



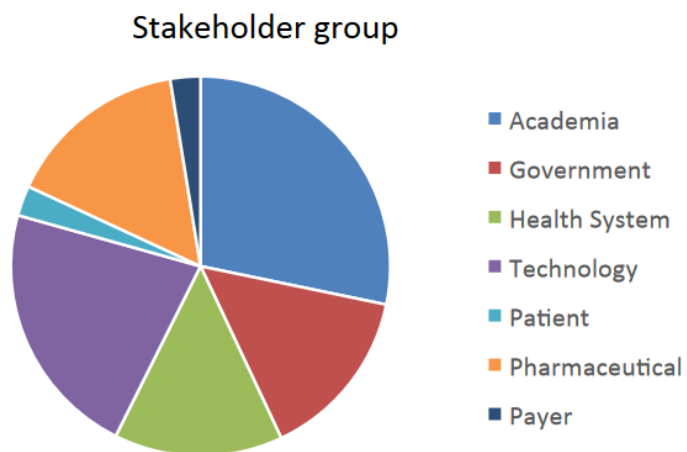
Oncology Module in the OMOP Standardized Vocabularies

21-Oct-2021



What OHDSI is:

- ✓ Open Source
- ✓ Community
- ✓ Data



OHDSI Collaborators:

- 2,100 users
- 25 workgroups
- 20 open-source applications

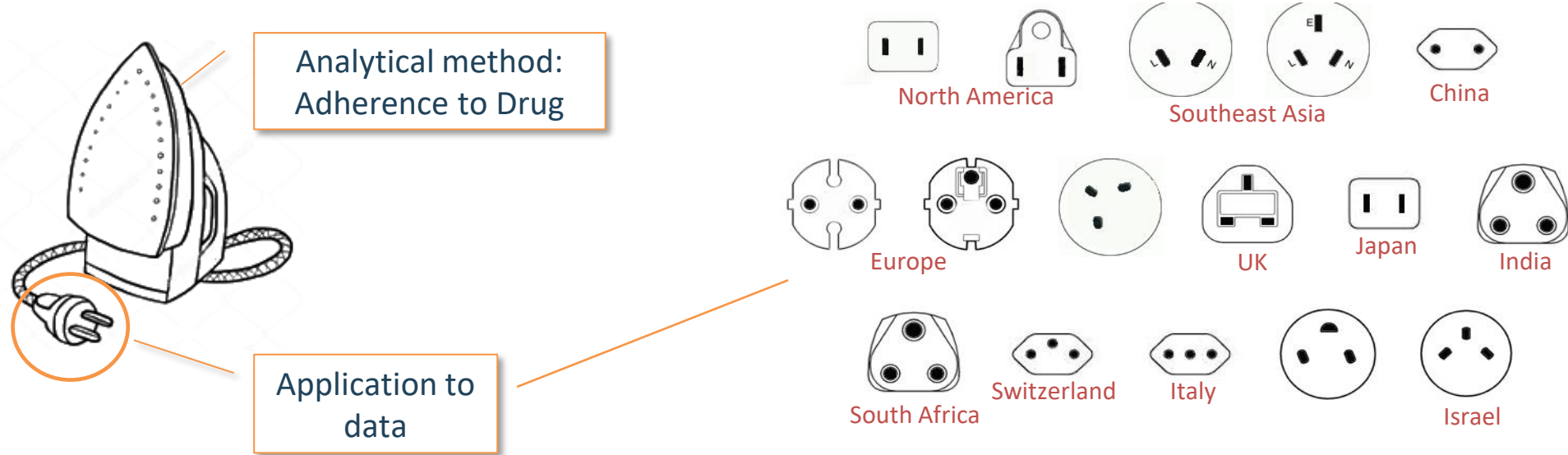
OHDSI Network:

- >160+ databases
- 23 countries
- 578M distinct patients
- 2.7B de-identified patient records

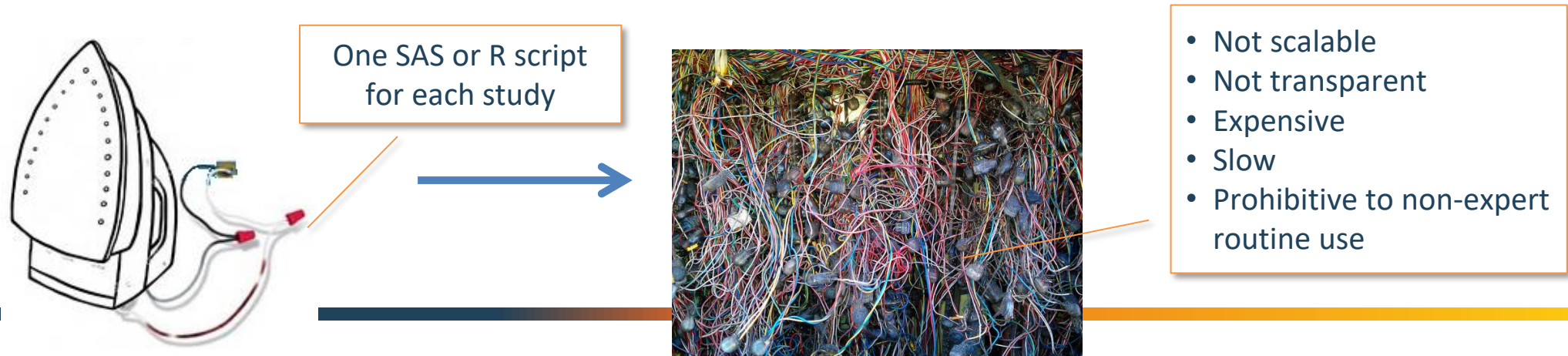


Current Approach: "One Study – One Script"

"What's the adherence to my drug in the data assets I own?"

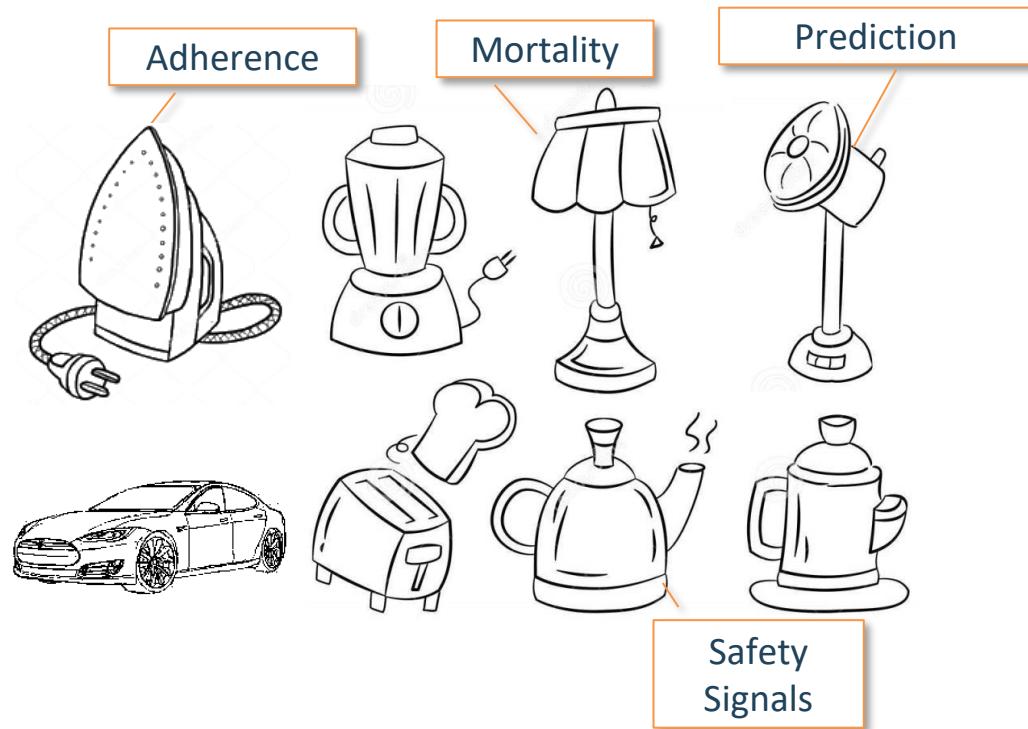


Current solution:

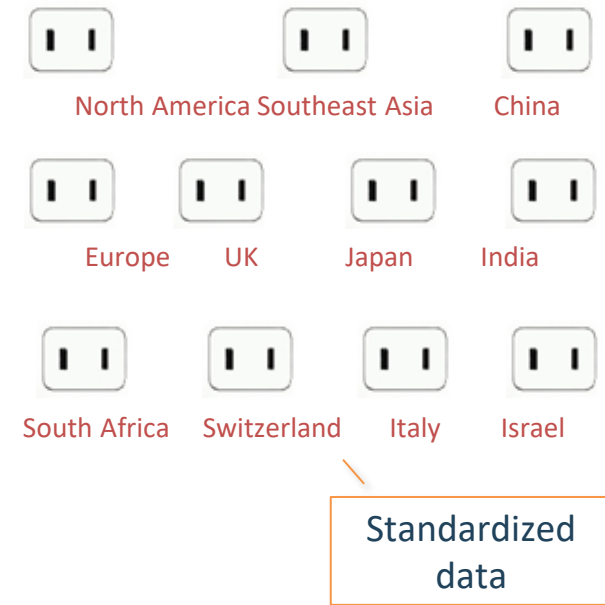




The OHDSI Approach



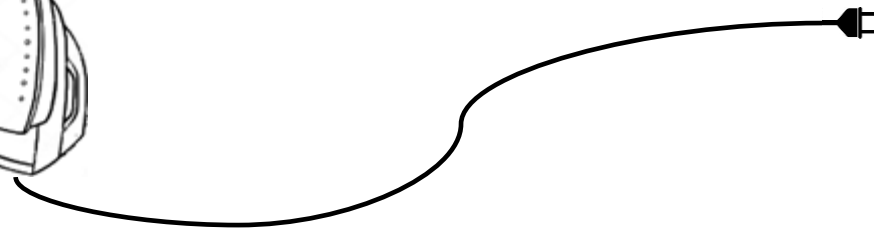
OHDSI Tools



OMOP CDM



Analytics can be remote



North America



Southeast Asia



China



Europe



UK



Japan



India



So Africa



Switzerland



Italy



Israel



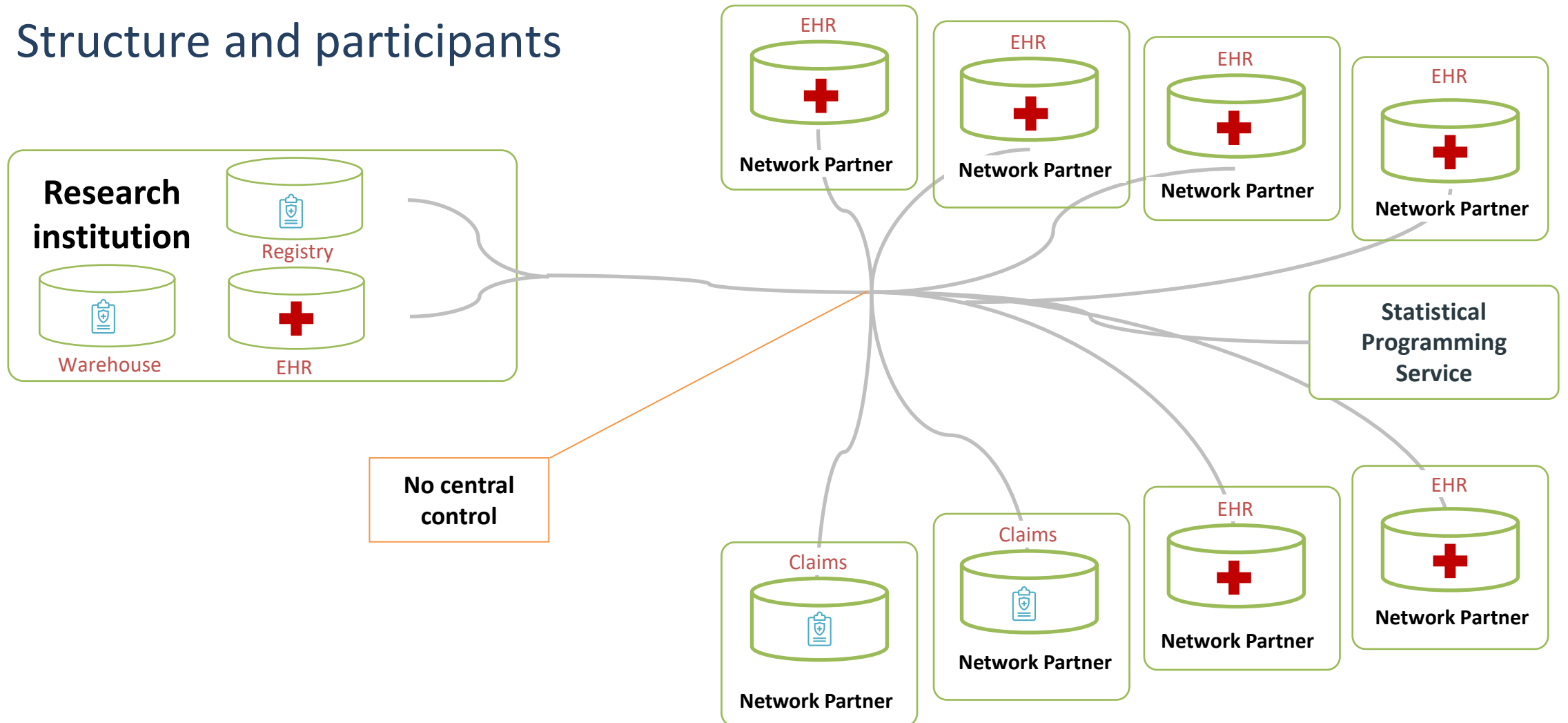
Analytics can be behind the firewall





OHDSI Research Network

Structure and participants



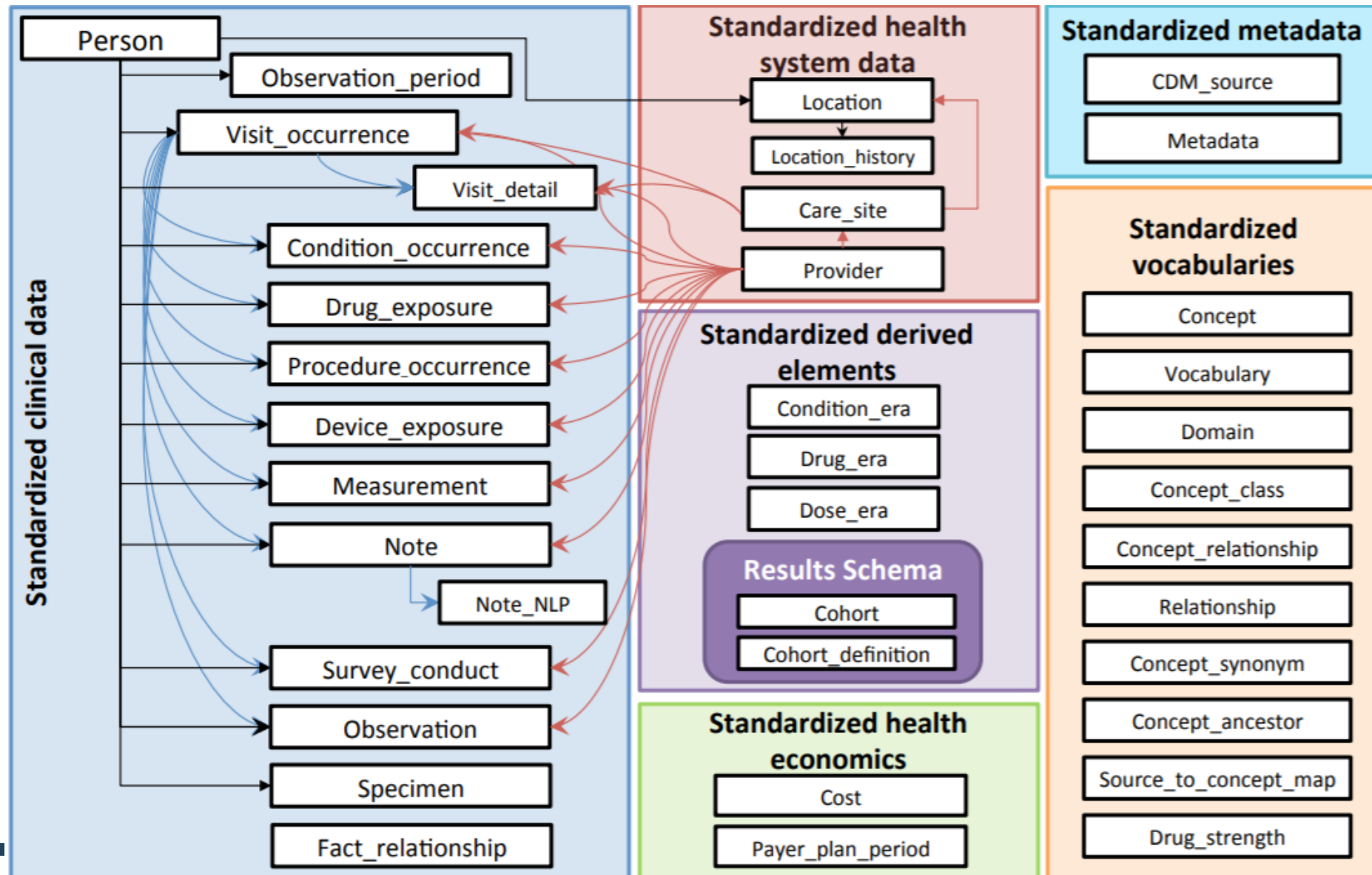


Opportunities for Standardization

Data structure	tables, fields, data types
Data content	vocabulary to codify clinical domains
Data semantics	conventions about meaning
Cohort definition	algorithms for identifying the set of patients who meet a collection of criteria for a given interval of time
Covariate construction	logic to define variables available for use in statistical analysis
Analysis	collection of decisions and procedures required to produce aggregate summary statistics from patient-level data
Results reporting	series of aggregate summary statistics presented in tabular and graphical form



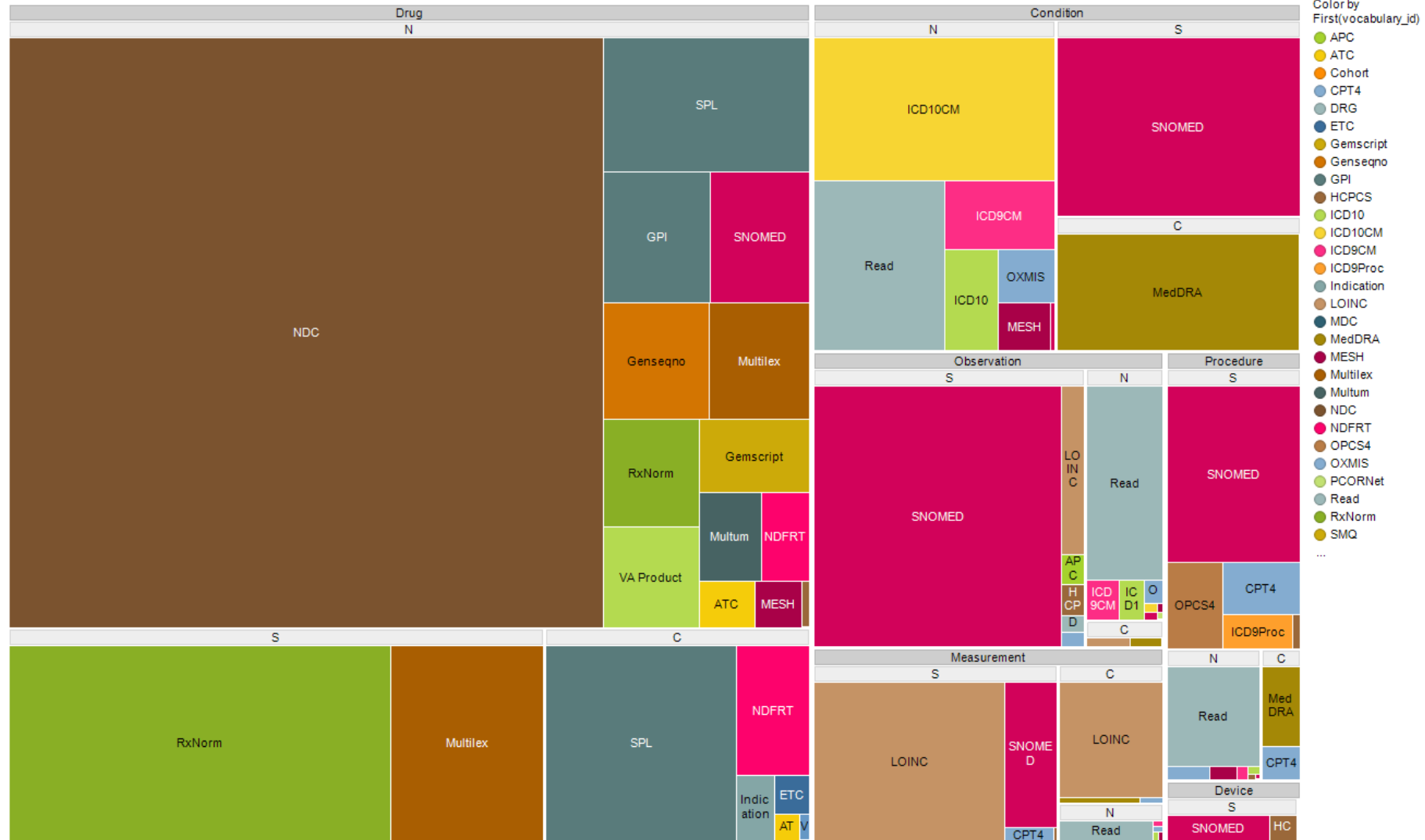
Standard Structure: OMOP Common Data Model



