



How to build an OHDSI Study

Mui Van Zandt



Put your idea in the Forums

OHDSI Women of OHDSI - CALL FOR RESEARCH QUESTIONS

General



Patrick_Ryan

May 19

WoO, I'm glad to see your workgroup coming together to support each other in moving forward a goal to generate reliable evidence.

I support whatever question you ultimately settle on and would be delighted to help in any way I can once you decide on a study.

To add some additional study ideas to the table, here's a type of prediction question that could be informative, for which I think our OHDSI data network could usefully contribute:

The US Preventative Services Task Force recommends regular screening for women for a variety of conditions, including breast, cervical, colorectal (colon) cancers. For each of these screenings, there is some diagnostic procedure performed which can detect the presence of the condition at that time. If a person tests positive, some additional diagnostics and then treatment intervention can be considered; if a person tests negative, the person is recommended to return in some time interval to be retested.

1. Woman aged 30 to 65 are recommended to be screened for cervical cancer every 3-5 years with cervical cytology and/or hrHPV testing.
(<https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/cervical-cancer-screening2> 1)
2. All women aged 50 to 74 are recommended to be screened for breast cancer every 2 years with mammography, but there remains debate about screening mammography when aged 40-49, as it can depend on patient preference toward the benefit-risk tradeoff
(<https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/breast-cancer-screening1> 3)
3. Adults (men or women) aged 50 to 75 are recommended for colorectal cancer screening through multiple methods under different frequency intervals.
(<https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/colorectal-cancer-screening2>)

<https://forums.ohdsi.org/c/researchers>



How to formulate a Question?

- Target cohort
- Outcome cohort
- Time At Risk

Among patients who **[insert patient cohort]**, which patients will go on to have **[insert outcome of interest]** within **[time window]**?

Women of OHDSI Study Initiated:

Amongst **[women aged 40-74 who undergo a screening mammography who do not have prior breast cancer]**, which patients will go on **[to develop breast cancer]** in the **[90d to 3 years following the screening mammography]**?



Why this question matters?

- Demonstrates the risk of developing breast cancer between screenings
 - Encourages patients who underestimate their risk to get regular mammograms
 - Helps patients who overestimate their risk to understand their true likelihood of developing breast cancer
- Ultimately allows patients to make informed, confident decisions about preventative care



Study Protocol

- Target cohort
- Outcome cohort
- Time At Risk

- Study Populations

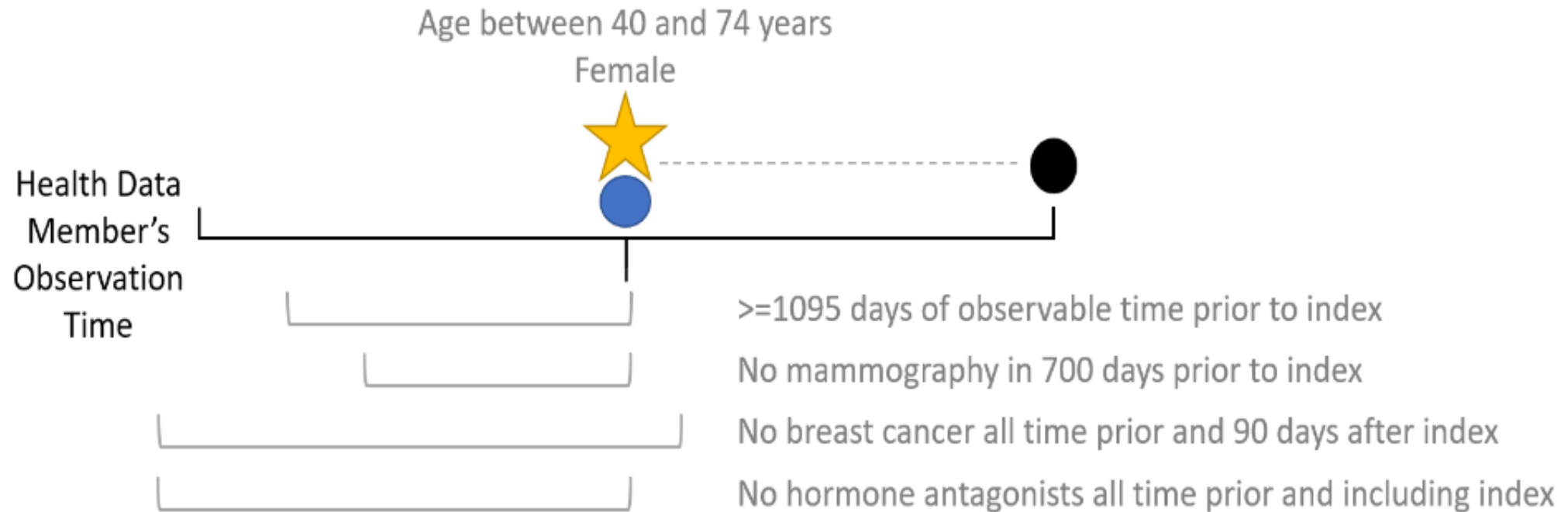
Projects Item	Definition
Target (T)	Women aged 40-74 who are undergo a screening mammography who do not have prior breast cancer.
Outcome (O)	Individuals who develop breast cancer
Time at Risk (TAR)	90 days after index day to 1095 days after index day



