□□□□□	NO	YES	NO
1317640	□□□□□□□	NO	YES	NO
1318137	□□□□□□	NO	YES	NO
1318853	□□□□□□	NO	YES	NO
1319880	□□□□□□	NO	YES	NO
1326012	□□□□□□	NO	YES	NO
1328165	□□□□□□	NO	YES	NO
1331235	□□□□□	NO	YES	NO
1332418	□□□□□□	NO	YES	NO
1334456	□□□□□	NO	YES	NO
1335471	□□□□□□	NO	YES	NO
1340128	□□□□□□	NO	YES	NO
1341927	□□□□□□	NO	YES	NO
1342439	□□□□□□□□	NO	YES	NO
1346686	□□□□□□□	NO	YES	NO

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1347384	□□□□□□□	NO	YES	NO
1351557	□□□□□□□	NO	YES	NO
1353776	□□□□□□	NO	YES	NO
1363749	□□□□□□	NO	YES	NO
1367500	□□□□□	NO	YES	NO
1373225	□□□□□□□	NO	YES	NO
1395058	□□□□□□□	NO	YES	NO
40226742	□□□□□□□	NO	YES	NO
40235485	□□□□□□□	NO	YES	NO

Table B.13: □□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
904542	□□□□□□□	NO	YES	NO
907013	□□□□□	NO	YES	NO
932745	□□□□□	NO	YES	NO
942350	□□□□□	NO	YES	NO
956874	□□□□□	NO	YES	NO
970250	□□□□□□□□	NO	YES	NO
974166	□□□□□□□□□□	NO	YES	NO
978555	□□□□□□	NO	YES	NO
991382	□□□□□	NO	YES	NO
1305447	□□□□□	NO	YES	NO
1307046	□□□□□□□	NO	YES	NO
1307863	□□□□□	NO	YES	NO
1308216	□□□□□□	NO	YES	NO
1308842	□□□□□□	NO	YES	NO
1309068	□□□□□□	NO	YES	NO
1309799	□□□□□□	NO	YES	NO
1310756	□□□□□□□	NO	YES	NO
1313200	□□□□□	NO	YES	NO
1314002	□□□□□□	NO	YES	NO
1314577	□□□□□□	NO	YES	NO
1317640	□□□□□□□	NO	YES	NO
1317967	□□□□□□	NO	YES	NO
1318137	□□□□□□	NO	YES	NO
1318853	□□□□□□	NO	YES	NO
1319880	□□□□□□	NO	YES	NO
1319998	□□□□□□□	NO	YES	NO
1322081	□□□□□□□	NO	YES	NO
1326012	□□□□□□	NO	YES	NO
1327978	□□□□□□□	NO	YES	NO

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1328165	□□□□□□	NO	YES	NO
1331235	□□□□□	NO	YES	NO
1332418	□□□□□□	NO	YES	NO
1334456	□□□□□	NO	YES	NO
1335471	□□□□□□	NO	YES	NO
1338005	□□□□□□□	NO	YES	NO
1340128	□□□□□□	NO	YES	NO
1341238	□□□□□	NO	YES	NO
1341927	□□□□□□	NO	YES	NO
1342439	□□□□□□□□	NO	YES	NO
1344965	□□□□□□□	NO	YES	NO
1345858	□□□□□□	NO	YES	NO
1346686	□□□□□□□	NO	YES	NO
1346823	□□□□□□□	NO	YES	NO
1347384	□□□□□□□	NO	YES	NO
1350489	□□□□□	NO	YES	NO
1351557	□□□□□□□	NO	YES	NO
1353766	□□□□□□□□	NO	YES	NO
1353776	□□□□□□	NO	YES	NO
1363053	□□□□□□	NO	YES	NO
1363749	□□□□□□	NO	YES	NO
1367500	□□□□□	NO	YES	NO
1373225	□□□□□□□	NO	YES	NO
1373928	□□□□□□	NO	YES	NO
1386957	□□□□□□	NO	YES	NO
1395058	□□□□□□□	NO	YES	NO
1398937	□□□□□	NO	YES	NO
40226742	□□□□□□□	NO	YES	NO
40235485	□□□□□□□	NO	YES	NO

Table B.14: □□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
316866	□□□□□□	NO	YES	NO

B.7 3

□□□□□□B.6 □□□□□□□□□□□□□□□□□□□□□□□365□□□□**1095**□□□□□□□□□□□□□

B.8 ACE

□□□□□□□□□□

□□□□□□□□□□□□

- *ACE*□□□□□ B.15□□□□□□□□

□□□□□□□□□□□□□□□□□□□□0□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□

□□□

ACE□□□□□□□□□□□□□□ B.15□

- □□□□□□□30□□□□□□□□
- □□□□□□□0□□□□□□□

□□□□□□□□□□

30□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.15: ACE□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1308216	□□□□□□	NO	YES	NO
1310756	□□□□□□□	NO	YES	NO
1331235	□□□□□	NO	YES	NO
1334456	□□□□□	NO	YES	NO
1335471	□□□□□□	NO	YES	NO
1340128	□□□□□□	NO	YES	NO
1341927	□□□□□□	NO	YES	NO
1342439	□□□□□□□□	NO	YES	NO
1363749	□□□□□□	NO	YES	NO
1373225	□□□□□□□	NO	YES	NO

B.9 ARB

□□□□□□B.8□□□□□□□□□□□□□□□□□□□□□□*ARB*□□□B.16□□*ACE*□□□□□ B.15□□□□□□□□□□□□□

□□□□□□□□□□

Table B.16: □□□□□□□□□□□□□□□ARB□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1308842	□□□□□□	NO	YES	NO
1317640	□□□□□□□	NO	YES	NO
1346686	□□□□□□□	NO	YES	NO
1347384	□□□□□□□	NO	YES	NO
1351557	□□□□□□□	NO	YES	NO
1367500	□□□□□	NO	YES	NO
40226742	□□□□□□□	NO	YES	NO
40235485	□□□□□□□	NO	YES	NO

B.10

□□□□□□B.8 □□□□□□□□□□□□□□□□□□□□□□□□□□□□B.17□□*ACE*□□□□□ B.15□□□□□□□□□□□□□

□□□□□□□□□□

Table B.17: □□□□□□□□□□□□□□□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
907013	□□□□□	NO	YES	NO
974166	□□□□□□□□□□	NO	YES	NO
978555	□□□□□□	NO	YES	NO
1395058	□□□□□□□	NO	YES	NO

B.11 DCCB

□□□□□□ B.8 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□*DCCB*□□□ B.18□□*ACE*□□□□□ B.15□□□□□□□□□□□□□

□□□□□□□□□□

Table B.18: □□□□□□□□□□□□□□□□□□□□□□DCCB□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1318137	□□□□□□	NO	YES	NO
1318853	□□□□□□	NO	YES	NO
1319880	□□□□□□	NO	YES	NO
1326012	□□□□□□	NO	YES	NO

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1332418	□□□□□□	NO	YES	NO
1353776	□□□□□□	NO	YES	NO

B.12 NDCCB

□□□□□□B.8 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□*NDCCB*□□□ B.19□□*ACE*□□□□□ B.15□□□□□□□□□□□□□□

□□□□□□□□□□

Table B.19: □□□□□□□□□□□□□□□□□□□□□□NDCCB□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1307863	□□□□□	NO	YES	NO
1328165	□□□□□□	NO	YES	NO

B.13

□□□□□□B.8 □□□□□□□*β*□□□ (□B.20) □*ACE*□□□ (□B.15)□□□□□□□□□□□□

□□□□□□□□□□

Table B.20: β□□□

□□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1307046	□□□□□□□	NO	YES	NO
1313200	□□□□□	NO	YES	NO
1314002	□□□□□□	NO	YES	NO
1314577	□□□□□□	NO	YES	NO
1319998	□□□□□□□	NO	YES	NO
1322081	□□□□□□□	NO	YES	NO
1327978	□□□□□□□	NO	YES	NO
1338005	□□□□□□□	NO	YES	NO
1345858	□□□□□□	NO	YES	NO
1346823	□□□□□□□	NO	YES	NO
1353766	□□□□□□□□	NO	YES	NO
1386957	□□□□□□	NO	YES	NO

B.14

ACE□□□□□ B.8 □□□□□□□□□□□□□ (□ B.21) □*ACE*□□□ (□ B.15) □□□□□□□□□□□□

□□□□□□□□□□

Table B.21: □□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
932745	□□□□□	NO	YES	NO
942350	□□□□□	NO	YES	NO
956874	□□□□□	NO	YES	NO

B.15

ACE□□□□□ B.8 □□□□□□□□□□□□□□□□□ (□ B.22) □ *ACE*□□□ (□ B.15) □□□□□□□□□□□□

□□□□□□□□□□□

Table B.22: □□□□□□□□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
904542	□□□□□□□	NO	YES	NO
991382	□□□□□□	NO	YES	NO

B.16 1

ACE□□□□□B.8□□□□□□□*α1*□□□(□B.23)□*ACE*□□□(□B.15)□□□□□□□□□□□□

□□□□□□□□□□□

Table B.23: α1□□□

□□□□□ID	□□□□□□	□□	□□□	□□□□□□
1341238	□□□□□	NO	YES	NO
1350489	□□□□□	NO	YES	NO
1363053	□□□□□□	NO	YES	NO

Chapter C

— —

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

C.1 ACE

Table C.1: ACE□□□□ACEi□□□□□□□□□□□□□□□□□□□□□□THZ□□□□□

□□□□□ID	□□□□□□□□
434165	□□□□□□□□□
436409	□□□□
199192	□□□□□□□□□□□□□□□□/□□□□□□□
4088290	□□□□
4092879	□□□□
44783954	□□□□
75911	□□□□□□□
137951	□□□□□□□□
77965	□□□□□□
376707	□□□□□
4103640	□□□
73241	□□□□□□□□□□□
133655	□□□□□
73560	□□□□□
434327	□□□□
4213540	□□□□□□□□□
140842	□□□□□□□□
81378	□□□□□□□□
432303	□□□□□□
4201390	□□□□□□
46269889	□□□□□□□□□□□

□□□□□ID	□□□□□□□□
134438	□□□□□□
78619	□□□□□
201606	□□□□□
76786	□□□□□□□
4115402	□□□□
45757370	□□□□□□□□
433111	□□□□□
433527	□□□□□
4170770	□□□□□
4092896	□□□□□□□
259995	□□□□□□□
40481632	□□□□□□□□
4166231	□□□□□
433577	□□□
4231770	□□□□□□
440329	□□□□□□□□□□□□
4012570	□□□□□□□□□
4012934	□□□□□□□□
441788	□□□□□□□□□□□□□
4201717	□□□□□□
374375	□□□□
4344500	□□□□□□□□□□□□□□
139099	□□□
444132	□□□
196168	□□□□
432593	□□□□□□□
434203	□□□□□□
438329	□□□□□□□□□
195873	□□□□
4083487	□□□□□□□
4103703	□□□□□□□□
4209423	□□□□□□□
377572	□□□□□□□□□□□
40480893	□□□□□□□□□□□□□□□
136368	□□□□□□□□□□□
140648	□□□□□□□□□□□
438130	□□□□□□□
4091513	□□
4202045	□□□□□□□□□□□□
373478	□□
46286594	□□□□□□□□□□□□□□
439790	□□□□
81634	□□□□

□□□□□ID	□□□□□□□□
380706	□□□
141932	□□□□□□
36713918	□□□□□□□□□
443172	□□□□□□□□□□□□
81151	□□□□□
72748	□□□□□□□□□□□
378427	□□□□
437264	□□□□□□□□
194083	□□□□□□□□□
140641	□□□□□
440193	□□□□
4115367	□□□□

Chapter D

— —

1. □□
2. □□□□
3. □□
4. □□□□□
5. □□□□□□□
6. □□□□□□□□
7. □□□□
    - □□□□
    - □□□□
    - □□□□
    - □□□□
8. □□□□
    - □□□□□□
    - □□□□□□
    - □□□□□□
    - □□
    - □□□□□□□□□
    - □□□
9. □□□□□□□
    - □□□□□□□□
    - □□□□□
    - □□□□□□□□□□□□□□□
    - □□□□□□
    - □□
    - □□□□□□□
10. □□□□
    - □□□□□□□□□□□
    - □□□□□□□□□□
    - □□□□□□□□

11. □□□□□□□□□□
12. □□□□□□□□
13. □□□□□□□□□□□□□□□□□
14. □□□□□□□□□□□□□□□□□□□□□
15. □□□□□□□
16. □□□□