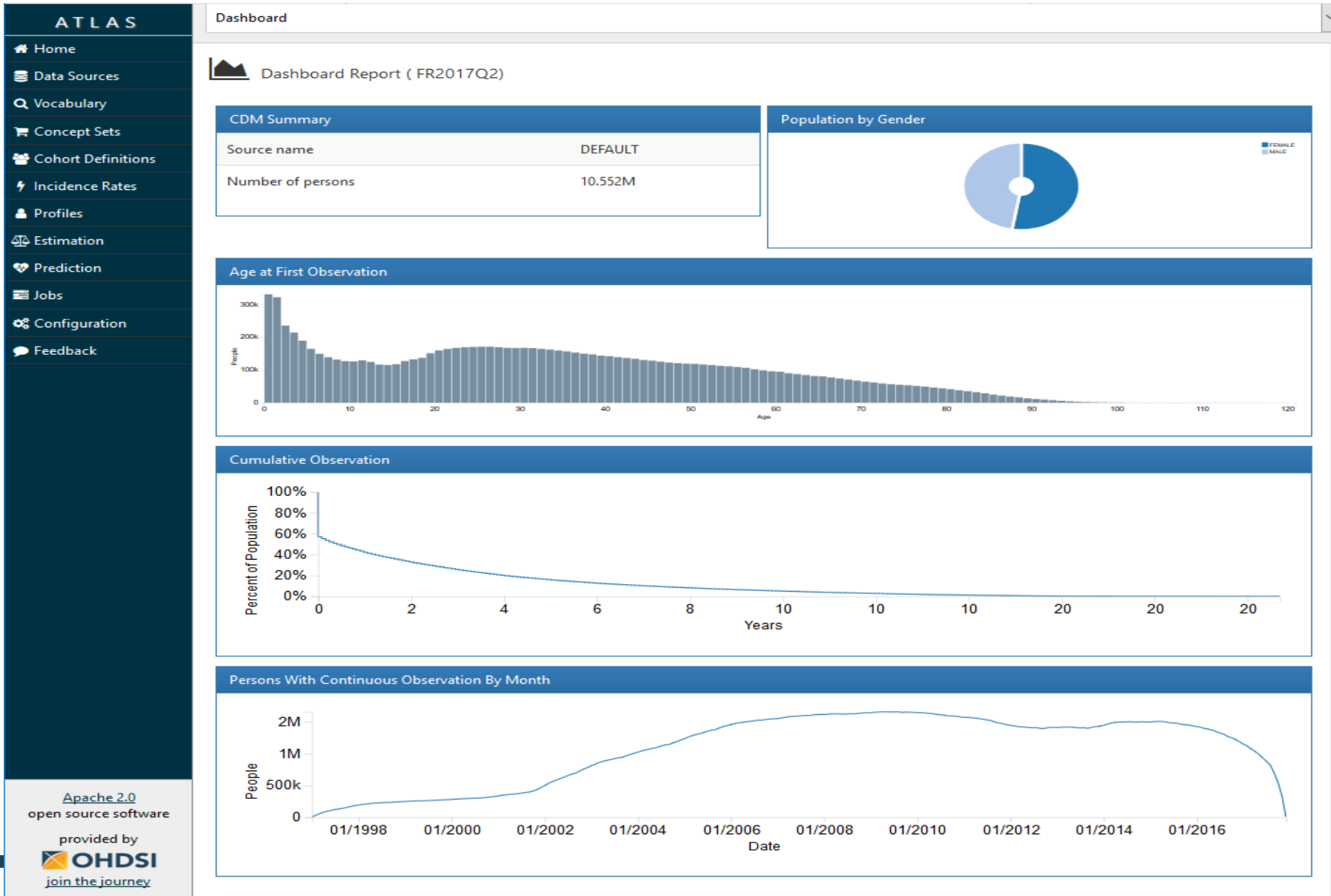


Standard Content & Semantics: OMOP Vocabularies

Breakdown of OHDSI concepts by domain, standard class, and vocabulary



Standard Analytic: ATLAS





Standard Cohort definition: ATLAS

ATLAS Cohort

OHDSI estimation tutorial: Graham replication: outcome cohort #1 - incident ischemic stroke

Definition Concept Sets Generation Reporting Explore Export

Cohort definition: A cohort is defined as the set of persons satisfying one or more inclusion criteria for a duration of time. One person may qualify for one cohort multiple times during non-overlapping time intervals. Cohorts are constructed in ATLAS by specifying cohort entry criteria and cohort exit criteria. Cohort entry criteria involve selecting one or more initial events, which determine the start date cohort entry, and optionally specifying additional inclusion criteria which filter to the qualifying events. Cohort exit criteria are applied to each cohort entry record to determine the end date when the person's episode no longer qualifies for the cohort.

Ischemic stroke, as defined in Graham et al, Circulation, 2015; <http://www.ncbi.nlm.nih.gov/pubmed/25359164>

All Cohort Entry Criteria Cohort Exit Criteria

Initial event cohort: Events are recorded time-stamped observations for the persons, such as drug exposures, conditions, procedures, measurements and visits. All events have a start date and end date though some events may have a start date and end date with the same value (such as procedures or measurements). The event index date is set to be equal to the event start date.

People having any of the following: Add Initial Event...

People having any of the following: Add Initial Event...

a condition occurrence of Ischemic stroke Add

Condition Type is: Inpatient detail - primary Inpatient header - primary Primary Condition

with a Visit occurrence of: Emergency Room Visit Inpatient Visit Add Import

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

Limit initial events to: earliest event per person.



Standard Methods: HADES

Estimation methods

Cohort Method
New-user cohort studies using large-scale regression for propensity and outcome models

Self-Controlled Case Series
Self-Controlled Case Series analysis using few or many predictors, includes splines for age and seasonality.

Self-Controlled Cohort
A self-controlled cohort design, where time preceding exposure is used as control.

Patient Level Prediction
Build and evaluate predictive models for user-specified outcomes, using a wide array of machine learning algorithms

Case-control
Case-control studies, matching controls on age, gender, provider, and visit date. Allows nesting of the study in another cohort.

Case-crossover
Case-crossover design including the options to adjust for time-trends in exposures (so-called case-time-control).

Method characterization

Empirical Calibration
Use negative control exposure-outcome pairs to profile and calibrate a particular analysis design.

Method Evaluation
Use real data and established reference sets as well as simulations injected in real data to evaluate the performance of methods.

Evidence Synthesis
Combining study diagnostics and results across multiple sites.

Supporting packages

Database Connector
Connect directly to a wide range of database platforms, including SQL Server, Oracle, and PostgreSQL.

Sql Render
Generate SQL on the fly for the various SQL dialects.

Cyclops
Highly efficient implementation of regularized logistic, Poisson and Cox regression.

ParallelLogger
Support for parallel computation with logging to console, disk, or e-mail

Feature Extraction
Automatically extract large sets of features for user-specified cohorts using data in the CDM.



**Why is oncology any different than
the rest of medicine?**

Problem 1: Cancer is a rare disease

HIGHLIGHTS OF PRESCRIBING INFORMATION

These highlights do not include all the information needed to use TABRECTA safely and effectively. See full prescribing information for TABRECTA.

TABRECTA™ (capmatinib) tablets, for oral use
Initial U.S. Approval: 2020

INDICATIONS AND USAGE

TABRECTA is a kinase inhibitor indicated for the treatment of adult patients with metastatic non-small cell lung cancer (NSCLC) whose tumors have a mutation that leads to mesenchymal-epithelial transition (MET) exon 14 skipping as detected by an FDA-approved test.

This indication is approved under accelerated approval based on overall response rate and duration of response. Continued approval for this indication may be contingent upon verification and description of clinical benefit in confirmatory trial(s). (1)

DOSAGE AND ADMINISTRATION

- Select patients for treatment with TABRECTA based on presence of a mutation that leads to MET exon 14 skipping. (2.1)
- Recommended dosage: 400 mg orally twice daily with or without food. (2.2)

DOSAGE FORMS AND STRENGTHS

Tablets: 150 mg and 200 mg (3)

CONTRAINDICATIONS

None. (4)

WARNINGS AND PRECAUTIONS

- **Interstitial Lung Disease (ILD)/Pneumonitis:** Monitor for new or worsening pulmonary symptoms indicative of ILD/pneumonitis. Permanently discontinue TABRECTA in patients with ILD/pneumonitis. (2.3, 5.1)
- **Hepatotoxicity:** Monitor liver function tests. Withhold, dose reduce, or permanently discontinue TABRECTA based on severity. (2.3, 5.2)
- **Risk of Photosensitivity:** May cause photosensitivity reactions. Advise patients to limit direct ultraviolet exposure. (5.3)
- **Embryo-Fetal Toxicity:** Can cause fetal harm. Advise patients of the potential risk to a fetus and to use effective contraception. (5.4, 8.1, 8.3)

ADVERSE REACTIONS

The most common adverse reactions (≥ 20%) are peripheral edema, nausea, fatigue, vomiting, dyspnea, and decreased appetite. (6)

To report SUSPECTED
Pharmaceuticals Cor
1088 or www.fda.gov

Strong and Moderate C

US
Lactation: Advise not

See 17 for PATIENT
approved patient label

HIGHLIGHTS OF PRESCRIBING INFORMATION

These highlights do not include all the information needed to use TABRECTA safely and effectively. See full prescribing information for TABRECTA.

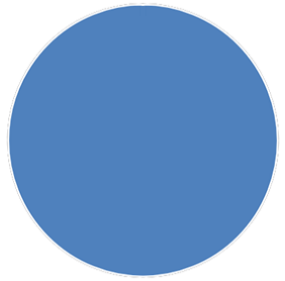
TABRECTA™ (capmatinib) tablets, for oral use
Initial U.S. Approval: 2020

INDICATIONS AND USAGE

TABRECTA is a kinase inhibitor indicated for the treatment of adult patients with metastatic non-small cell lung cancer (NSCLC) whose tumors have a mutation that leads to mesenchymal-epithelial transition (MET) exon 14 skipping as detected by an FDA-approved test.



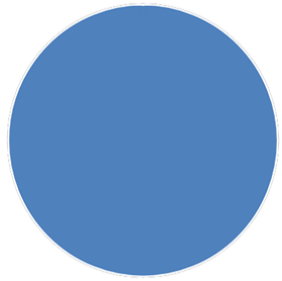
Lung Cancer: most frequent cancer and cause of death