Study Protocol

-  Target cohort
-  Outcome cohort
-  Time At Risk

- Target Cohort



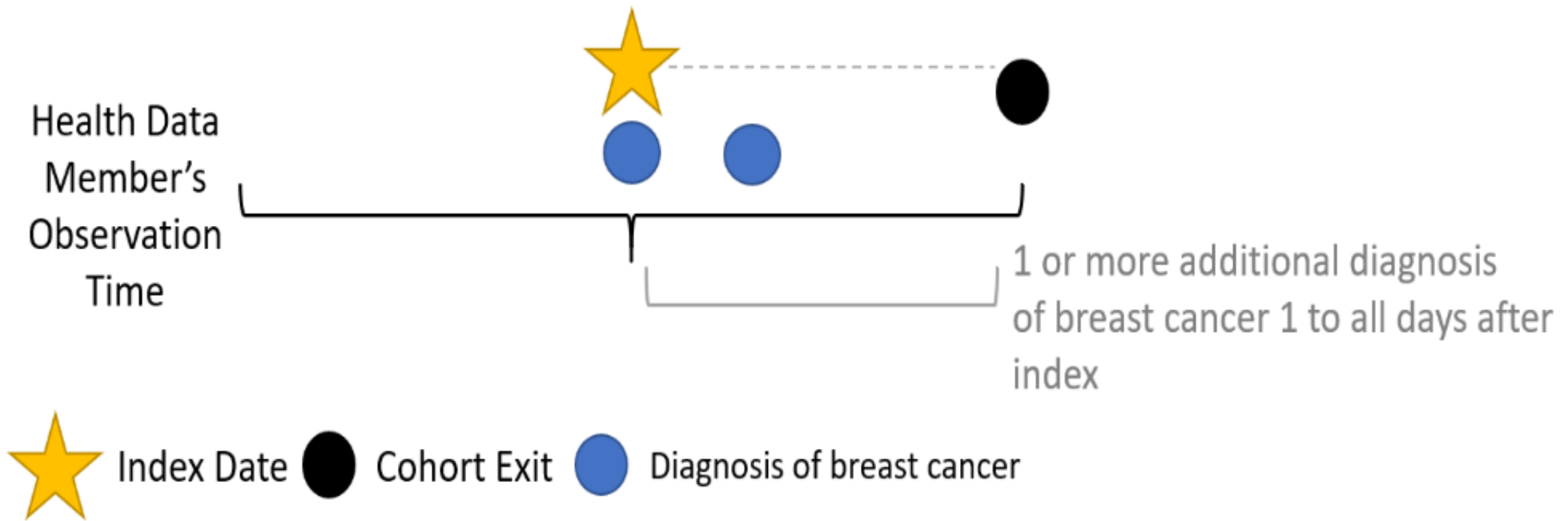
 Index Date  Cohort Exit  Procedure for screening mammography




Study Protocol

- Outcome Cohort

- Target cohort
- Outcome cohort
- Time At Risk



Build your cohorts



ATLAS

- Home
- Data Sources
- Search
- Concept Sets
- Cohort Definitions**
- Characterizations
- Cohort Pathways
- Incidence Rates
- Profiles
- Estimation
- Prediction
- Jobs
- Configuration
- Feedback

Apache 2.0
open source software

Cohort #1771537

[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms (1)

Definition **?** Concept Sets Generation Reporting Export Messages **5**

enter a cohort definition description here

Cohort Entry Events **?**

Events having any of the following criteria:

+ Add Initial Event ▾

a procedure occurrence of Screening mammography ▾ + Add attribute... ▾

✗ with age Between ▾ 40 and 74

✗ with a gender of: **✗ FEMALE** Add Import

with continuous observation of at least 1095 ▾ days before and 0 ▾ days after event index date

Limit initial events to: all events ▾ per person.

Restrict initial events

Inclusion Criteria **?**

New inclusion criteria

1. No mammography in prior 700 days
2. No neoplasm disease prior or up to 90 days following
3. No history of breast cancer all days prior and 90 days after
4. No prior hormone antagonists
21603829



Put your code in github

StudyProtocols/finalWoo at mast

github.com/OHDSI/StudyProtocols/tree/master/finalWoo

Why GitHub? Enterprise Explore Marketplace Pricing Search Sign in Sign up

OHDSI / StudyProtocols

Watch 31 Star 21 Fork 37

Code Issues 7 Pull requests 1 Projects 0 Security Insights

Branch: master StudyProtocols / finalWoo / Create new file Find file History

jreps moved document for woo to document folder Latest commit 452fa4a on Aug 30

..		
R	Adding WoO 2019 study	3 months ago
documents	moved document for woo to document folder	2 months ago
extras	Adding WoO 2019 study	3 months ago
inst	Adding WoO 2019 study	3 months ago
man	Adding WoO 2019 study	3 months ago
vignettes	Adding WoO 2019 study	3 months ago
.Rbuildignore	Adding WoO 2019 study	3 months ago
.Rprofile	Adding WoO 2019 study	3 months ago
.gitignore	Adding WoO 2019 study	3 months ago
DESCRIPTION	Adding WoO 2019 study	3 months ago