Chapter E

— —

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

E.1

□□ 4.1

□□□□□□□□□□□□□□□□□□□□□□□□□□□ E.1 □□□□□□□□□□□

Table E.1: PERSON□□□□□

□□□□	□	□□
PERSON_ID	2	□□□□□□
GENDER_CONCEPT_ID	8507	□□□□□□□□ ID □ 8507□
YEAR_OF_BIRTH	1974	
MONTH_OF_BIRTH	8	
DAY_OF_BIRTH	4	
BIRTH_DATETIME	1974-08-04 00:00:00	□□□□□□□□□0□□00:00:00□□□□
DEATH_DATETIME	NULL	
RACE_CONCEPT_ID	8516	□□□□□□□□□□□□□□□□ ID □ 8516□
ETHNICITY_CONCEPT_ID	38003564	38003564 □□□□□□□□□□□□□□
LOCATION_ID		□□□□□□□
PROVIDER_ID		□□□□□□□
CARE_SITE		□□□□□□□□□□□
PERSON_SOURCE_VALUE	NULL	□□□□□□□□□
GENDER_SOURCE_VALUE	Man	□□□□□□□□□□□□□

□□□□	□	□□
GENDER_SOURCE_ CONCEPT_ID	0	
RACE_SOURCE_ VALUE	African American	□□□□□□□□□□□□□
RACE_SOURCE_ CONCEPT_ID	0	
ETHNICITY_SOURCE_ VALUE	NULL	
ETHNICITY_SOURCE_ CONCEPT_ID	0	

□□ 4.2

□□□□□□□□□□□□□□□□□□□□□□□□□□□□ E.2 □□□□□□□□□□□

Table E.2: OBSERVATION_PERIOD□□□□□

□□□□	□	□□
OBSERVATION_ PERIOD_ID	2	□□□□□□
PERSON_ID	2	□□□PERSON□□□□□□□□□□□□□□□
OBSERVATION_PERIOD_ START_DATE	2015-01-01	□□□□□□□
OBSERVATION_PERIOD_ END_DATE	2019-07-01	□□□□□□□□□□□□□□□□□□□□□□□
PERIOD_TYPE_ CONCEPT_ID	44814722	44814724 □□□□□□□□□□□□□

□□ 4.3

□□□□□□□□□□□□□□□□□□□□□□□□□□□□ E.3 □□□□□□□□□□□

Table E.3: DRUG_EXPOSURE□□□□□

□□□□	□	□□
DRUG_EXPOSURE_ID	1001	□□□□□□
PERSON_ID	2	□□□PERSON□□□□□□□□□□□□□□□
DRUG_CONCEPT_ID	19078461	□□□□□NDC□□□□□□□□□□□ 19078461 □□□□□□□□□□
DRUG_EXPOSURE_ START_DATE	2019-05-01	□□□□□□□□□

□□□□	□	□□
DRUG_EXPOSURE_START_DATETIME	2019-05-01 00:00:00	□□□□□□□□□0□□□□□□
DRUG_EXPOSURE_END_DATE	2019-05-31	□□□ + □□□□□□□□□□
DRUG_EXPOSURE_END_DATETIME	2019-05-31 00:00:00	□□□□□□□□□0□□□□□□
VERBATIM_END_DATE	NULL	□□□□□□□□□□
DRUG_TYPE_CONCEPT_ID	38000177	38000177 □□□□□□□□□□□□□□□□
STOP_REASON	NULL	
REFILLS	NULL	
QUANTITY	NULL	□□□□□□□□□□
DAYS_SUPPLY	30	□□□□□□□□□□□□□□□
SIG	NULL	□□□□□□□□□□
ROUTE_CONCEPT_ID	4132161	4132161 □□□□□□□□□□
LOT_NUMBER	NULL	□□□□□□□□□□
PROVIDER_ID	NULL	□□□□□□□□□□
VISIT_OCCURRENCE_ID	NULL	□□□□□□□□□□□□□□□□□□□□□□□
VISIT_DETAIL_ID	NULL	
DRUG_SOURCE_VALUE	76168009520	□□□□□□NDC□□□□□
DRUG_SOURCE_CONCEPT_ID	583945	583945 □□□□□□□□□□□□□□□□□ (NDC□□□□76168009520□)□
ROUTE_SOURCE_VALUE	NULL	

□□ 4.4

□□□□□□□□□□□□□□□CONDITION_OCCURRENCE□□□□□□□□□□□□□□□□□□

```r
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT *
FROM @cdm.condition_occurrence
WHERE condition_concept_id = 192671;"

result <- renderTranslateQuerySql(connection, sql, cdm = "main")
head(result)
```

```
##   CONDITION_OCCURRENCE_ID PERSON_ID CONDITION_CONCEPT_ID ...
```

```
## 1                  4657       273              192671 ...
## 2                  1021        61              192671 ...
## 3                  5978       351              192671 ...
## 4                  9798       579              192671 ...
## 5                  9301       549              192671 ...
## 6                  1997       116              192671 ...
```

□□ 4.5

□□□□□□□□□□□□□□□CONDITION_OCCURRENCE□□□□□CONDITION_SOURCE_VALUE

```
sql <- "SELECT *
FROM @cdm.condition_occurrence
WHERE condition_source_value = 'K92.2';"

result <- renderTranslateQuerySql(connection, sql, cdm = "main")
head(result)
```

```
##   CONDITION_OCCURRENCE_ID PERSON_ID CONDITION_CONCEPT_ID ...
## 1                  4657       273              192671 ...
## 2                  1021        61              192671 ...
## 3                  5978       351              192671 ...
## 4                  9798       579              192671 ...
## 5                  9301       549              192671 ...
## 6                  1997       116              192671 ...
```

□□ 4.6

□□□□□OBSERVATION_PERIOD□□□□□□□□□□□□□□

```
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT *
FROM @cdm.observation_period
WHERE person_id = 61;"

renderTranslateQuerySql(connection, sql, cdm = "main")
```

```
##   OBSERVATION_PERIOD_ID PERSON_ID OBSERVATION_PERIOD_START_DATE ...
## 1                    61        61                    1968-01-21 ...
```

E.2

□□ 5.1

□□□□□ ID 192671 (“□□□□”)

□□ 5.2

ICD-10CM□□□□□

- K29.91 “□□□□□□□□□□□□□□□□□□□□□”
- K92.2 “□□□□□□□□□□□□□”

ICD-9CM□□□□□

- 578 “□□□□”
- 578.9 “□□□□□□□□□□□□□”

□□ 5.3

MedDRA□□□□□□

- “□□□□” (□□□□□ID 35707864)
- “□□□” (□□□□□ID 35707858)

E.3

□□ 6.1

A) □□□□□□□□CDM□□□□□□□□□ETL□□□□□□□
B) □□□□□□□□□□□□□□□□□□□□□□□□□
C) □□□□ETL□□□□□
D) □□□□□□□□□□□□□

□□ 6.2

□□□	□	□□
PERSON_ID	A123B456	□□□□□□□□□□□□□□□□□□□□
GENDER_CONCEPT_ID	8532	
YEAR_OF_BIRTH	NULL	□□□□□□□□□□□□□□□□□□□□
MONTH_OF_BIRTH	NULL	
DAY_OF_BIRTH	NULL	
RACE_CONCEPT_ID	0	□□□WHITE□□□□□□□8527□□□
ETHNICITY_CONCEPT_ID	8527	□□□□□□□□□□□□□□□□□□□0□
PERSON_SOURCE_VALUE	A123B456	
GENDER_SOURCE_VALUE	F	
RACE_SOURCE_VALUE	WHITE	
ETHNICITY_SOURCE_VALUE	NONE PROVIDED	

□□ 6.3

□□□	□
VISIT_OCCURRENCE_ID	1
PERSON_ID	11
VISIT_START_DATE	2004-09-26
VISIT_END_DATE	2004-09-30
VISIT_CONCEPT_ID	9201
VISIT_SOURCE_VALUE	inpatient

E.4

□□ 7.1

1. □□□□
2. □□□□□□□□
3. □□□□□□□□

□□ 7.2

□□□

E.5 SQL R

□□ 9.1

□□□□□□□□□□□□□PERSON□□□□□□□□□□□□□□□:

```r
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT COUNT(*) AS person_count
FROM @cdm.person;"

renderTranslateQuerySql(connection, sql, cdm = "main")
```

```
##   PERSON_COUNT
## 1         2694
```

□□ 9.2

□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□DRUG_EXPOSURE□□□□□□□□□□□

```r
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT COUNT(DISTINCT(person_id)) AS person_count
FROM @cdm.drug_exposure
INNER JOIN @cdm.concept_ancestor
  ON drug_concept_id = descendant_concept_id
INNER JOIN @cdm.concept ingredient
  ON ancestor_concept_id = ingredient.concept_id
WHERE LOWER(ingredient.concept_name) = 'celecoxib'
  AND ingredient.concept_class_id = 'Ingredient'
  AND ingredient.standard_concept = 'S';"

renderTranslateQuerySql(connection, sql, cdm = "main")
```

```
##   PERSON_COUNT
## 1         1844
```

COUNT(DISTINCT(person_id))□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□DRUG_ERA□□□□□□□□□□□□□□□□:

```r
library(DatabaseConnector)
connection <- connect(connectionDetails)

sql <- "SELECT COUNT(DISTINCT(person_id)) AS person_count
FROM @cdm.drug_era
INNER JOIN @cdm.concept ingredient
  ON drug_concept_id = ingredient.concept_id
WHERE LOWER(ingredient.concept_name) = 'celecoxib'
  AND ingredient.concept_class_id = 'Ingredient'
  AND ingredient.standard_concept = 'S';"

renderTranslateQuerySql(connection, sql, cdm = "main")
```

```
##   PERSON_COUNT
## 1         1844
```

□□ 9.3

□□□□□□□□□□□□□□□□□□□□□□□□□□□□CONDITION_OCCURRENCE□□□□□□□

```r
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT COUNT(*) AS diagnose_count
FROM @cdm.drug_era
INNER JOIN @cdm.concept ingredient
  ON drug_concept_id = ingredient.concept_id
```

```
INNER JOIN @cdm.condition_occurrence
  ON condition_start_date >= drug_era_start_date
    AND condition_start_date <= drug_era_end_date
INNER JOIN @cdm.concept_ancestor
  ON condition_concept_id =descendant_concept_id
WHERE LOWER(ingredient.concept_name) = 'celecoxib'
  AND ingredient.concept_class_id = 'Ingredient'
  AND ingredient.standard_concept = 'S'
  AND ancestor_concept_id = 192671;"

renderTranslateQuerySql(connection, sql, cdm = "main")
```

```
##   DIAGNOSE_COUNT
## 1             41
```

□□□□□DRUG_EXPOSURE□□□□□□□□DRUG_ERA□□□□□□□□□□□□□□□□□□□□□□

E.6

□□ 10.1

□□□□□□□□□□□□□□□□□□□□□□□□□□□:

- □□□□□□□□□□□□□
- 16□□□
- □□□□□□□□□365□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□E.1□□□□□□□□□

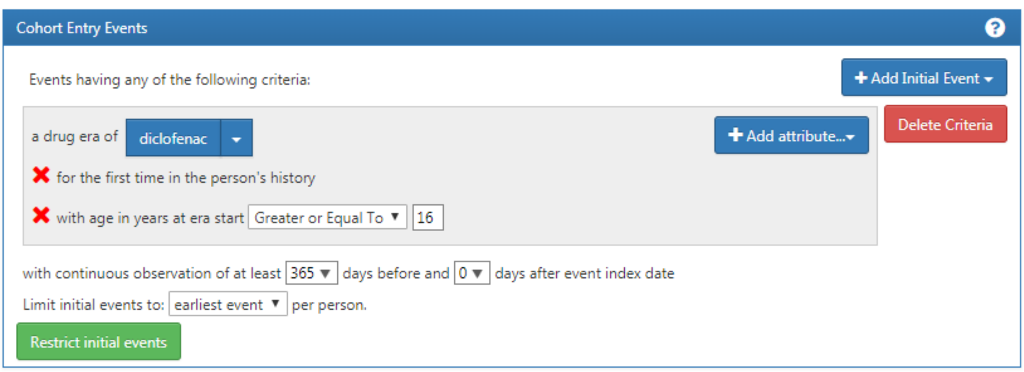


Figure E.1: □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□E.2□□□□□□□□□□Diclofenac□□□□□□□□□□□□□□□□□