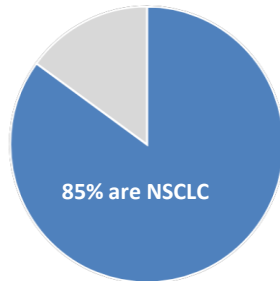
236,000
cases



But: Not all are Non-small cell



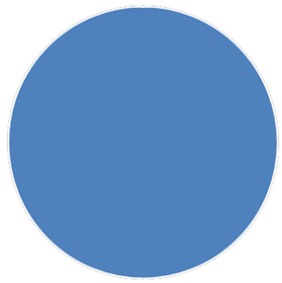
236,000
cases



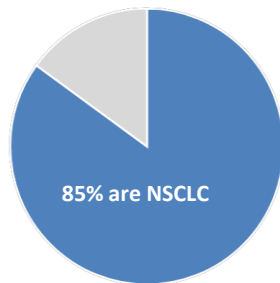
200,600
cases



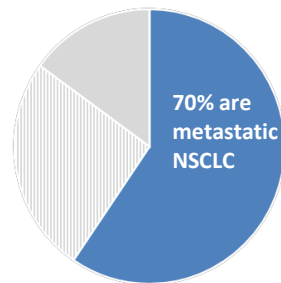
Not all are metastatic



236,000
cases



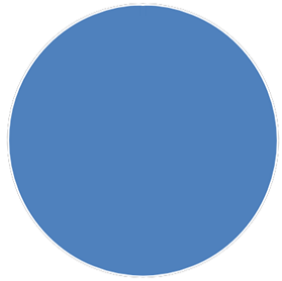
200,600
cases



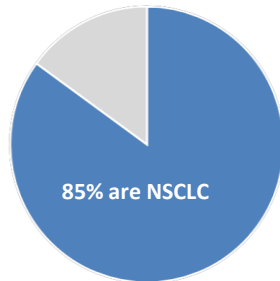
140,400
cases



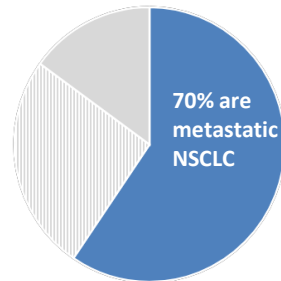
Not all are adenocarcinomas



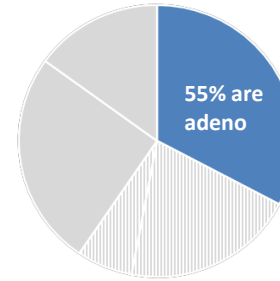
236,000
cases



200,600
cases



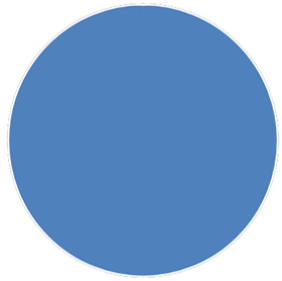
140,400
cases



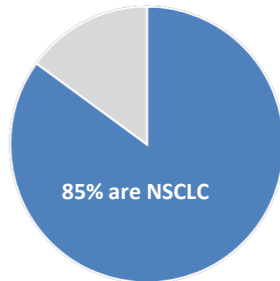
77,200
cases



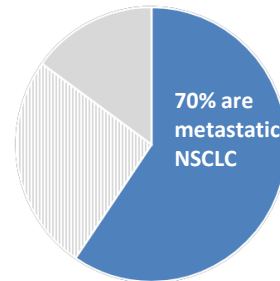
Few have Met ex14 mutation



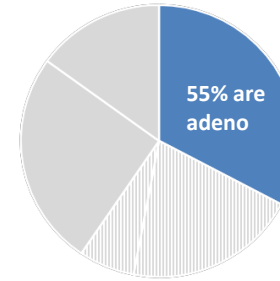
236,000
cases



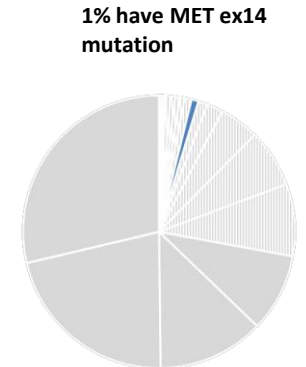
200,600
cases



140,400
cases



77,200
cases

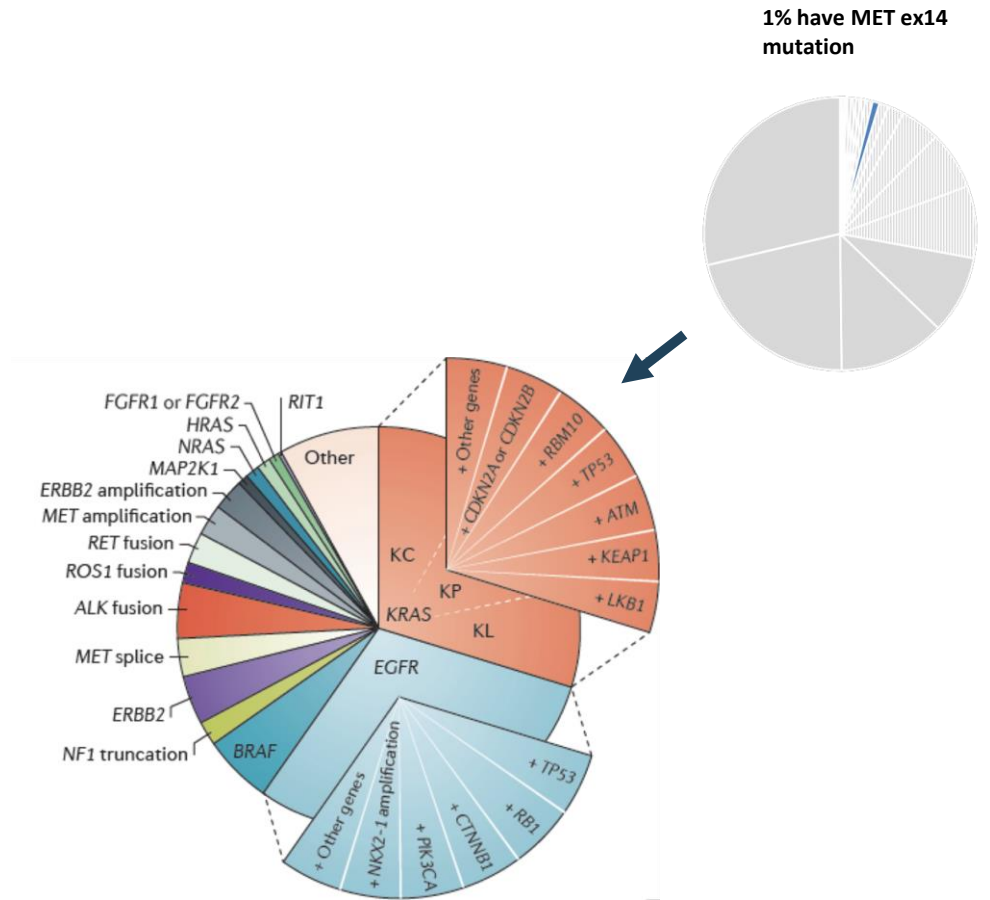


**1% have MET ex14
mutation**

7,700
cases



There are many such “rare diseases”





Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology
Lung	Anatomical site



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology
Lung	Anatomical site
Metastatic disease	Tumor attribute



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology
Lung	Anatomical site
Metastatic disease	Tumor attribute
MET exon 14 skipping	Genomic Variant



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology
Lung	Anatomical site
Metastatic disease	Tumor attribute
MET exon 14 skipping	Genomic Variant
First line treatment	Treatment Episode



Problem 2: Cancer needs more detail

“What is the Overall Survival or Progression-free Survival of patients with metastatic Non-small Cell Lung Cancer with confirmed MET exon 14 skipping who received oral capmatinib as first line?”

Concept	Category
Non-small Cell	Histology
Lung	Anatomical site
Metastatic disease	Tumor attribute
MET exon 14 skipping	Genomic Variant
First line treatment	Treatment Episode
Capmatinib	Regimen



Problem 3: Terminologies

There are no common or even good terminologies

Concept	Category	
Non-small Cell	Histology	ICDO, SNOMED
Lung	Anatomical site	ICDO, SNOMED
Metastatic disease	Tumor attribute	
MET exon 14 skipping	Genomic Variant	CiVIC, OncoKB, ClinVar, NCIt, CAP, LOINC, SNOMED
First line treatment	Treatment Episode	
Capmatinib	Regimen	RxNorm, HemOnc



Lymph Nodes (CAP and NAACCR)

Vocabularies are badly curated

- Lymph Node Status:
 - Lymph Node Status
 - Nodal Status: Para-Aortic, Mediastinal, Pelvic, Femoral Inguinal and Distant (Mediastinal, Scalene)
 - LN Status: Femoral-Inguinal, Para-Aortic, Pelvic
 - Clinical Status of Lymph Node Mets
 - Clinical Status of Lymph Nodes
- Lymph node size:
 - LN Size
 - Size of Lymph Nodes
- Mets at DX-Distant LN
- LN Distant: Mediastinal, Scalene
- Adenopathy
- Nodal Stations Involved
- Laterality
 - Laterality
 - LN Laterality
 - Regional Lymph Node – Laterality



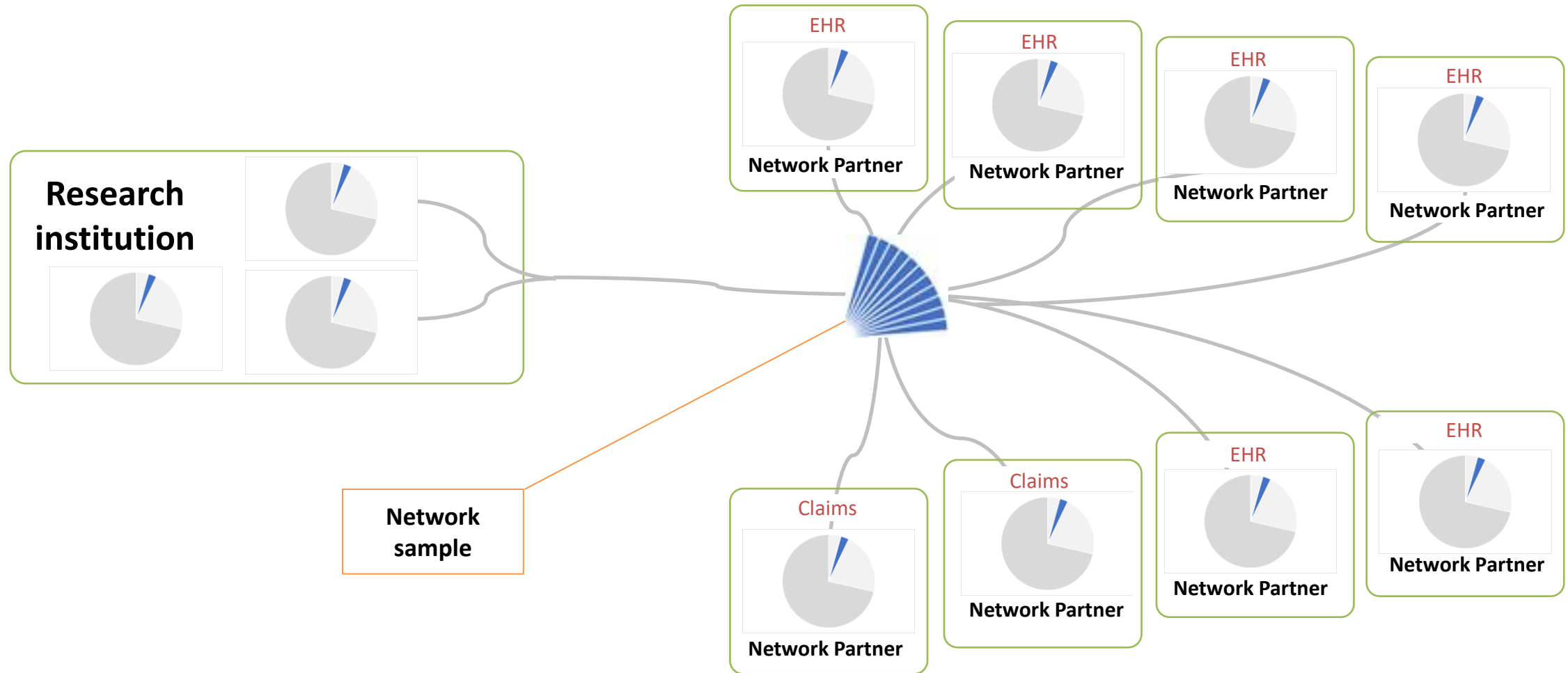
The OHDSI Oncology Working Group Has Worked toward a Solution

- Oncology Network
- Oncology Module for the OMOP CDM
- Example studies



OHDSI Oncology Network

Data from many institutions can be analyzed together





OMOP CDM: Oncology Module

Solves all problems of oncology research

1 Cancer Disease Model

Cancer Diagnosis: Base Diagnosis + Diagnostic Modifiers
(One-to-many connection between them)

2 Cancer Treatment Model

Composite Level (Treatment Episodes) or Individual Level (standard OMOP)