Ask organizations to run your study

Reps, Jenna [JRDGB] <jreps@ITS.JNJ.com> | [Maura Beaton](#); [Blacketer, Clair \[JRDUS\]](#); [Noemie Elhadad](#); [Anupama E. Gururaj](#); [Hardin, Jill \[JRDUS\]](#); [Hester, Laura \[JRDUS\]](#); [Jiang, Xinzhuo](#); [Shinyoung Ju](#); [Kristin Kostka](#); [Chun Yee Lau](#); + 12

Re: WoO at the symposium

You replied to this message on 8/8/2019 5:31 PM.

Our study is now live on github!

Study Atlas Design: <http://www.ohdsi.org/web/atlas/#/prediction/78>

Study package: <https://github.com/OHDSI/StudyProtocols/tree/master/finalWoo>

To run the package:

- 1) Open an Rstudio session
- 2) Make sure to install PatientLevelPrediction see <https://github.com/OHDSI/PatientLevelPrediction>
- 3) Install the package:

```
install.packages("devtools")
# install the network package
devtools::install_github("OHDSI/StudyProtocolSandbox/finalWoo")
```
- 4) Run the code in codeToRun.R found here: <https://github.com/OHDSI/StudyProtocols/blob/master/finalWoo/extras/CodeToRun.R>

Then you will have a compressed folder with the results (sensitive data removed) that can be shared back to me (jreps@its.jnj.com). If you have any questions, send me an email.

I'm currently running this on 3 claims datasets and 1 EHR we have at JNJ. If you have lots of data it can take a day to run.

Best wishes,
Jenna



Gather results from the sites

← → ↻ Not secure | data.ohdsi.org/WoO2019/ ☆ [Extensions] [M]

Multiple PLP Viewer ☰

- Summary
- Performance
- Model
- Log
- Help

Filters

Development Database
All

Validation Database
All

Target Cohort
All

Outcome Cohort
All

Time-at-risk start:
All

Time-at-risk end:
All

Model:
All

Results | Model Settings | Population Settings | Covariate Settings

Show 10 entries Search:

Analysis	Dev	Val	T	O	Model	TAR start	TAR end	AUC	AUPRC	T Size	O Count	O Incidence (%)
Analysis_1001	CCAE	CCAE	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.61081	0.00672	412572	1767	0.42829
Analysis_1002	Optum claims	Optum claims	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.64097	0.01078	216239	1131	0.52303
Analysis_1003	MDCD	MDCD	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.6698	0.0125	22254	138	0.6201
Analysis_1004	Optum EHR	Optum EHR	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.64183	0.02834	819560	9221	1.12512
Analysis_1005	MDCR	MDCR	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.568	0.012	19459	186	0.956
Analysis_1006	CUMC	CUMC	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.5172	0.0047	8634	31	0.359
Analysis_1007	CDM	CDM	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.60976	0.00732	250000	879	0.3516
Analysis_1008	AmbEMR	AmbEMR	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.6647	0.01113	250000	1427	0.5708
Analysis_1009	LRDX	LRDX	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.64533	0.0115	250000	1395	0.558
Analysis_1010	STaRR	STaRR	[WoO 2019] T1 - first mammography in 2 years and no prior neoplasms	[WoO 2019] O: Breast cancer (2 occurrences)	Lasso Logistic Regression	1	1095	0.617	0.013	3345	27	0.807

Showing 1 to 10 of 22 entries Previous 1 2 3 Next

<http://data.ohdsi.org/WoO2019/>



Create the paper (from Atlas)

Patient-Level Prediction: *Predicting breast cancer 90 days to 3 years after a mammography*

Prepared on: 2019-08-08

Created by: ()

Table of Contents

1. List of Abbreviations	3
2. Responsible Parties.....	3
3. Executive Summary	3
4. Rational & Background.....	3
5. Objective	3
6. Methods.....	4
6.1. Study Design.....	4
6.2. Data Source(s).....	5
6.3. Study Populations.....	5
6.4. Statistical Analysis Method(s).....	6
6.5. Quality Control	7
6.6. Tools	8
7. Diagnostics.....	8
8. Data Analysis Plan	8
8.1. Algorithm Settings.....	8
8.2. Covariate Settings.....	8
8.3. Model Development & Evaluation	11
8.4. Analysis Execution Settings.....	11
9. Strengths & Limitations	11
10. Protection of Human Subjects.....	12
11. Plans for Disseminating & Communicating Study Results	12
12. Tables & Figures.....	13
12.1. Incidence Rate of Target & Outcome.....	13



Verify results and get IRB approval



Publish Paper!