□□□NSAID□□□□□□□□□□□□□□□□□□□□□□E.3□□□□□□□□□

NSAIDs□□□□□□□□□□□□□E.4□□□□□□□NSAIDs□□□□□□□□□□□□□□□□NSAID□□□□

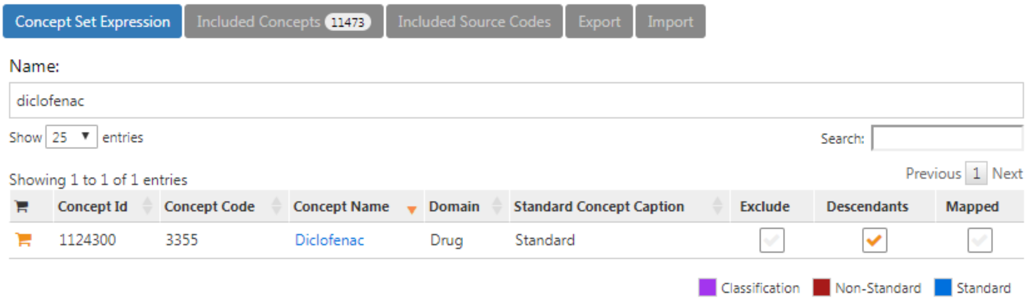


Figure E.2: □□□□□□□□□□□□□□□□□□□□

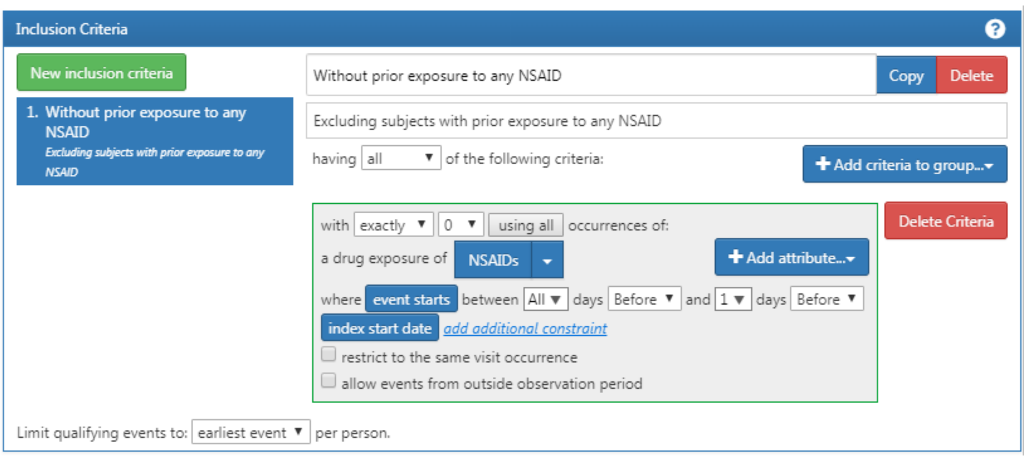


Figure E.3: NSAID□□□□□□□□□□□□

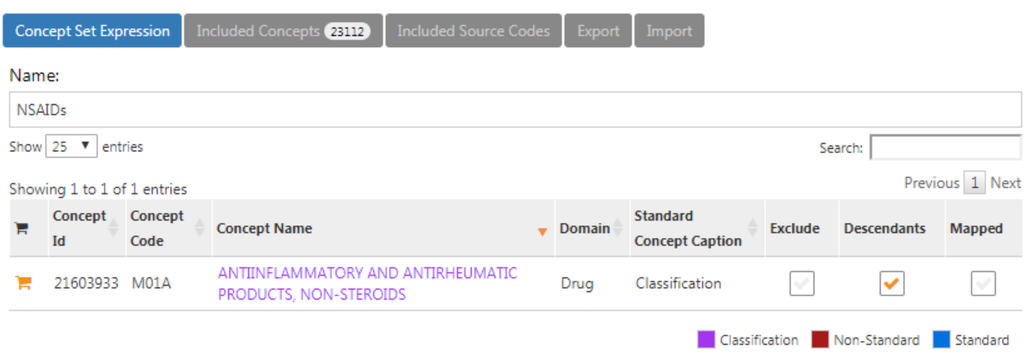


Figure E.4: NSAIDs□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□E.5□□□□□□□□□

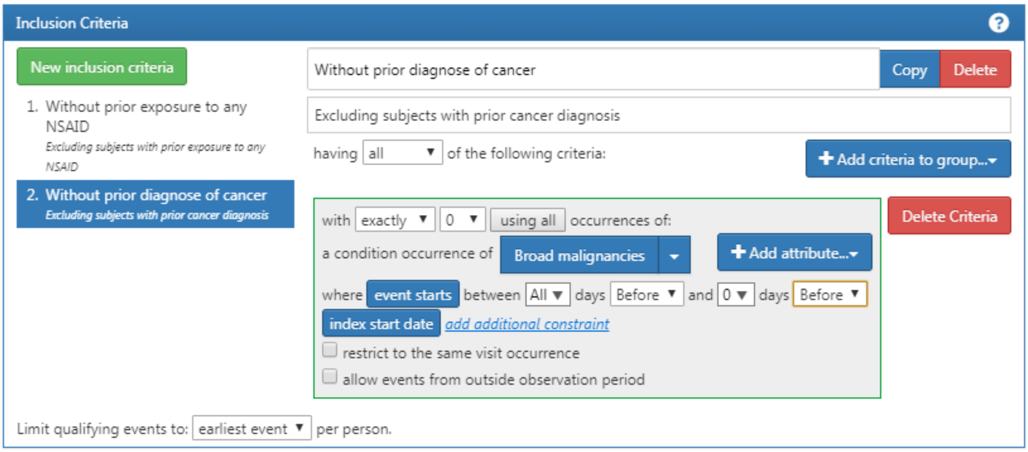


Figure E.5: □□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□E.6□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

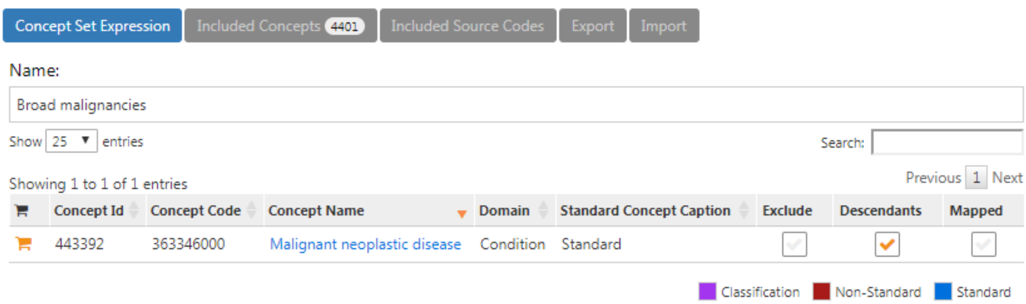


Figure E.6: □□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□30□□□□□□□□□□□□□□□□□□□□E.7□□□□□□

□□ 10.2

□□□□□□□□□SQL□2□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
library(DatabaseConnector)
connection <- connect(connectionDetails)
sql <- "SELECT person_id AS subject_id,
  condition_start_date AS cohort_start_date
INTO #diagnoses
FROM @cdm.condition_occurrence
WHERE condition_concept_id IN (
    SELECT descendant_concept_id
    FROM @cdm.concept_ancestor
```

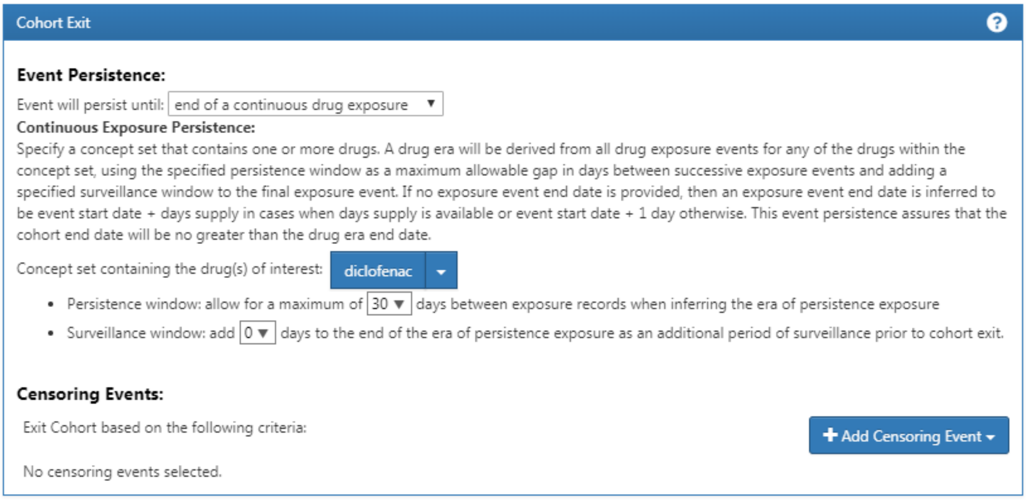


Figure E.7: □□□□□□□□□□□

```
    WHERE ancestor_concept_id = 4329847 --
)
  AND condition_concept_id NOT IN (
    SELECT descendant_concept_id
    FROM @cdm.concept_ancestor
    WHERE ancestor_concept_id = 314666 --
);"

renderTranslateExecuteSql(connection, sql, cdm = "main")
```

□□□□□□□□ER□□□□□□□□□□□□□□□□□□□□□□□□COHORT_DEFINITION_ID□□□□

```
sql <- "INSERT INTO @cdm.cohort (
  subject_id,
  cohort_start_date,
  cohort_definition_id
  )
SELECT subject_id,
  cohort_start_date,
  CAST (1 AS INT) AS cohort_definition_id
FROM #diagnoses
INNER JOIN @cdm.visit_occurrence
  ON subject_id = person_id
    AND cohort_start_date >= visit_start_date
    AND cohort_start_date <= visit_end_date
WHERE visit_concept_id IN (9201, 9203, 262); --        ;"

renderTranslateExecuteSql(connection, sql, cdm = "main")
```

□□□□□□VISIT_OCCURRENCE_ID□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□:

```
sql <- "TRUNCATE TABLE #diagnoses;
DROP TABLE #diagnoses;"