3 Cancer Episode Model

Continuous periods of disease or treatment with distinct clinical meaning
Composed of multiple events
Essential for conducting cancer research



1

Oncology Module: Cancer Disease Model

Cancer Disease Model

Cancer Diagnosis: Base Diagnosis + Diagnostic Modifiers

Base Diagnosis in **CONDITION**:

Histology + Anatomic Site

Diagnostic Modifiers in **MEASUREMENT**:

Topography

Staging/Grading

Histological pattern

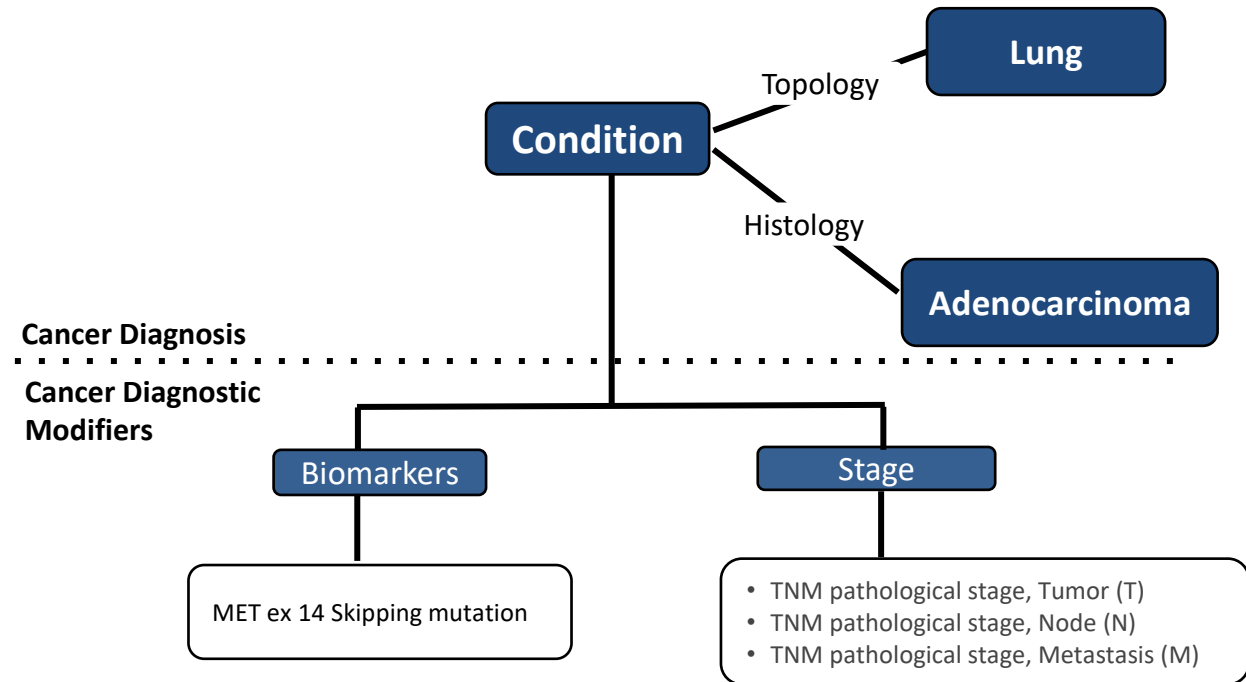
Genomic Variant

Margin

Metastasis

Dimension

Extension/Invasion





Standard OMOP Variant Concepts

1. Genes

- Protein: **SNP**, expression
- Transcript: **SNP**, amplification
- Genomic: **SNP**, mutation, rearrangement, gene fusion, aneu/polysomy

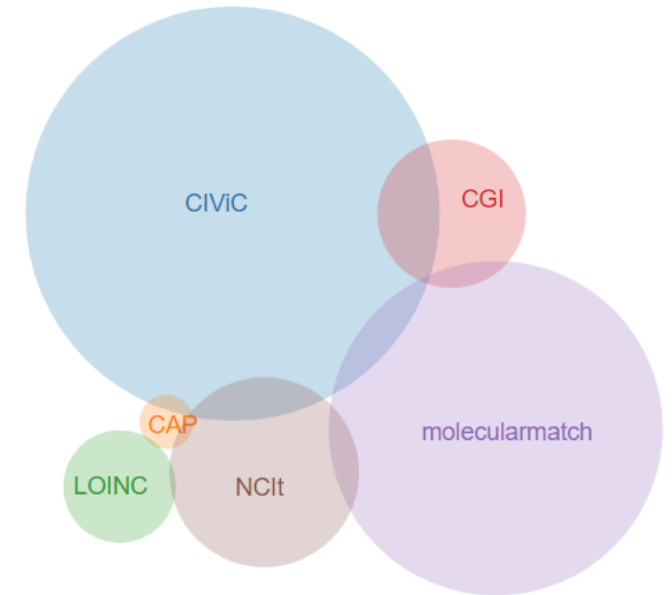
2. Karyotypes

- Chromosomes, bands: tr, del, mono/polysomy, +, -

3. Microsatellite stability

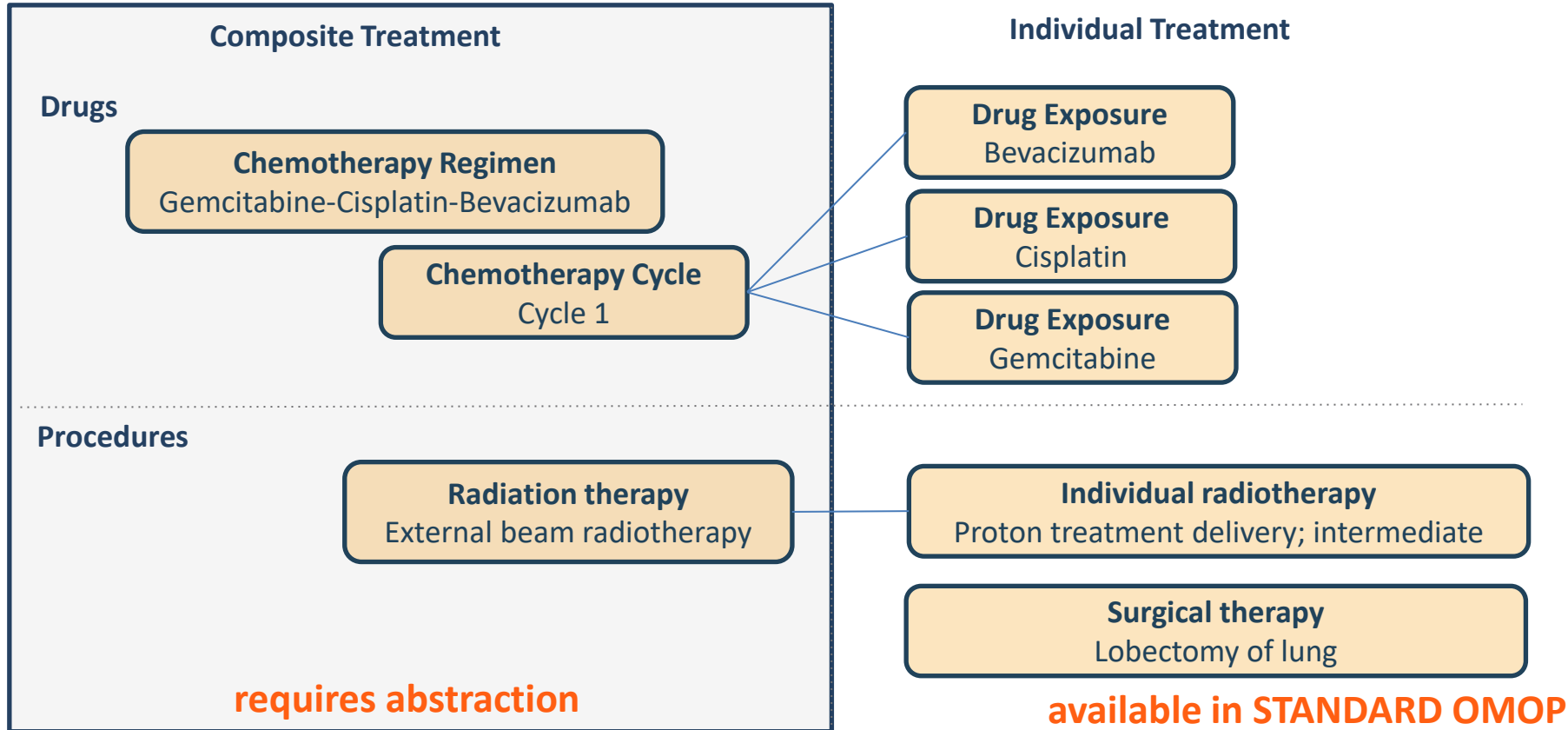
4. Histone Trimethylation

5. Epigenetics





2 Oncology Module: Cancer Treatment Model





3 Oncology Module: Cancer Episode Model

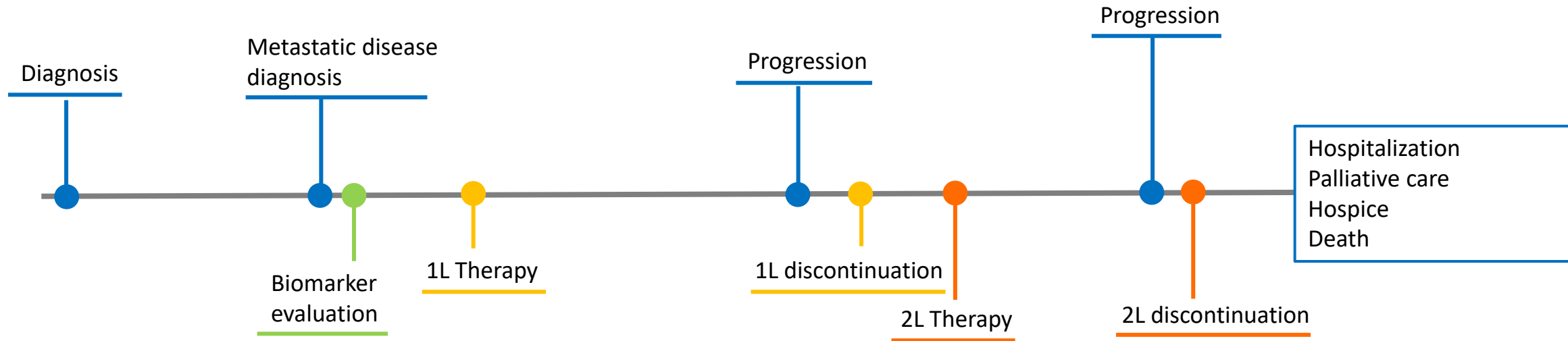
Episodes

- ✓ Continuous periods of disease or treatment with distinct clinical meaning
- ✓ Composed of multiple events
- ✓ Essential for conducting cancer research
- ✓ Obtained directly from source data (e.g., registries) or algorithmically derived

- Parent Episode:
 - **Overarching disease episode:** Covers the entire cancer duration
- Children Episodes:
 - **Disease dynamic** (remission, stable, progression)
 - **Disease extent** (confined, invasive, metastatic)

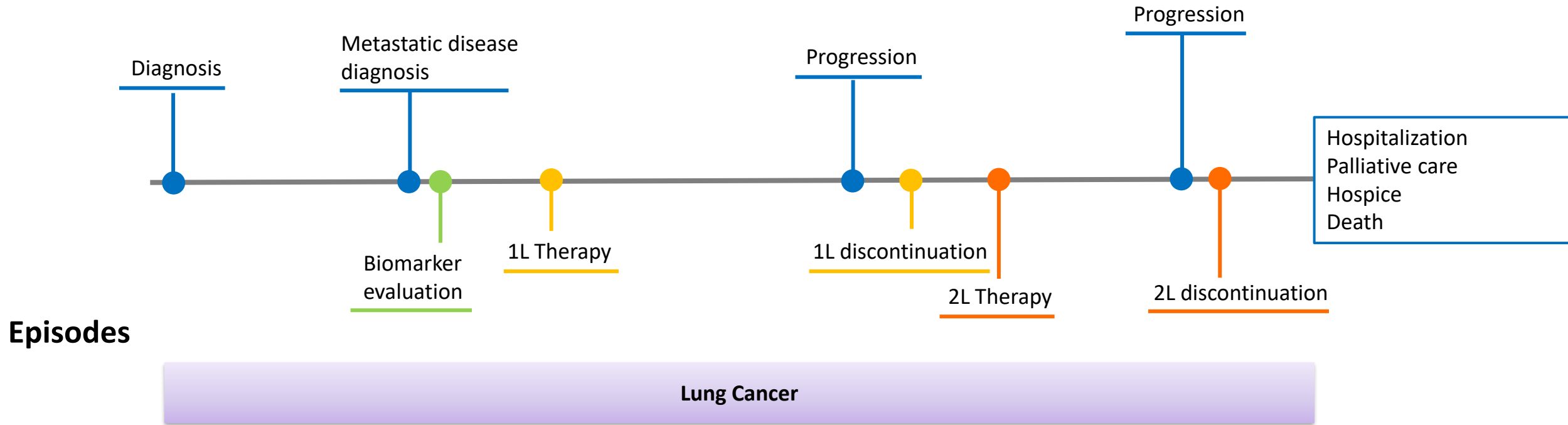


Cancer Episode Model: Schematic Patient Journey



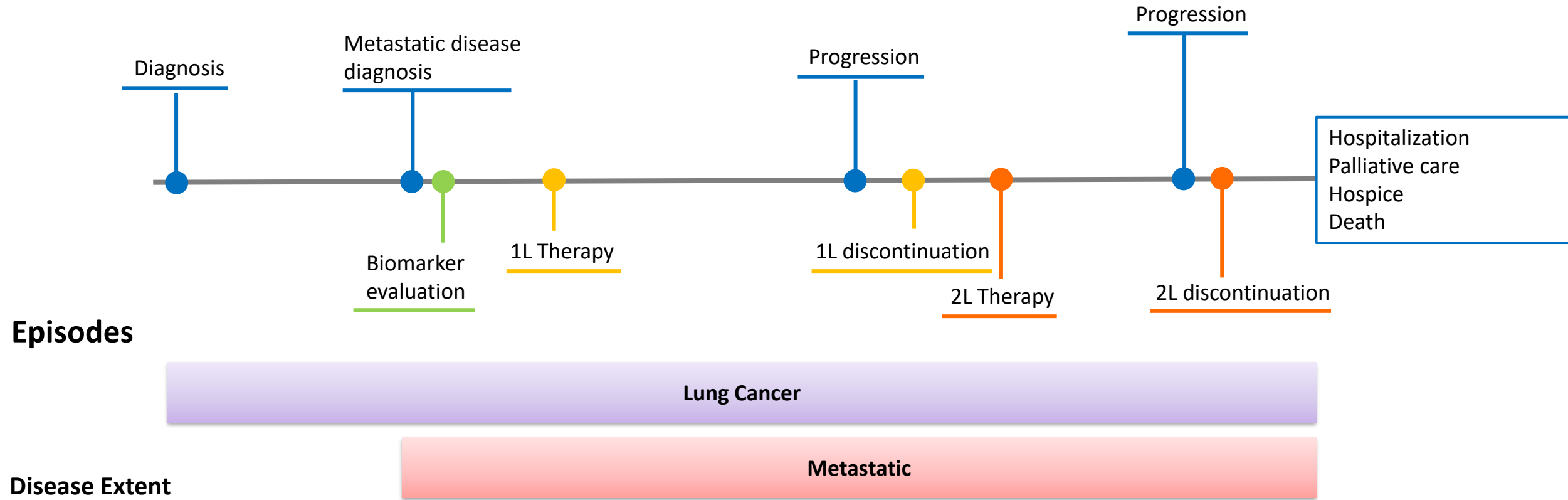


Cancer Episode Model: Schematic Patient Journey



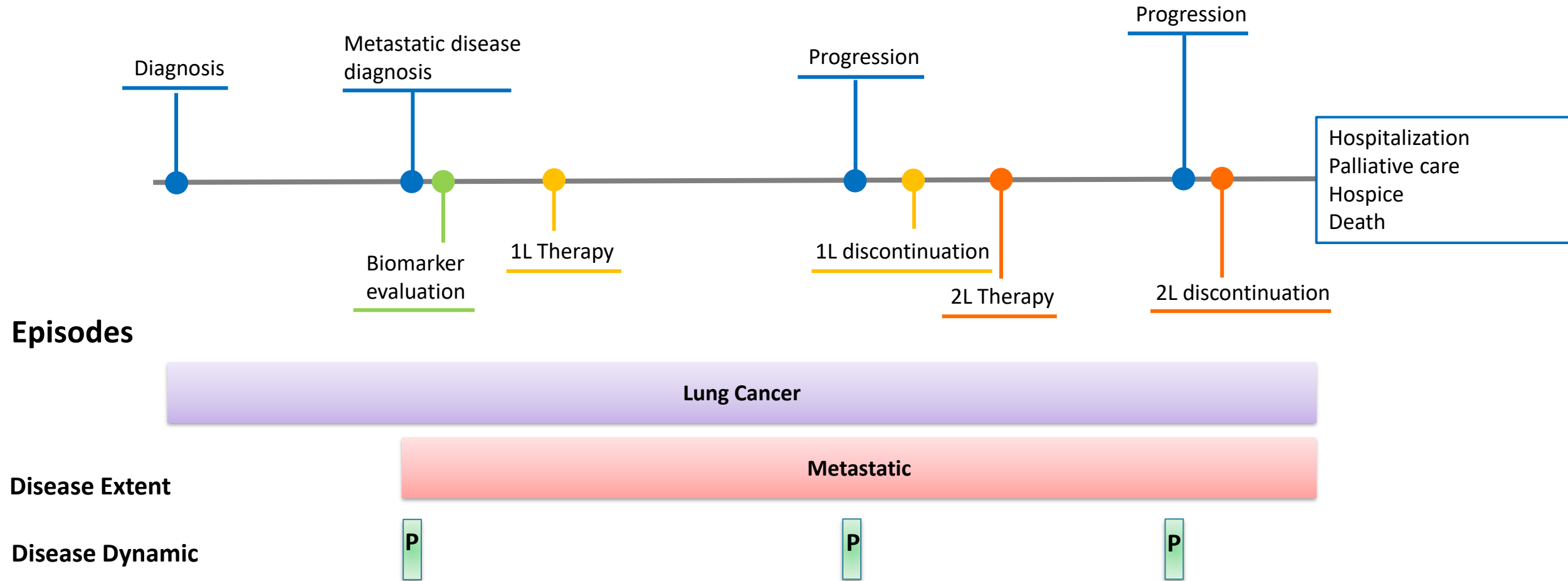


Cancer Episode Model: Schematic Patient Journey