renderTranslateExecuteSql(connection, sql)
```

E.7

□□ 11.1

ATLAS□ Data Sources □□□ E.8 □□□

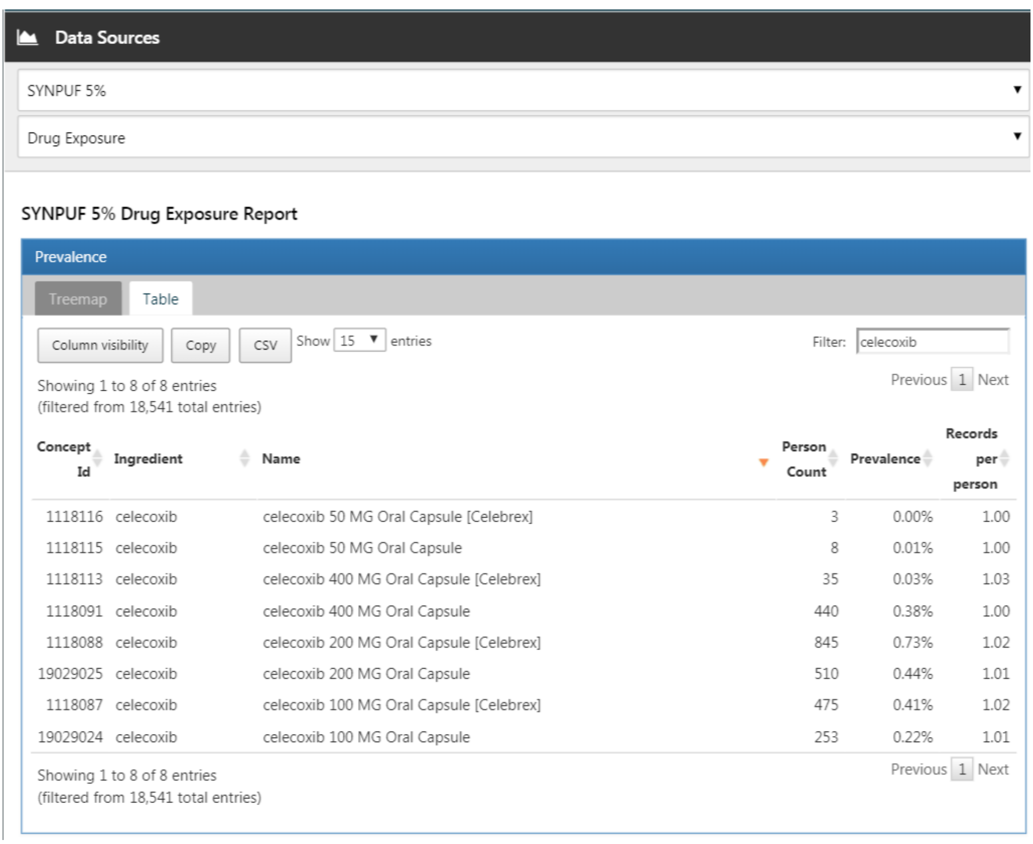


Figure E.8: □□□□□□□□□□□□

□□ 11.2

Cohort Definitions □□ Search □□□ □□□□□□□□□□□□□□□□□□□□□□□□□□□□□ E.9 □□□□

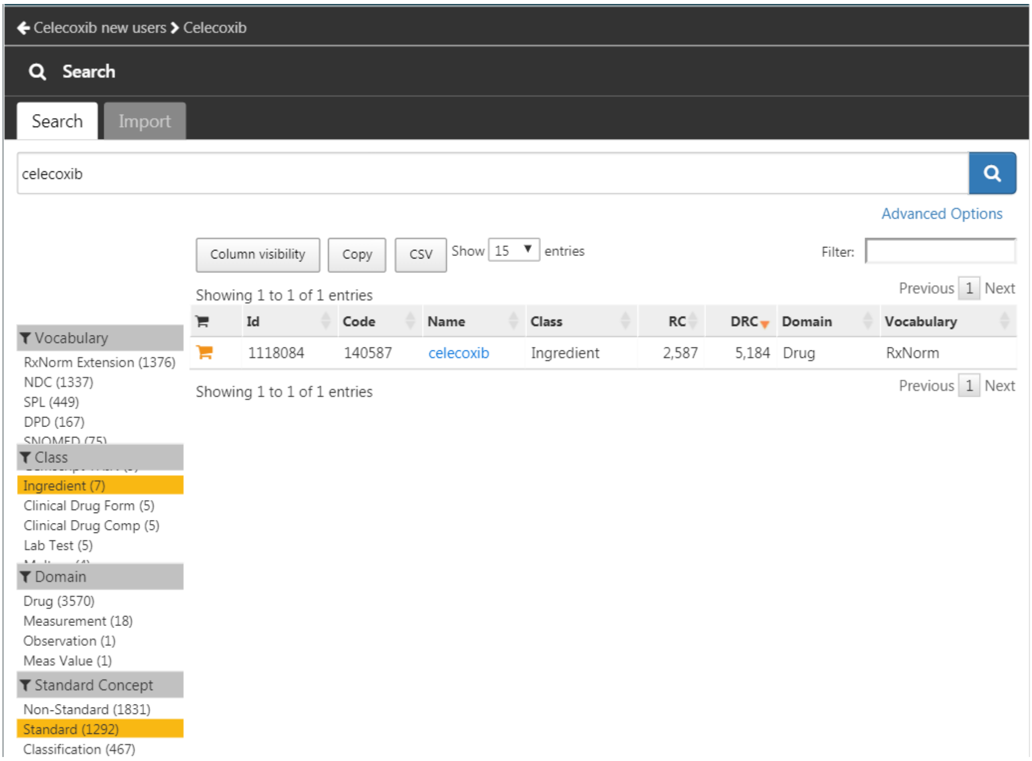


Figure E.9: □□□□□□□□□□□□□□□□□□□□□□

□ E.9 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□+□□□□□□□□□□□□□ E.10 □□ □□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□ Characterizations □□□□□□□□□□□□□ E.11 □□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ E.12 □□□□

□□ 11.3

Cohort Definitions □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□G □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□ E.13 □□□□

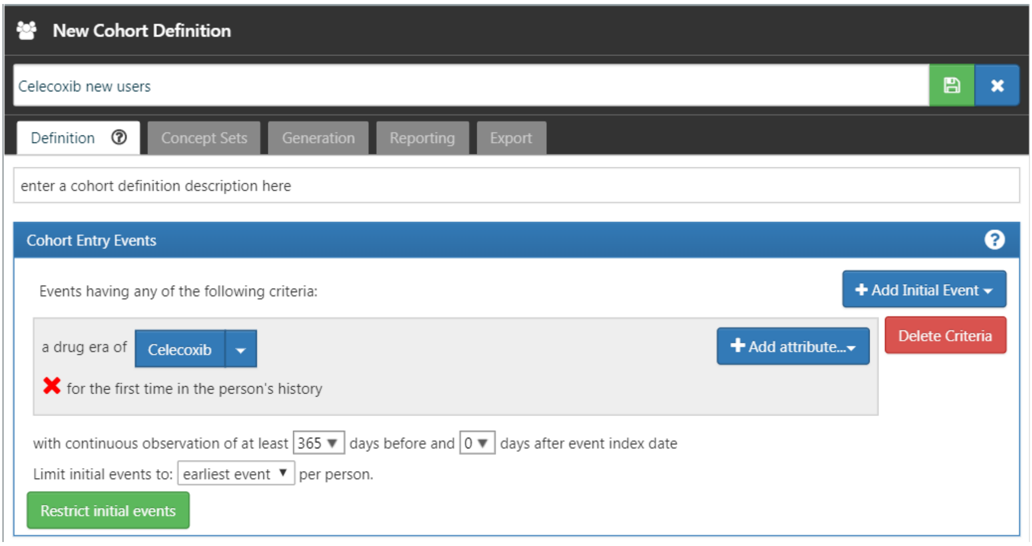


Figure E.10: □□□□□□□□□□□□□□□□□□□□□□

□ E.13 □□
E.14 □□□□

□□□□□□□□□□□+□□
E.15 □□□
□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□ Incidence Rates □□□□□□□□□□□□□□□□□□□
E.16 □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□1□□□□□□□□□□□□□□□□□□□□□□□□
E.17 □□□□

E.8

□□ 12.1

□□□□□□□□□□□□□□□□□□□□□□□□□□2□□□□□□□□□□□□□□□□□□□□□□□□

```r
library(CohortMethod)
nsaids <- c(1118084, 1124300) #
covSettings <- createDefaultCovariateSettings(
  excludedCovariateConceptIds = nsaids,
  addDescendantsToExclude = TRUE)

#
cmData <- getDbCohortMethodData(
  connectionDetails = connectionDetails,
```

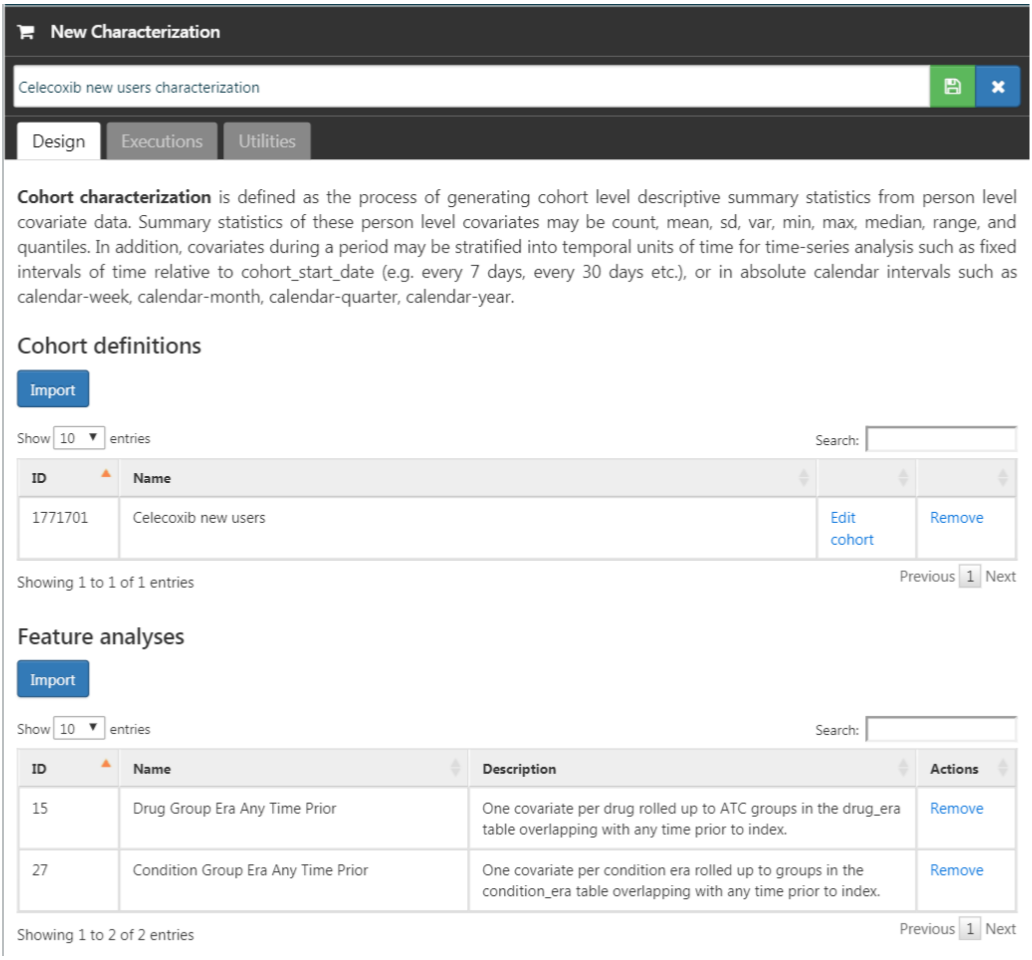

New Characterization
Celecoxib new users characterization
Design
Executions
Utilities
Cohort characterization is defined as the process of generating cohort level descriptive summary statistics from person level covariate data. Summary statistics of these person level covariates may be count, mean, sd, var, min, max, median, range, and quantiles. In addition, covariates during a period may be stratified into temporal units of time for time-series analysis such as fixed intervals of time relative to cohort_start_date (e.g. every 7 days, every 30 days etc.), or in absolute calendar intervals such as calendar-week, calendar-month, calendar-quarter, calendar-year.
Cohort definitions
Import
Show 10 entries
Search:
ID
Name
1771701
Celecoxib new users
Edit cohort
Remove
Showing 1 to 1 of 1 entries
Previous 1 Next
Feature analyses
Import
Show 10 entries
Search:
ID
Name
Description
Actions
15
Drug Group Era Any Time Prior
One covariate per drug rolled up to ATC groups in the drug_era table overlapping with any time prior to index.
Remove
27
Condition Group Era Any Time Prior
One covariate per condition era rolled up to groups in the condition_era table overlapping with any time prior to index.
Remove
Showing 1 to 2 of 2 entries
Previous 1 Next

Figure E.11: □□□□□□□

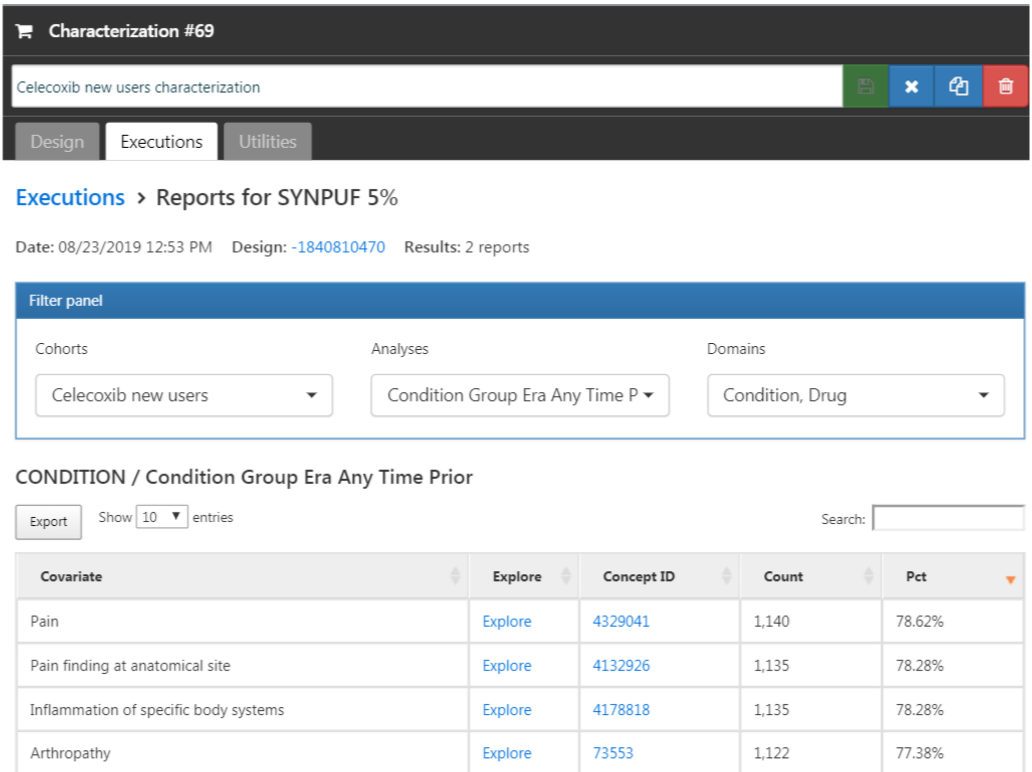


Figure E.12: □□□□□□□

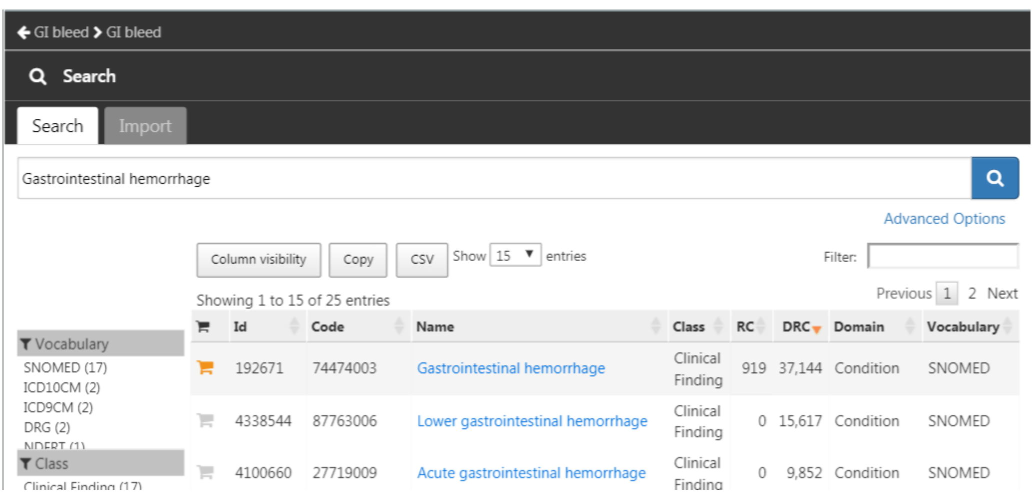


Figure E.13: □□□□□□□□□□□□□□□□□□□□

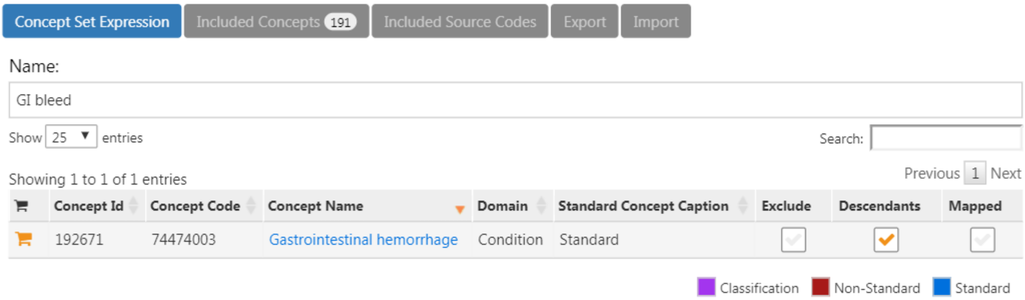


Figure E.14: □□□□□□□□□□□□□□□□□□

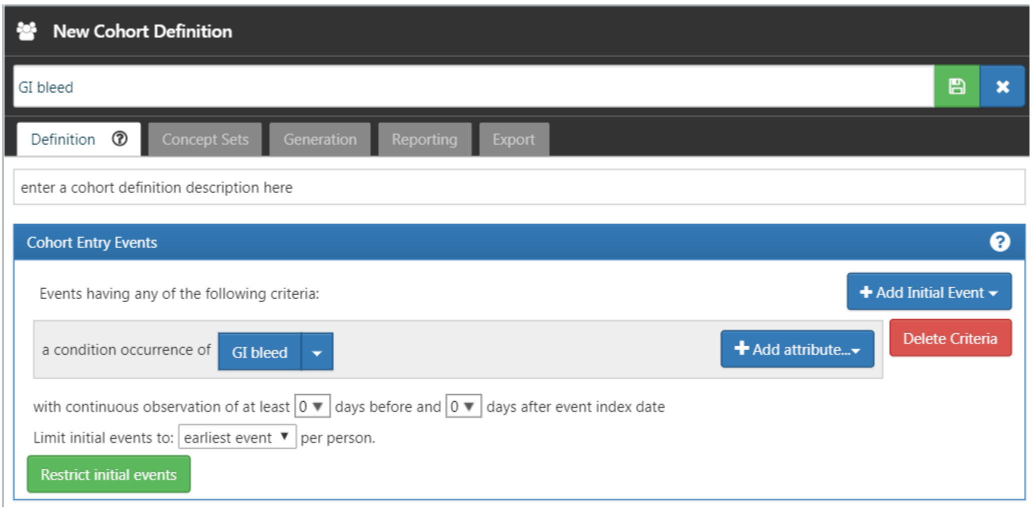


Figure E.15: □□□□□□□□□□□□□□□

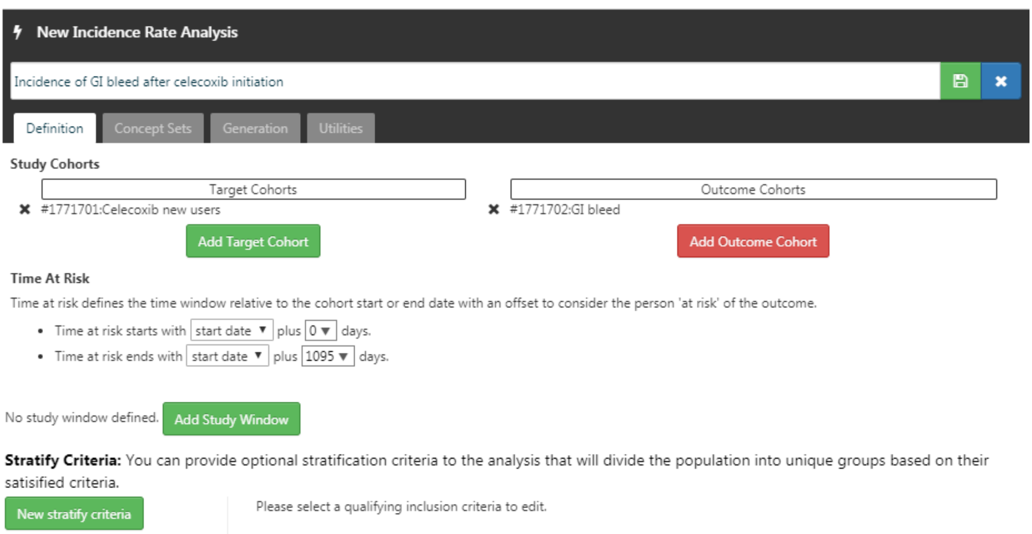


Figure E.16: □□□□□□

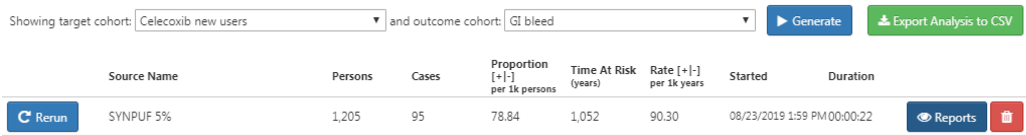


Figure E.17: □□□□□

```
  cdmDatabaseSchema = "main",
  targetId = 1,
  comparatorId = 2,
  outcomeIds = 3,
  exposureDatabaseSchema = "main",
  exposureTable = "cohort",
  outcomeDatabaseSchema = "main",
  outcomeTable = "cohort",
  covariateSettings = covSettings)
summary(cmData)
```

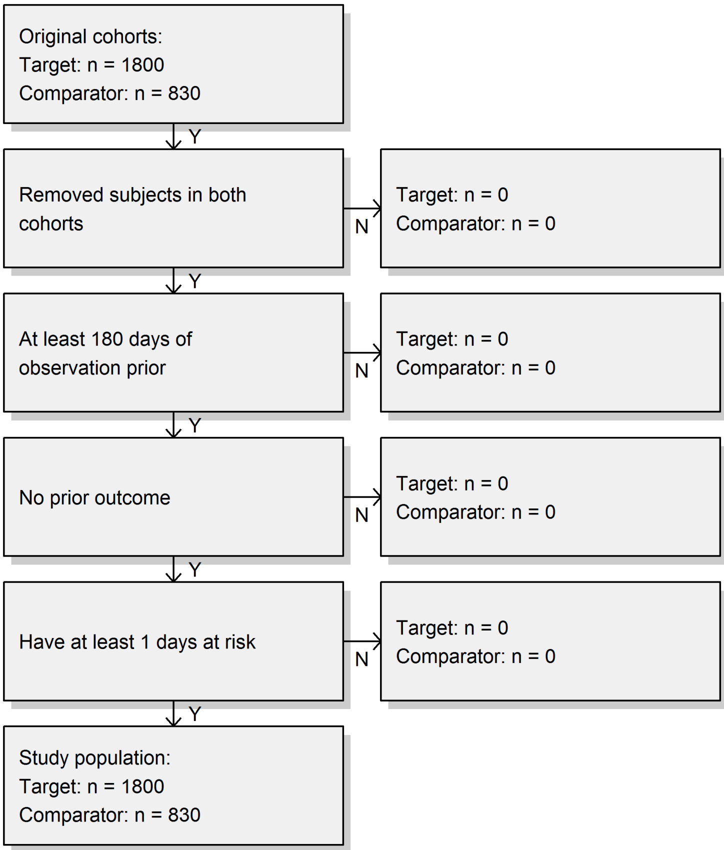
```
## CohortMethodData
##
##      ID 1
##      ID 2
##      ID 3
##
##       1800
##      830
##
##
##
## 3          479        479
##
##
##     389
##       26923
```

□□ 12.2

□□□□□□□□□□□□□□□□□□□□□□□□□

```
studyPop <- createStudyPopulation(
  cohortMethodData = cmData,
  outcomeId = 3,
  washoutPeriod = 180,
  removeDuplicateSubjects = "remove all",
  removeSubjectsWithPriorOutcome = TRUE,
```

```
  riskWindowStart = 0,
  startAnchor = "cohort start",
  riskWindowEnd = 99999)
drawAttritionDiagram(studyPop)
```



□□□

□□ 12.3

Cox □□□□□□□□□□□□□□□□□□□□□□□□□

```
model <- fitOutcomeModel(population = studyPop,
                         modelType = "cox")
model
```

```
##     cox
##   FALSE
```

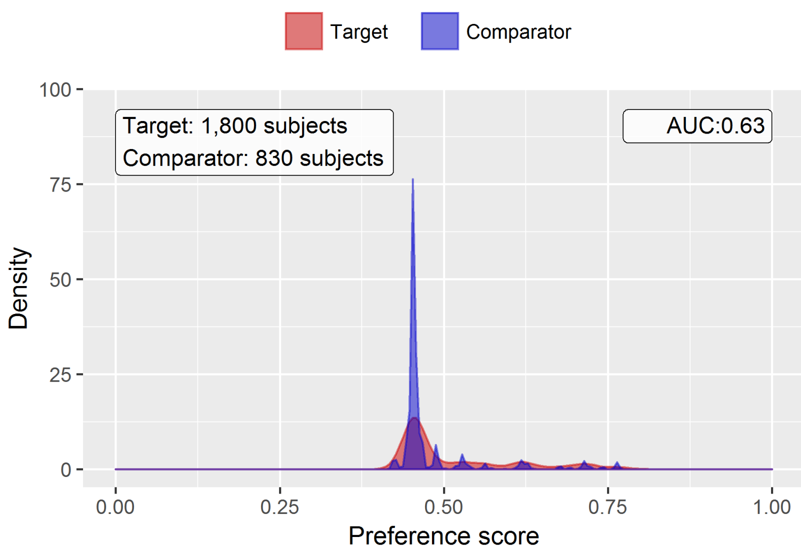
```
##     FALSE
##           FALSE
##     OK
##
##                 .95    .95
##          1.34612     1.10065     1.65741  0.29723     0.1044
```

□□□

□□ **12.4**

□□□

```
ps <- createPs(cohortMethodData = cmData,
               population = studyPop)
plotPs(ps, showCountsLabel = TRUE, showAucLabel = TRUE)
```



□□□

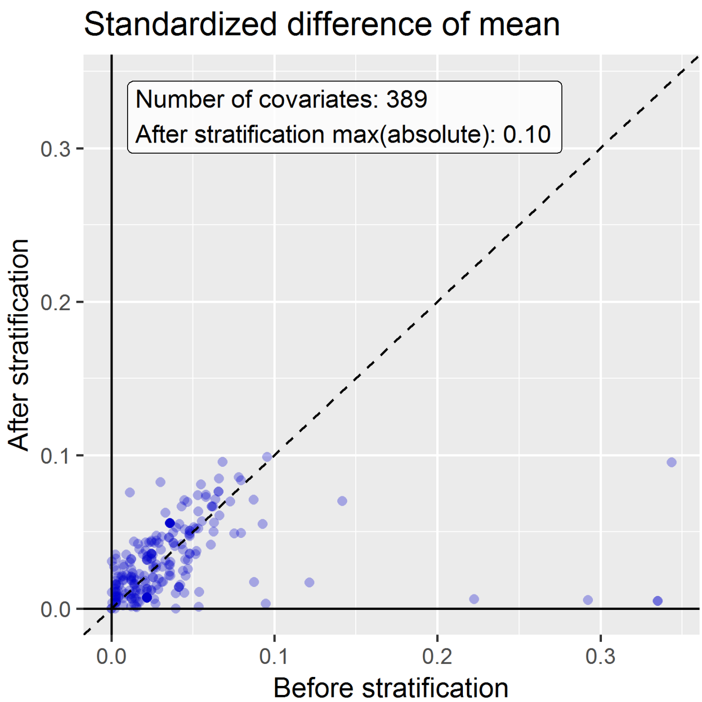
□□□□□□□□□0.63□AUC□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ **12.5**

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