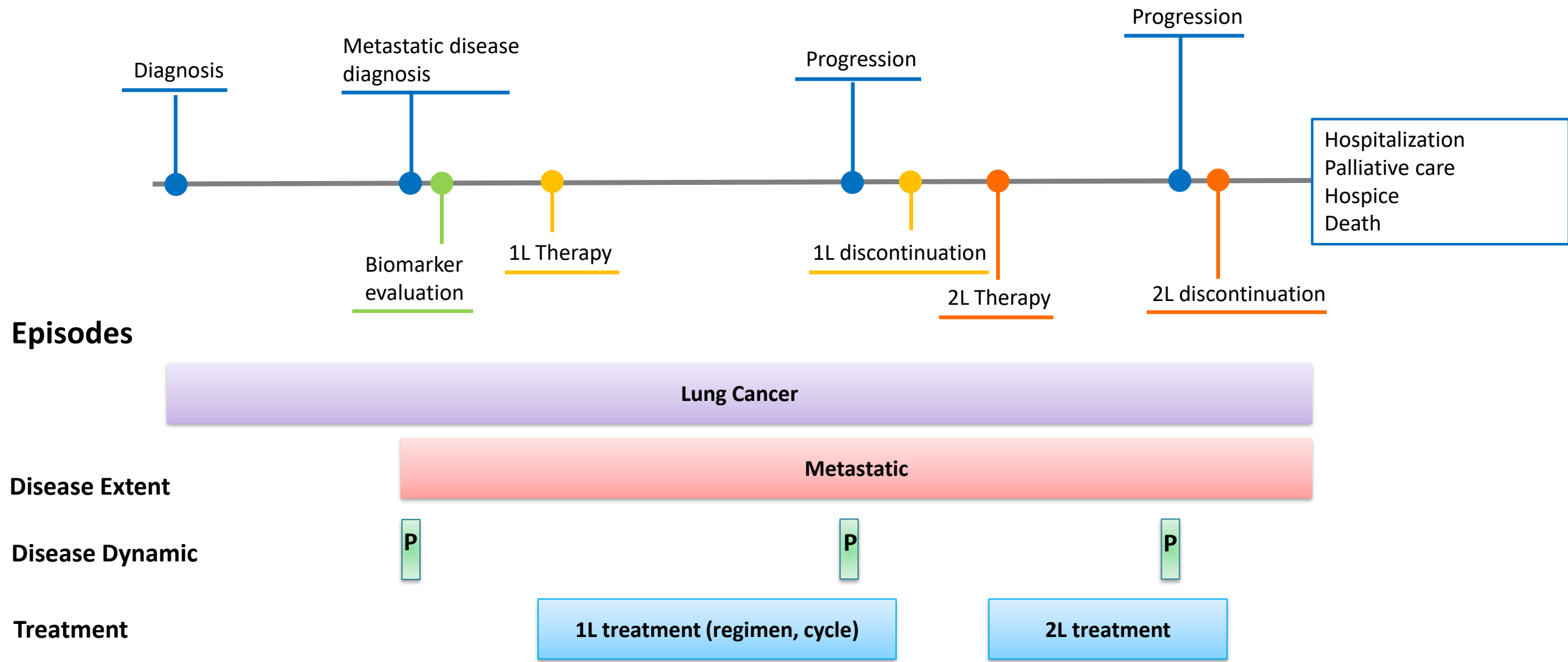


Cancer Episode Model: Schematic Patient Journey



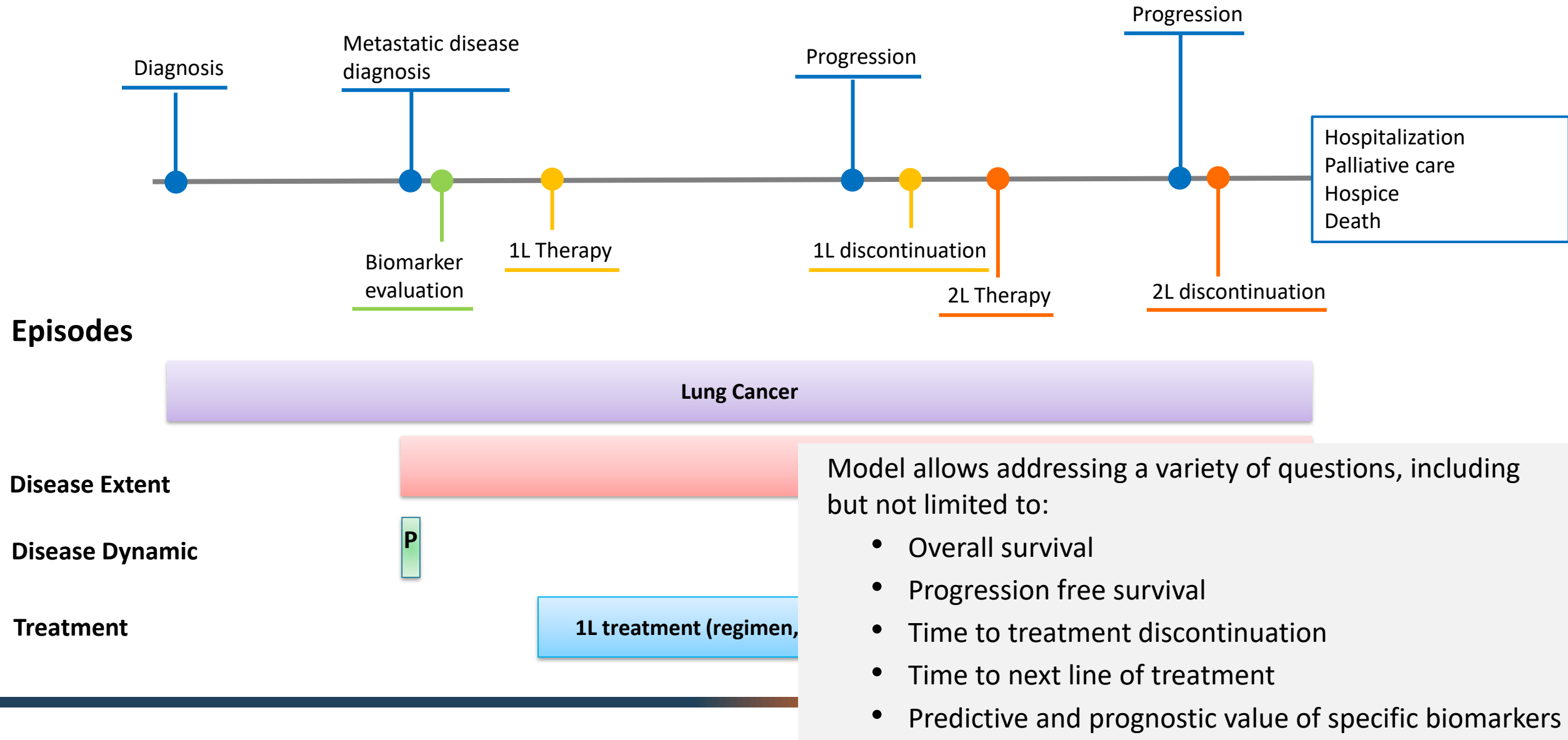


Cancer Episode Model: Schematic Patient Journey





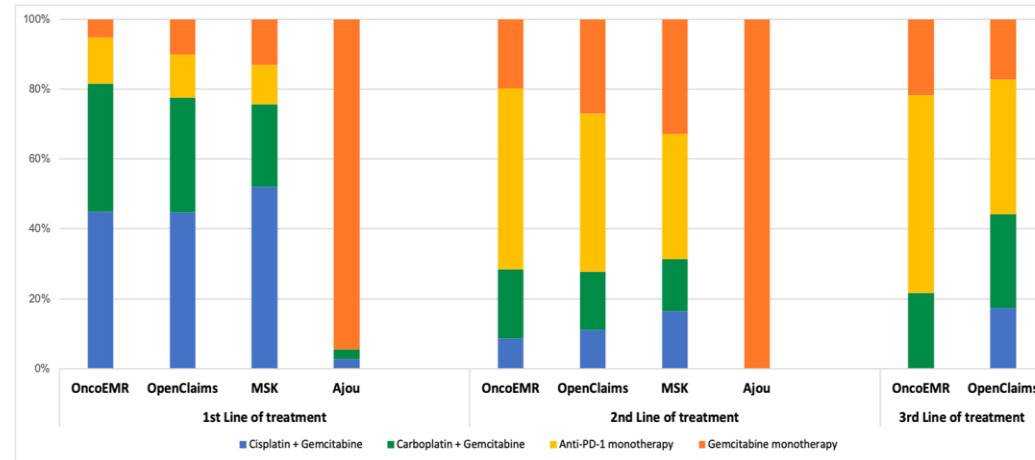
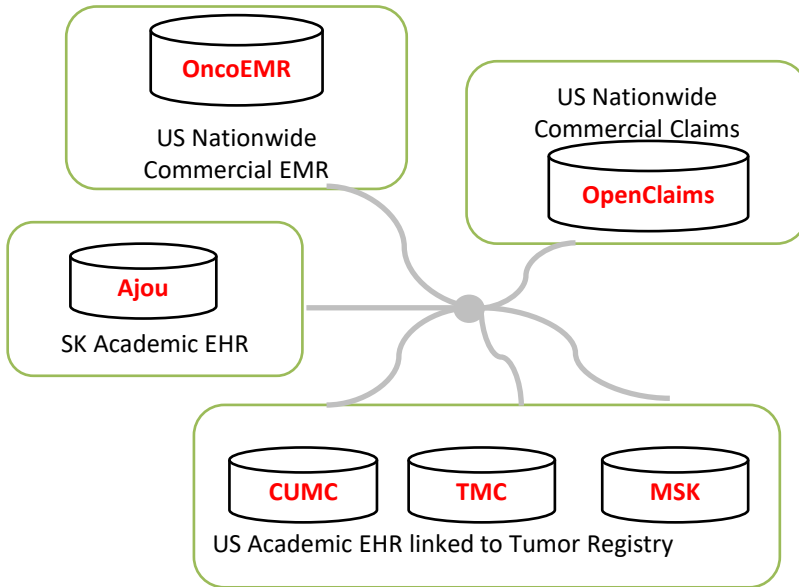
Cancer Episode Model: Schematic Patient Journey