```
strataPop <- stratifyByPs(ps, numberOfStrata = 5)
bal <- computeCovariateBalance(strataPop, cmData)
plotCovariateBalanceScatterPlot(bal,
                                showCovariateCountLabel = TRUE,
```

```
                                showMaxLabel = TRUE,
                                beforeLabel = "   ",
                                afterLabel = "   ")
```



□□□□□□□□□□□□□□□□□□□□□x□□□□□□□>0.3□□□□□□□□□□□□□□□□□□□□
0.1□□□

□□ 12.6

PS□□□□ Cox □□□□□□□□□□□□□□□

```
adjModel <- fitOutcomeModel(population = strataPop,
                            modelType = "cox",
                            stratified = TRUE)
adjModel
```

```
##       cox
##     TRUE
##       FALSE
##             FALSE
##      OK
##
##                     .95    .95
##           1.13211     0.92132     1.40008  0.12409     0.1068
```

□□□□□□□□□□□□□□□□□□□□□□95□□□□□□1□□□□□□□□□□□□□□□□□□□□□□2□

E.9

□□ 13.1

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□getPlpData□□□□□□□□□

```
library(PatientLevelPrediction)
covSettings <- createCovariateSettings(
  useDemographicsGender = TRUE,
  useDemographicsAge = TRUE,
  useConditionGroupEraLongTerm = TRUE,
  useConditionGroupEraAnyTimePrior = TRUE,
  useDrugGroupEraLongTerm = TRUE,
  useDrugGroupEraAnyTimePrior = TRUE,
  useVisitConceptCountLongTerm = TRUE,
  longTermStartDays = -365,
  endDays = -1)

plpData <- getPlpData(connectionDetails = connectionDetails,
                      cdmDatabaseSchema = "main",
                      cohortDatabaseSchema = "main",
                      cohortTable = "cohort",
                      cohortId = 4,
                      covariateSettings = covSettings,
                      outcomeDatabaseSchema = "main",
                      outcomeTable = "cohort",
                      outcomeIds = 3)

summary(plpData)
```

```
## plpData
##
##          ID -1
##      ID 3
##
##   2630
##
##
##
## 3           479        479
##
##
##     245
##        54079
```

□□ **13.2**

□□□

```r
population <- createStudyPopulation(plpData = plpData,
                                    outcomeId = 3,
                                    washoutPeriod = 364,
                                    firstExposureOnly = FALSE,
                                    removeSubjectsWithPriorOutcome = TRUE,
                                    priorOutcomeLookback = 9999,
                                    riskWindowStart = 1,
                                    riskWindowEnd = 365,
                                    addExposureDaysToStart = FALSE,
                                    addExposureDaysToEnd = FALSE,
                                    minTimeAtRisk = 364,
                                    requireTimeAtRisk = TRUE,
                                    includeAllOutcomes = TRUE)
nrow(population)
```

```
## [1] 2578
```

□□□□□□□□□□□□□□□□□□□□□□□□□□364□□□□□□□□□□□□□□□□□□

□□ **13.3**

LASSO □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□runPlp□□□□□□□□□□

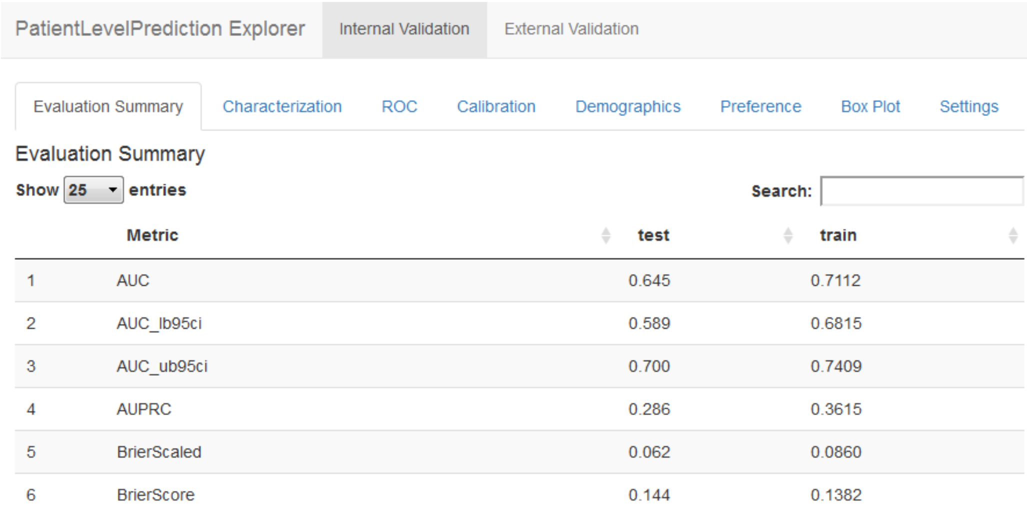
```r
lassoModel <- setLassoLogisticRegression(seed = 0)

lassoResults <- runPlp(population = population,
                       plpData = plpData,
                       modelSettings = lassoModel,
                       testSplit = 'person',
                       testFraction = 0.25,
                       nfold = 2,
                       splitSeed = 0)
```

□□□□□□LASSO□□□

Shiny□□□□□□□□□□□□□□□□□□□□□□□□

```r
viewPlp(lassoResults)
```

□□□□□□□E.18□□□□□□□□□□□□□□□□□□□□□□□□□□□□AUC□0.645□□□□□□□



	Metric	test	train
1	AUC	0.645	0.7112
2	AUC_lb95ci	0.589	0.6815
3	AUC_ub95ci	0.700	0.7409
4	AUPRC	0.286	0.3615
5	BrierScaled	0.062	0.0860
6	BrierScore	0.144	0.1382

Figure E.18: □□□□□□□Shiny□□□□

E.10

□□ 15.1

ACHILLES□□□□□□□:

```
library(ACHILLES)
result <- achilles(connectionDetails,
                   cdmDatabaseSchema = "main",
                   resultsDatabaseSchema = "main",
                   sourceName = "Eunomia",
                   cdmVersion = "5.3.0")
```

□□ 15.2

□□□□□□□□□□□□□□□□□□□□:

```
DataQualityDashboard::executeDqChecks(
  connectionDetails,
  cdmDatabaseSchema = "main",
  resultsDatabaseSchema = "main",
  cdmSourceName = "Eunomia",
  outputFolder = "C:/dataQualityExample")
```

□□ 15.3

□□□□□□□□□□□□□□□□□□□:

```r
DataQualityDashboard::viewDqDashboard(
  "C:/dataQualityExample/Eunomia/results_Eunomia.json")
```

E.11

Bibliography

Allison, D. B., Brown, A. W., George, B. J., and Kaiser, K. A. (2016). Reproducibility: A tragedy of errors. *Nature*, 530(7588):27–29.

Arnold, B. F., Ercumen, A., Benjamin-Chung, J., and Colford, J. M. (2016). Brief Report: Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies. *Epidemiology*, 27(5):637–641.

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics*, 10(2):150–161.

Banda, J. M., Halpern, Y., Sontag, D., and Shah, N. H. (2017). Electronic phenotyping with APHRODITE and the Observational Health Sciences and Informatics (OHDSI) data network. *AMIA Jt Summits Transl Sci Proc*, 2017:48–57.

Boland, M. R., Parhi, P., Li, L., Miotto, R., Carroll, R., Iqbal, U., Nguyen, P. A., Schuemie, M., You, S. C., Smith, D., Mooney, S., Ryan, P., Li, Y. J., Park, R. W., Denny, J., Dudley, J. T., Hripcsak, G., Gentine, P., and Tatonetti, N. P. (2017). Uncovering exposures responsible for birth season - disease effects: a global study. *J Am Med Inform Assoc*.

Botsis, T., Hartvigsen, G., Chen, F., and Weng, C. (2010). Secondary use of ehr: data quality issues and informatics opportunities. *Summit on Translational Bioinformatics*, 2010:1.

Byrd, J. B., Adam, A., and Brown, N. J. (2006). Angiotensin-converting enzyme inhibitor-associated angioedema. *Immunol Allergy Clin North Am*, 26(4):725–737.

Callahan, T. J., Bauck, A. E., Bertoch, D., Brown, J., Khare, R., Ryan, P. B., Staab, J., Zozus, M. N., and Kahn, M. G. (2017). A comparison of data quality assessment checks in six data sharing networks. *eGEMs*, 5(1).

Cepeda, M. S., Reps, J., Fife, D., Blacketer, C., Stang, P., and Ryan, P. (2018). Finding treatment-resistant depression in real-world data: How a data-driven approach compares with expert-based heuristics. *Depress Anxiety*, 35(3):220–228.

Chen, X., Dallmeier-Tiessen, S., Dasler, R., Feger, S., Fokianos, P., Gonzalez, J. B., Hirvonsalo, H., Kousidis, D., Lavasa, A., Mele, S., Rodriguez, D. R., Šimko, T., Smith, T., Trisovic, A., Trzcinska, A., Tsanaktsidis, I., Zimmermann, M., Cranmer, K., Heinrich, L., Watts, G., Hildreth, M., Iglesias, L. L., Lassila-Perini, K., and Neubert, S. (2018). Open is not enough. *Nature Physics*, 15(2):113–119.

Cicardi, M., Zingale, L. C., Bergamaschini, L., and Agostoni, A. (2004). Angioedema associated with angiotensin-converting enzyme inhibitor use: outcome after switching to a different treatment. *Arch. Intern. Med.*, 164(8):910–913.

Dasu, T. and Johnson, T. (2003). *Exploratory data mining and data cleaning*, volume 479. John Wiley & Sons.

Defalco, F. J., Ryan, P. B., and Soledad Cepeda, M. (2013). Applying standardized drug terminologies to observational healthcare databases: a case study on opioid exposure. *Health Serv Outcomes Res Methodol*, 13(1):58–67.

Duke, J. D., Ryan, P. B., Suchard, M. A., Hripcsak, G., Jin, P., Reich, C., Schwalm, M. S., Khoma, Y., Wu, Y., Xu, H., Shah, N. H., Banda, J. M., and Schuemie, M. J. (2017). Risk of angioedema associated with levetiracetam compared with phenytoin: Findings of the observational health data sciences and informatics research network. *Epilepsia*, 58(8):e101–e106.

Farrington, C. P. (1995). Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*, 51(1):228–235.

Farrington, C. P., Anaya-Izquierdo, K., Whitaker, H. J., Hocine, M. N., Douglas, I., and Smeeth, L. (2011). Self-controlled case series analysis with event-dependent observation periods. *Journal of the American Statistical Association*, 106(494):417–426.

Garza, M., Del Fiol, G., Tenenbaum, J., Walden, A., and Zozus, M. N. (2016). Evaluating common data models for use with a longitudinal community registry. *J Biomed Inform*, 64:333–341.

Hernan, M. A. and Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am. J. Epidemiol.*, 183(8):758–764.

Hersh, W. R., Weiner, M. G., Embi, P. J., Logan, J. R., Payne, P. R., Bernstam, E. V., Lehmann, H. P., Hripcsak, G., Hartzog, T. H., Cimino, J. J., et al. (2013). Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical care*, 51(8 0 3):S30.

Hill, A. B. (1965). THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc. R. Soc. Med.*, 58:295–300.

Hripcsak, G. and Albers, D. J. (2017). High-fidelity phenotyping: richness and freedom from bias. *J Am Med Inform Assoc*.

Hripcsak, G., Duke, J. D., Shah, N. H., Reich, C. G., Huser, V., Schuemie, M. J., Suchard, M. A., Park, R. W., Wong, I. C. K., Rijnbeek, P. R., van der Lei, J., Pratt, N., Norén, G. N., Li, Y.-C., Stang, P. E., Madigan, D., and Ryan, P. B. (2015). Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Studies in health technology and informatics*, 216:574–578.

Hripcsak, G., Levine, M. E., Shang, N., and Ryan, P. B. (2018). Effect of vocabulary mapping for conditions on phenotype cohorts. *J Am Med Inform Assoc*, 25(12):1618–1625.

Hripcsak, G., Ryan, P. B., Duke, J. D., Shah, N. H., Park, R. W., Huser, V., Suchard, M. A., Schuemie, M. J., DeFalco, F. J., Perotte, A., Banda, J. M., Reich, C. G., Schilling, L. M., Matheny, M. E., Meeker, D., Pratt, N., and Madigan, D. (2016). Characterizing treatment pathways at scale using the OHDSI network. *Proceedings of the National Academy of Sciences*, 113(27):7329–7336.

Hripcsak, G., Shang, N., Peissig, P. L., Rasmussen, L. V., Liu, C., Benoit, B., Carroll, R. J., Carrell, D. S., Denny, J. C., Dikilitas, O., Gainer, V. S., Marie Howell, K., Klann, J. G., Kullo, I. J., Lingren, T., Mentch, F. D., Murphy, S. N., Natarajan, K., Pacheco, J. A., Wei, W. Q., Wiley, K., and Weng, C. (2019). Facilitating phenotype transfer using a common data model. *J Biomed Inform*, page 103253.

Huser, V., DeFalco, F. J., Schuemie, M., Ryan, P. B., Shang, N., Velez, M., Park, R. W., Boyce, R. D., Duke, J., Khare, R., Utidjian, L., and Bailey, C. (2016). Multisite Evaluation of a Data Quality Tool for Patient-Level Clinical Data Sets. *EGEMS (Washington, DC)*, 4(1):1239.

Huser, V., Kahn, M. G., Brown, J. S., and Gouripeddi, R. (2018). Methods for examining data quality in healthcare integrated data repositories. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 23:628–633.

Johnston, S. S., Morton, J. M., Kalsekar, I., Ammann, E. M., Hsiao, C. W., and Reps, J. (2019). Using Machine Learning Applied to Real-World Healthcare Data for Predictive Analytics: An Applied Example in Bariatric Surgery. *Value Health*, 22(5):580–586.

Kahn, M. G., Brown, J. S., Chun, A. T., Davidson, B. N., Meeker, D., Ryan, P. B., Schilling, L. M., Weiskopf, N. G., Williams, A. E., and Zozus, M. N. (2015). Transparent reporting of data quality in distributed data networks. *EGEMS (Washington, DC)*, 3(1):1052.

Kahn, M. G., Callahan, T. J., Barnard, J., Bauck, A. E., Brown, J., Davidson, B. N., Estiri, H., Goerg, C., Holve, E., Johnson, S. G., Liaw, S.-T., Hamilton-Lopez, M., Meeker, D., Ong, T. C., Ryan, P. B., Shang, N., Weiskopf, N. G., Weng, C., Zozus, M. N., and Schilling, L. (2016). A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS (Washington, DC)*, 4(1):1244.

Kahn, M. G., Raebel, M. A., Glanz, J. M., Riedlinger, K., and Steiner, J. F. (2012). A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. *Medical care*, 50.

Liaw, S.-T., Rahimi, A., Ray, P., Taggart, J., Dennis, S., de Lusignan, S., Jalaludin, B., Yeo, A., and Talaei-Khoei, A. (2013). Towards an ontology for data quality in integrated chronic disease management: a realist review of the literature. *International journal of medical informatics*, 82(1):10–24.

Lipsitch, M., Tchetgen Tchetgen, E., and Cohen, T. (2010). Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*, 21(3):383–388.

Maclure, M. (1991). The case-crossover design: a method for studying transient effects on the risk of acute events. *Am. J. Epidemiol.*, 133(2):144–153.

Madigan, D., Ryan, P. B., and Schuemie, M. (2013a). Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies. *Ther Adv Drug Saf*, 4(2):53–62.

Madigan, D., Ryan, P. B., Schuemie, M., Stang, P. E., Overhage, J. M., Hartzema, A. G., Suchard, M. A., DuMouchel, W., and Berlin, J. A. (2013b). Evaluating the impact of database heterogeneity on observational study results. *Am. J. Epidemiol.*, 178(4):645–651.

Magid, D. J., Shetterly, S. M., Margolis, K. L., Tavel, H. M., O'Connor, P. J., Selby, J. V., and Ho, P. M. (2010). Comparative effectiveness of angiotensin-converting enzyme inhibitors versus beta-blockers as second-line therapy for hypertension. *Circ Cardiovasc Qual Outcomes*, 3(5):453–458.

Makadia, R. and Ryan, P. B. (2014). Transforming the Premier Perspective Hospital Database into the Observational Medical Outcomes Partnership (OMOP) Common Data Model. *EGEMS (Wash DC)*, 2(1):1110.

Martin, R. C. (2008). *Clean Code: A Handbook of Agile Software Craftsmanship*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1 edition.

Matcho, A., Ryan, P., Fife, D., and Reich, C. (2014). Fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model. *Drug Saf*, 37(11):945–959.

Noren, G. N., Caster, O., Juhlin, K., and Lindquist, M. (2014). Zoo or savannah? Choice of training ground for evidence-based pharmacovigilance. *Drug Saf*, 37(9):655–659.

Norman, J. L., Holmes, W. L., Bell, W. A., and Finks, S. W. (2013). Life-threatening ACE inhibitor-induced angioedema after eleven years on lisinopril. *J Pharm Pract*, 26(4):382–388.

Olsen, L., Aisner, D., McGinnis, J. M., et al. (2007). *The learning healthcare system: workshop summary*. Natl Academy Pr.

O'Mara, N. B. and O'Mara, E. M. (1996). Delayed onset of angioedema with angiotensin-converting enzyme inhibitors: case report and review of the literature. *Pharmacotherapy*, 16(4):675–679.

Overhage, J. M., Ryan, P. B., Reich, C. G., Hartzema, A. G., and Stang, P. E. (2012). Validation of a common data model for active safety surveillance research. *J Am Med Inform Assoc*, 19(1):54–60.

Powers, B. J., Coeytaux, R. R., Dolor, R. J., Hasselblad, V., Patel, U. D., Yancy, W. S., Gray, R. N., Irvine, R. J., Kendrick, A. S., and Sanders, G. D. (2012). Updated report on comparative effectiveness of ACE inhibitors, ARBs, and direct renin inhibitors for patients with essential hypertension: much more data, little new information. *J Gen Intern Med*, 27(6):716–729.

Ramcharran, D., Qiu, H., Schuemie, M. J., and Ryan, P. B. (2017). Atypical Antipsychotics and the Risk of Falls and Fractures Among Older Adults: An Emulation Analysis and an Evaluation of Additional Confounding Control Strategies. *J Clin Psychopharmacol*, 37(2):162–168.

Rassen, J. A., Shelat, A. A., Myers, J., Glynn, R. J., Rothman, K. J., and Schneeweiss, S. (2012). One-to-many propensity score matching in cohort studies. *Pharmacoepidemiol Drug Saf*, 21 Suppl 2:69–80.

Reps, J. M., Rijnbeek, P. R., and Ryan, P. B. (2019). Identifying the DEAD: Development and Validation of a Patient-Level Model to Predict Death Status in Population-Level Claims Data. *Drug Saf*.

Reps, J. M., Schuemie, M. J., Suchard, M. A., Ryan, P. B., and Rijnbeek, P. R. (2018). Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data. *Journal of the American Medical Informatics Association*, 25(8):969–975.

Roebuck, K. (2012). *Data quality: high-impact strategies-what you need to know: definitions, adoptions, impact, benefits, maturity, vendors*. Emereo Publishing.

Rosenbaum, P. and Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55.

Rubbo, B., Fitzpatrick, N. K., Denaxas, S., Daskalopoulou, M., Yu, N., Patel, R. S., Hemingway, H., Danesh, J., Allen, N., Atkinson, M., Blaveri, E., Brannan, R., Brayne, C., Brophy, S., Chaturvedi, N., Collins, R., deLusignan, S., Denaxas, S., Desai, P., Eastwood, S., Gallacher, J., Hemingway, H., Hotopf, M., Landray, M., Lyons, R., O'Neil, T., Pringle, M., Sprosen, T., Strachan, D., Sudlow, C., Sullivan, F., Zhang, Q., and Flaig, R. (2015). Use of electronic health records to ascertain, validate and phenotype acute myocardial infarction: A systematic review and recommendations. *Int. J. Cardiol.*, 187:705–711.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3-4):169–188.

Ryan, P. B., Buse, J. B., Schuemie, M. J., DeFalco, F., Yuan, Z., Stang, P. E., Berlin, J. A., and Rosenthal, N. (2018). Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A real-world meta-analysis of 4 observational databases (OBSERVE-4D). *Diabetes Obes Metab*, 20(11):2585–2597.

Ryan, P. B., Madigan, D., Stang, P. E., Overhage, J. M., Racoosin, J. A., and Hartzema, A. G. (2012). Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership. *Stat Med*, 31(30):4401–4415.

Ryan, P. B., Schuemie, M. J., and Madigan, D. (2013a). Empirical performance of a self-controlled cohort method: lessons for developing a risk identification and analysis system. *Drug Saf*, 36 Suppl 1:95–106.

Ryan, P. B., Schuemie, M. J., Ramcharran, D., and Stang, P. E. (2017). Atypical Antipsychotics and the Risks of Acute Kidney Injury and Related Outcomes Among Older Adults: A Replication Analysis and an Evaluation of Adapted Confounding Control Strategies. *Drugs Aging*, 34(3):211–219.

Ryan, P. B., Stang, P. E., Overhage, J. M., Suchard, M. A., Hartzema, A. G., DuMouchel, W., Reich, C. G., Schuemie, M. J., and Madigan, D. (2013b). A comparison of the empirical performance of methods for a risk identification system. *Drug Saf*, 36 Suppl 1:S143–158.