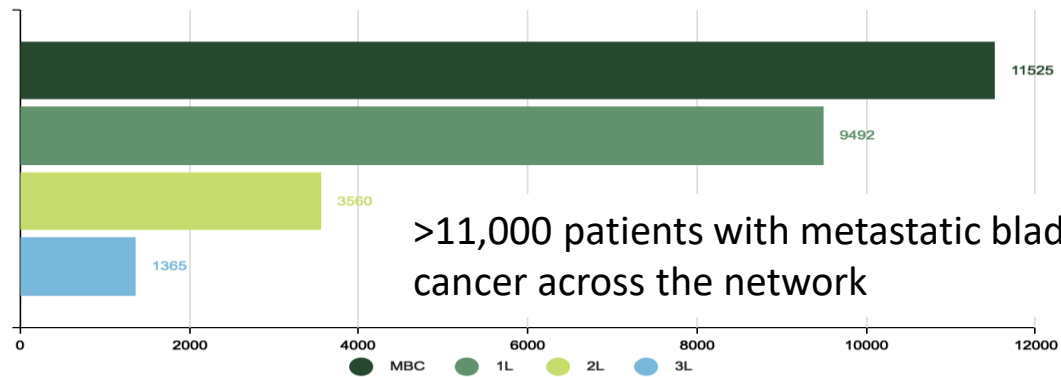


Network Study 1: Bladder Cancer

Lines of Treatment and Timing

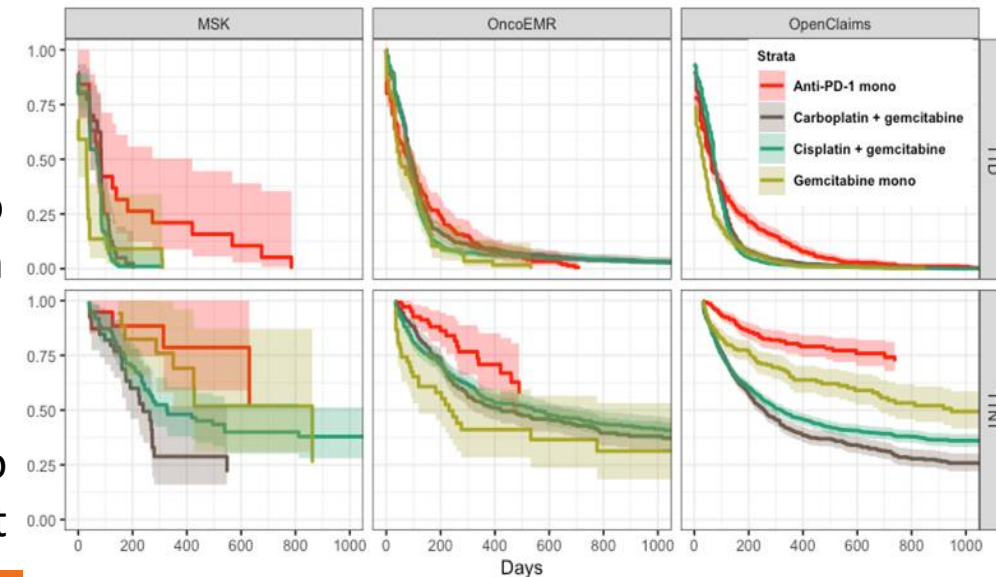


Distribution of the 1st, 2nd and 3rd lines of treatment



Time to discontinuation

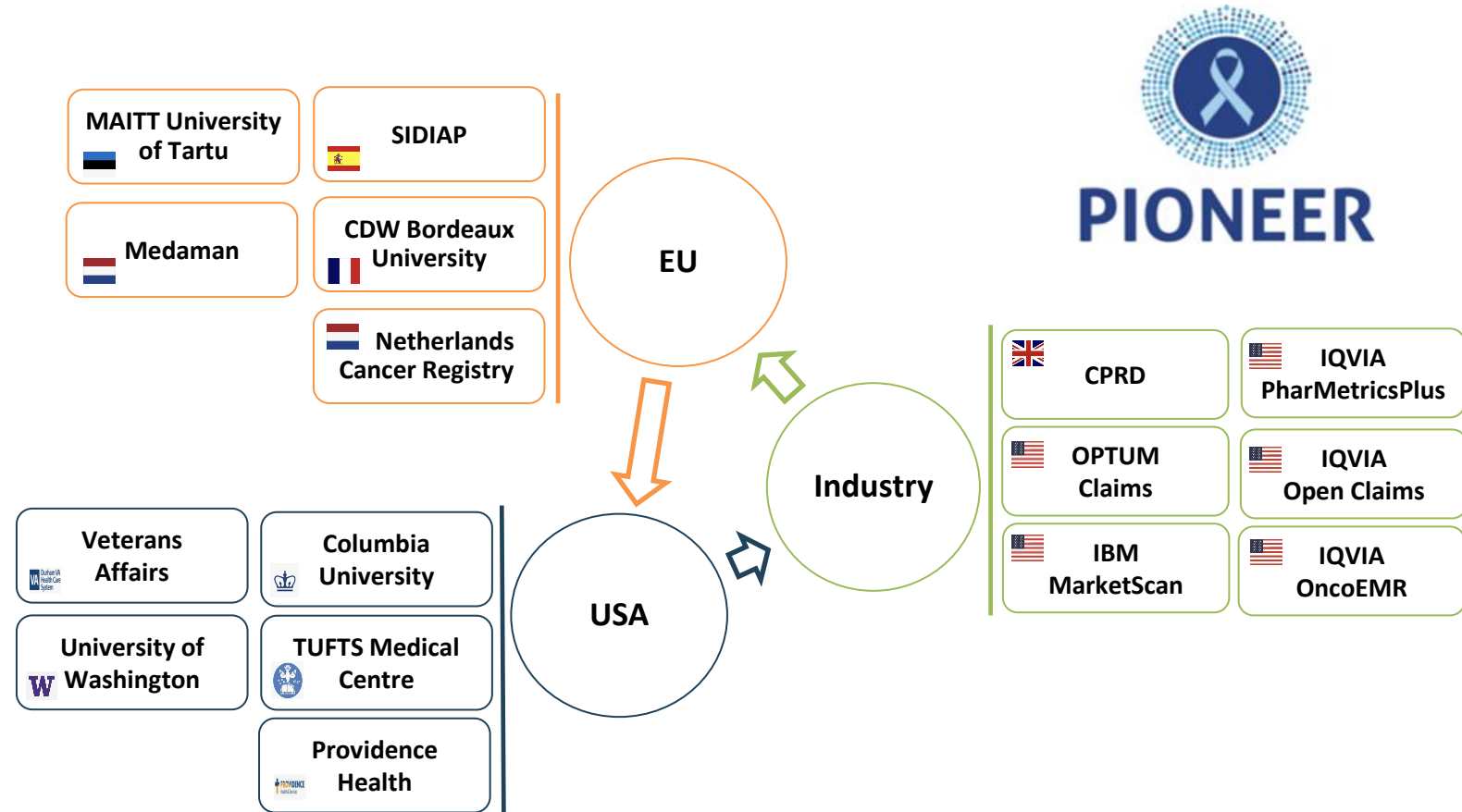
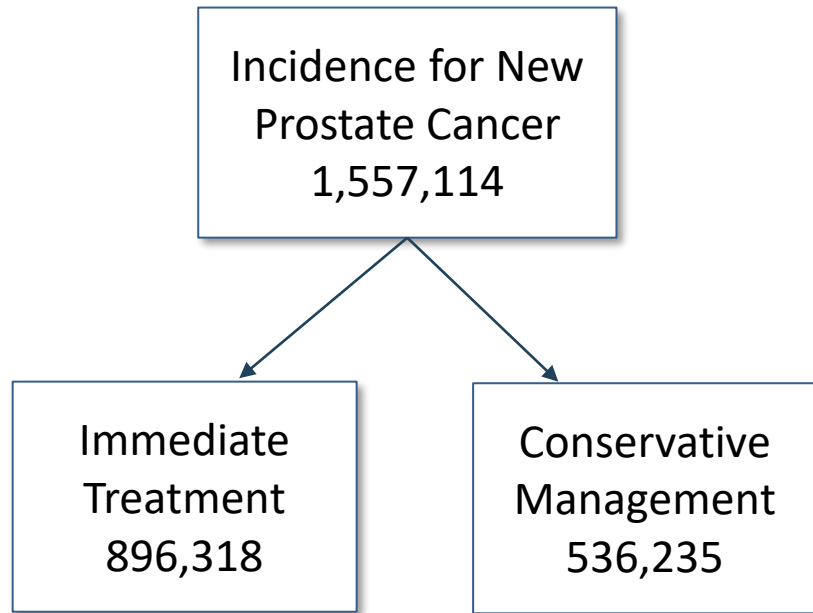
Time to next treatment





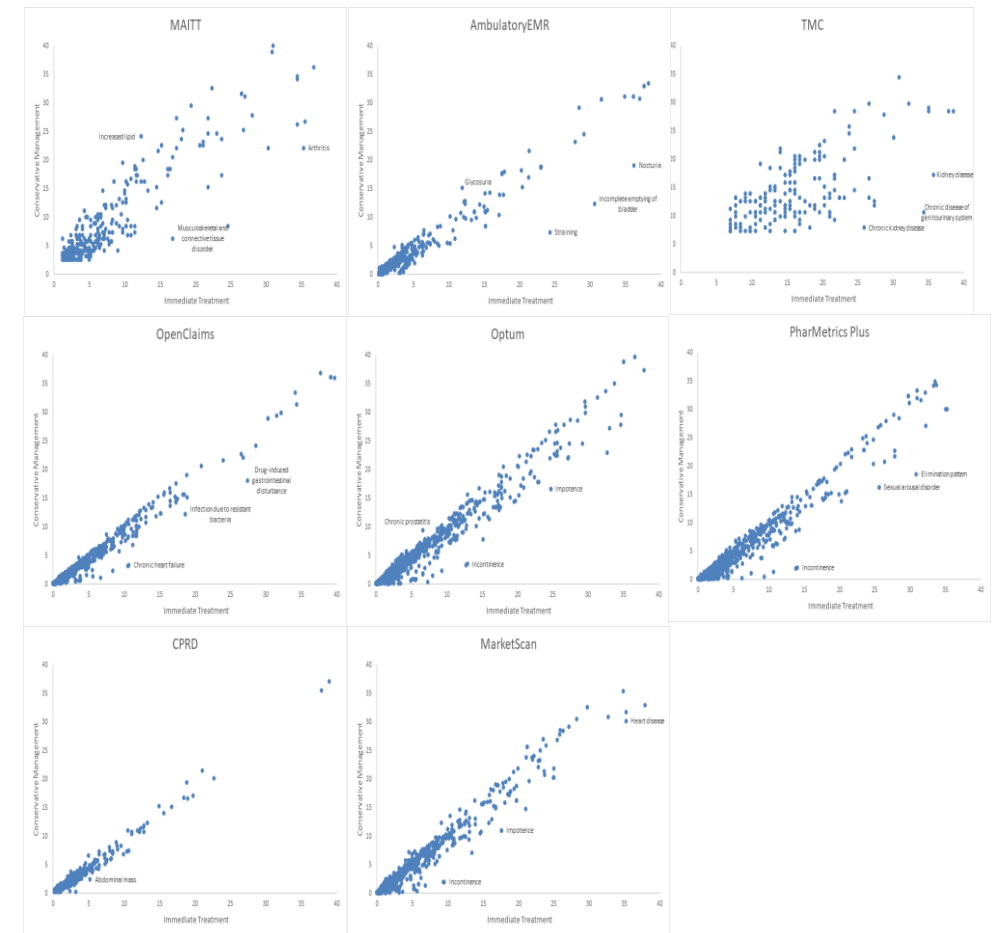
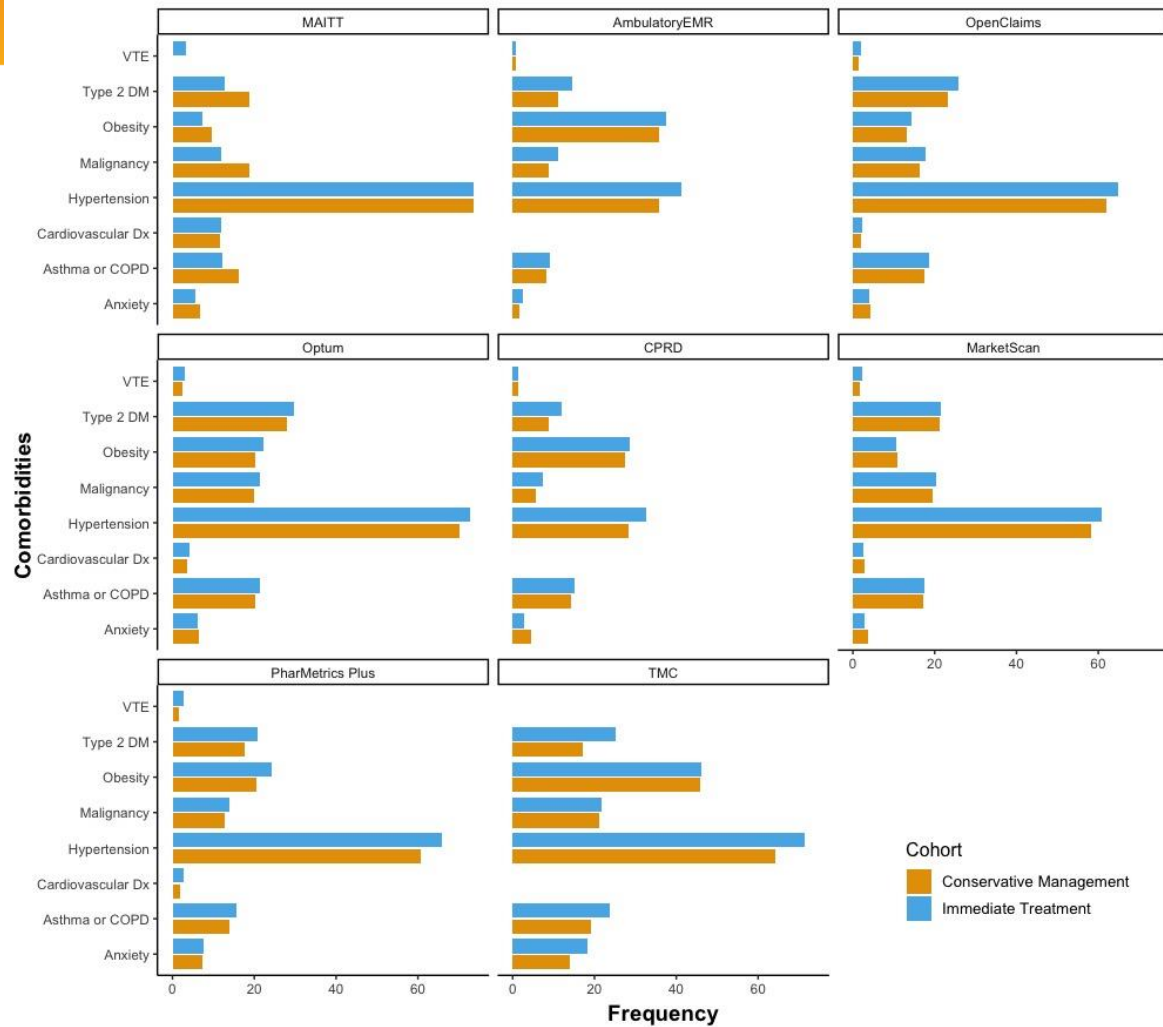
Network Study 2: Prostate Cancer

Characterization of Treatment Choice and Prediction of Symptomatic Disease





Baseline Characterization



Baseline conditions: slightly higher proportion of genitourinary conditions in immediate treatment group

Comorbidities: Hypertension: the most common condition in both cohorts



Network Study 3: Lung Cancer

Prognostic & Predictive Significance of Liver metastasis in Non-Small Cell



- Collaboration between NCI, UAB and OHDSI
 1. Characterization of NSCLC patients with and without liver metastasis
 2. Evaluating the prognostic and predictive significance of liver metastasis in patients with NSCLC



Summary

Oncology module enables observational cancer study in a network setting

Details and foundation to answer series of questions in cancer:

1. Cancer burden (e.g., relative 5-year survival, incidence and mortality)
2. Characteristics of patients
3. Treatment pattern, dosing, regimens and sequence
4. Treatment administration
5. Benchmarking
6. Real world outcomes