Sabroe, R. A. and Black, A. K. (1997). Angiotensin-converting enzyme (ACE) inhibitors and angio-oedema. *Br. J. Dermatol.*, 136(2):153–158.

Schuemie, M. J., Hripcsak, G., Ryan, P. B., Madigan, D., and Suchard, M. A. (2016). Robust empirical calibration of p-values using observational data. *Stat Med*, 35(22):3883–3888.

Schuemie, M. J., Hripcsak, G., Ryan, P. B., Madigan, D., and Suchard, M. A. (2018a). Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data. *Proc. Natl. Acad. Sci. U.S.A.*, 115(11):2571–2577.

Schuemie, M. J., Ryan, P. B., DuMouchel, W., Suchard, M. A., and Madigan, D. (2014). Interpreting observational studies: why empirical calibration is needed to correct p-values. *Stat Med*, 33(2):209–218.

Schuemie, M. J., Ryan, P. B., Hripcsak, G., Madigan, D., and Suchard, M. A. (2018b). Improving reproducibility by using high-throughput observational studies with empirical calibration. *Philos Trans A Math Phys Eng Sci*, 376(2128).

Sherman, R. E., Anderson, S. A., Dal Pan, G. J., Gray, G. W., Gross, T., Hunter, N. L., LaVange, L., Marinac-Dabic, D., Marks, P. W., Robb, M. A., et al. (2016). Real-world evidence—what is it and what can it tell us. *N Engl J Med*, 375(23):2293–2297.

Simpson, S. E., Madigan, D., Zorych, I., Schuemie, M. J., Ryan, P. B., and Suchard, M. A. (2013). Multiple self-controlled case series for large-scale longitudinal observational databases. *Biometrics*, 69(4):893–902.

Slater, E. E., Merrill, D. D., Guess, H. A., Roylance, P. J., Cooper, W. D., Inman, W. H. W., and Ewan, P. W. (1988). Clinical Profile of Angioedema Associated With Angiotensin Converting-Enzyme Inhibition. *JAMA*, 260(7):967–970.

Stang, P. E., Ryan, P. B., Racoosin, J. A., Overhage, J. M., Hartzema, A. G., Reich, C., Welebob, E., Scarnecchia, T., and Woodcock, J. (2010). Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann. Intern. Med.*, 153(9):600–606.

Suchard, M. A., Simpson, S. E., Zorych, I., Ryan, P. B., and Madigan, D. (2013). Massive parallelization of serial inference algorithms for a complex generalized linear model. *ACM Trans. Model. Comput. Simul.*, 23(1):10:1–10:17.

Suissa, S. (1995). The case-time-control design. *Epidemiology*, 6(3):248–253.

Swerdel, J. N., Hripcsak, G., and Ryan, P. B. (2019). PheValuator: Development and Evaluation of a Phenotype Algorithm Evaluator. *J Biomed Inform*, page 103258.

Thompson, T. and Frable, M. A. (1993). Drug-induced, life-threatening angioedema revisited. *Laryngoscope*, 103(1 Pt 1):10–12.

Tian, Y., Schuemie, M. J., and Suchard, M. A. (2018). Evaluating large-scale propensity score performance through real-world and synthetic data experiments. *Int J Epidemiol*, 47(6):2005–2014.

Toh, S., Reichman, M. E., Houstoun, M., Ross Southworth, M., Ding, X., Hernandez, A. F., Levenson, M., Li, L., McCloskey, C., Shoaibi, A., Wu, E., Zornberg, G., and Hennessy, S. (2012). Comparative risk for angioedema associated with the use of drugs that target the renin-angiotensin-aldosterone system. *Arch. Intern. Med.*, 172(20):1582–1589.

van der Lei, J. (1991). Use and abuse of computer-stored medical records. *Methods of information in medicine*, 30(02):79–80.

Vandenbroucke, J. P. and Pearce, N. (2012). Case-control studies: basic concepts. *Int J Epidemiol*, 41(5):1480–1489.

Vashisht, R., Jung, K., Schuler, A., Banda, J. M., Park, R. W., Jin, S., Li, L., Dudley, J. T., Johnson, K. W., Shervey, M. M., Xu, H., Wu, Y., Natrajan, K., Hripcsak, G., Jin, P., Van Zandt, M., Reckard, A., Reich, C. G., Weaver, J., Schuemie, M. J., Ryan, P. B., Callahan, A., and Shah, N. H. (2018). Association of Hemoglobin A1c Levels With Use of Sulfonylureas, Dipeptidyl Peptidase 4 Inhibitors, and Thiazolidinediones in Patients With Type 2 Diabetes Treated With Metformin: Analysis From the Observational Health Data Sciences and Informatics Initiative. *JAMA Netw Open*, 1(4):e181755.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., and Vandenbroucke, J. P. (2008). The strengthening the reporting of observational studies in epidemiology (strobe) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 61(4):344 – 349.

Voss, E. A., Boyce, R. D., Ryan, P. B., van der Lei, J., Rijnbeek, P. R., and Schuemie, M. J. (2016). Accuracy of an Automated Knowledge Base for Identifying Drug Adverse Reactions. *J Biomed Inform*.

Voss, E. A., Ma, Q., and Ryan, P. B. (2015a). The impact of standardizing the definition of visits on the consistency of multi-database observational health research. *BMC Med Res Methodol*, 15:13.

Voss, E. A., Makadia, R., Matcho, A., Ma, Q., Knoll, C., Schuemie, M., DeFalco, F. J., Londhe, A., Zhu, V., and Ryan, P. B. (2015b). Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *J Am Med Inform Assoc*, 22(3):553–564.

Walker, A. M., Patrick, A. R., Lauer, M. S., Hornbrook, M. C., Marin, M. G., Platt, R., Roger, V. L., Stang, P., and Schneeweiss, S. (2013). A tool for assessing the feasibility of comparative effectiveness research. *Comp Eff Res*, 3:11–20.

Wang, Y., Desai, M., Ryan, P. B., DeFalco, F. J., Schuemie, M. J., Stang, P. E., Berlin, J. A., and Yuan, Z. (2017). Incidence of diabetic ketoacidosis among patients with type 2 diabetes mellitus treated with SGLT2 inhibitors and other antihyperglycemic agents. *Diabetes Res. Clin. Pract.*, 128:83–90.

Weinstein, R. B., Ryan, P., Berlin, J. A., Matcho, A., Schuemie, M., Swerdel, J., Patel, K., and Fife, D. (2017). Channeling in the Use of Nonprescription Paracetamol and Ibuprofen in an Electronic Medical Records Database: Evidence and Implications. *Drug Saf*, 40(12):1279–1292.

Weiskopf, N. G. and Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *Journal of the American Medical Informatics Association: JAMIA*, 20(1):144–151.

Whelton, P. K., Carey, R. M., Aronow, W. S., Casey, D. E., Collins, K. J., Dennison Himmelfarb, C., DePalma, S. M., Gidding, S., Jamerson, K. A., Jones, D. W., MacLaughlin, E. J., Muntner, P., Ovbiagele, B., Smith, S. C., Spencer, C. C., Stafford, R. S., Taler, S. J., Thomas, R. J., Williams, K. A., Williamson, J. D., and Wright, J. T. (2018). 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*, 138(17):e426–e483.

Whitaker, H. J., Farrington, C. P., Spiessens, B., and Musonda, P. (2006). Tutorial in biostatistics: the self-controlled case series method. *Stat Med*, 25(10):1768–1797.

Who, A. (2013). Global brief on hypertension. *World Health Organization*.

Wikipedia (2019a). Open science — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Open%20science&oldid=900178688. [Online; accessed 24-June-2019].

Wikipedia (2019b). Science 2.0 — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Science%202.0&oldid=887565958. [Online; accessed 09-July-2019].

Wikiquote (2019). Ronald fisher — wikiquote,. [Online; accessed 2-August-2019].

Yoon, D., Ahn, E. K., Park, M. Y., Cho, S. Y., Ryan, P., Schuemie, M. J., Shin, D., Park, H., and Park, R. W. (2016). Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for Distributed Network Research. *Healthc Inform Res*, 22(1):54–58.

Yuan, Z., DeFalco, F. J., Ryan, P. B., Schuemie, M. J., Stang, P. E., Berlin, J. A., Desai, M., and Rosenthal, N. (2018). Risk of lower extremity amputations in people with type 2 diabetes mellitus treated with sodium-glucose co-transporter-2 inhibitors in the USA: A retrospective cohort study. *Diabetes Obes Metab*, 20(3):582–589.

Zaadstra, B. M., Chorus, A. M., van Buuren, S., Kalsbeek, H., and van Noort, J. M. (2008). Selective association of multiple sclerosis with infectious mononucleosis. *Mult. Scler.*, 14(3):307–313.

Zaman, M. A., Oparil, S., and Calhoun, D. A. (2002). Drugs targeting the renin-angiotensin-aldosterone system. *Nat Rev Drug Discov*, 1(8):621–636.

Index

ACE□□□, 159
ACHILLES, 184
adaboost, 155
agnostic SQL, *see* SqlRender
analysis implementation, 63
APHRODITE, 95
arachne, 222
ATHENA, 33
ATLAS, 64, 160
    characterization features, 120
    □□□□□□, 66
    □□□□□□□□□, 65
    □□□□□□, 65
    □□□□□□□□, 65
    □□□, 65
    □□□□□□, 65
    □□□□□□, 64
    □□□□□□, 66
    □□□□□□□, 65
    □□□□□□, 65
    □□□□□□□□, 64
    □□□□□□□, 65
    □□□□□, 65
    □□□, 65
    □□, 65
    □□□□□□□, 65
attrition diagram, 146
AUC, 158

back-propagation, 156
baseline time, 111
best practice for network research, 222
bigknn, 155
Bill of Mortality, 31

CDM , *see* Common Data Model
characterization, 57, 111
    cohort, 111
    database level, 111
    treatment pathways, 112
classification concept, 35
Clem McDonald, 182
cohort
    entry event, 94
    exit criteria, 94
    inclusion criteria, 94
    probabilistic design, 95
    rule-based design, 94
Common Data Model, 19
    □□□□□□□□, 21
    □□□□□□, 21
    □□□□□□□, 19
    □□□□□, 19
    □□□□□□□, 21
    □□□□, 21
    □□□□□, 21
    □□□□□□, 21
    □□□□□, 19
    □□□□, 19
community, 5, 182
comparative effect estimation, 133
comparative effectiveness, *see* comparative effect estimation
concept, 33
    class, 34
    code, 35
    hierarchy, 38
    identifier, 33
    mapping, 37

relationship, 37
concept set, 95
control hypotheses, 66
convolutional neural network, 156
counterfactual, 133
covariate balance
example, 146
cross-validation, 154
Cyclops, 153

data profiling, *see* White Rabbit
data quality, 182, 183
DatabaseConnector, 71
creating a connection, 79
querying, 79
decision boundary, 153
decision tree, 155
deep learning, 156
descriptive statistics, *see* characterization
design considerations for network research, 220
direct effect estimation, 133
disease natural history, *see* characterization
domain
concept, 34
drug utilization, 112

ETL, *see* □□□□□□□□□□ETL□
implementations, 51
quality control, 51
□□□□□, 185
ETL design, *see* Rabbit-In-A-Hat
evidence quality, 181, 182

FAIR, 14
feature analyses, 116
FeatureExtraction, 120
forum, 7

gradient boosting, 154

hyper-parameter, 154

incidence, 112
proportion, 113
rate, 113

index date, 111

k-nearest neighbors, 155
Kaplan-Meier plot, 146

LASSO, 153
logistic regression, 153
logistics of network research, 220

methods library, 66
minimum detectable relative risk (MDRR), 146
mission, 4

naive bayes, 155
network study, 219
neural network, 156
no free lunch, 153

objectives, 5
OHDSI Methods Benchmark, 210
OHDSI SQL, *see* SqlRender

p-value calibration, 207
Pallas system, 31
patient-level prediction, 59
perceptron, 156
person-time, 113
phenotype library, 96
PheValuator, 195
population-level estimation, 58, 133
post-index time, 111
power, 146
propensity model
example, 146
python, 154, 155

quality improvement, *see* characterization
Query Library, 71
QueryLibrary, 84

R, 71
installation, 68
Rabbit-In-A-Hat, 47
random forest, 154
recurrent neural networks, 156
regularization, 153

regulatory decision-making, 182
reliable evidence, 182
ROC, 158
running network research, 221

safety surveillance, 133
sklearn, 154
source code mapping, *see* Usagi
SQL, 71
SQL Query Library, *see* Query Library
SqlRender, 71
  debugging, 78
  parameterization, 72
  supported functions, 74
  translation, 73
standard concept, 35
Standard SQL Dialect, *see* SqlRender
standardized vocabularies, 31
  download, 33
  search, 33
structured query language, *see* SQL
study-a-thon, 13
supervised learning(□□□□), 152
survival plot, *see* Kaplan-Meier plot
system requirements, 68

tools deployment, 68
  Amazon AWS, 68
  Broadsea, 68
treatment utilization, *see* characterization
TRIPOD, 151

Usagi, 49

variance, 154
vignette, 68
vision, 4
vocabulary, 34

White Rabbit, 43

xgboost, 154
xSens cohort, 196
xSpec cohort, 196

□□□□□□□□□, 151
  SCCS □□□□, 137
  □□□□□□□□□, 133
  □□□□□□□□□, 136
  □□□□□□□□□, 136
  □□□□□□□□□□□□□, 135
□□□□□□□□□□, 152
□□□□□□□, 159
□□□□□□□, 158
□□□□□□□□□, 13
  open standards, 14
  □□□□□□□, 14
  □□□□□□□□□□, 14
  □□□□□□□, 14
□□□□, 135
□□□□□□□□□, 158
□□□, 152
□□□□□□, 134
□□□□□□□□□□□□□, *see* □□□□□□
□□□□, 93
□□□□□□□□, 133
□□□□□□, 93
□□□□□□
  □□□□□□□□□, 8
□□□□□□□□□, 201
□□□□□
  □□, 39
□□□□□□, 93
□□□□□□□□□□, 201
□□□□□□□□□□□□, 201
□□□□□□, *see* □□□□
□□□□□□□□□, 194
□□□□□□□□□, 151
  SCCS □□□□, 137
  □□□□□□□□□, 133
  □□□□□□□□□, 136
  □□□□□□□□□, 136
  □□□□□□□□□□□□□, 135
□□□□□
  □□□□□□□□□, 187
  □□□□, 184
  □□□□□□□□□, 184
  □□□□□□□, 184
  □□□, 184
  □□□, 184
  □□, 184

  □□, 184
□□□□□□□□□□□□, 184
□□□□□□□□, *see* □□□□
□□□□□□□□□□□, 206
□□□□□□□□□□
  □□□□□□□□, 136
□□□□, *see* □□□□□□□
□□□□□□□□□, 157
□□□□□□, 93
□□□□□□□□□□□□□□□□□, 201
□□□□□, 215
□□□□□□, 134
□□□□□□□□□□□, 206
  □□, 206
□□□, 152
□□□□□□□, 134
□□□□□, 151
□□□□□□□□□□□□□, *see* Common Data Model
□□□□□□□□□, 134

□□□□□□□, 151
□□□□□, 151
□□□□□□□□, 156
□□□□, 157
□□□□, 134
□□□□□□, 8

□□□□□, 135
□□□, 157
□□□, 157
□□□□□, 134
  □□□□□, 134
  □□, 134
  □□□□, 134
□□□□□□□□, 134
□□□, 152
□□□□□□□, 135
□□□□□□□□
  □□□□□□□, 23
  □□, 21
□□□□□□□□, 134
□□□, 152
□□□□□□□□□□□□, 207
□□□□□, 207
□□□□□□□, 134

□□□□□□□, 151
□□, 157
□□□□□□, 205
□□
  □□□□, 156
  □□□□, 156
  □□□□□, 157
  □□□□□, 157
□□□□, 151
□□□□□, 60, 153
□□□□□□□□, 133

□□□, 157
□□□□, 43
□□-□□-□□□□□□□□□□, 136
□□□□□□□□, 136
□□□□□□□□, 136
□□□, 157
□□□, 157
□□□□□□□□, 219
□□□□□□□, 216
□□□□, 205
□□, 157

□□□□, 135
□□□□□□, 151
□□□□□□, 193
□□□□□□□□□□□□, 135
□□□□□□□□□□ (SCCS) □□□□, 137
□□□□□□□, 60
□□□□□□□, 151
□□□, 158

□□□□, 134
□□□-□□□□□□□□□, 158
□□□□□, 157