

Single Day Event

Examples of utilizing real-world data in different drug development stages

Xiaoxuan Zhu, BeiGene
Pu Huang, BeiGene



Outlines

- What is RWx?
- Real-World Data (RWD) Characteristics
- The Opportunities of RWE in Drug Development
- Case Sharing
 - Case #1 1L HCC Treatment Pattern
 - Case #2 B-Cell Lymphoma Patient Recruitment
 - Case #3 CLL Cardiovascular AEs (CVAEs) Investigation
- Conclusion



What is RWx?

Real-World Evidence (RWE) ¹

is the clinical evidence regarding the usage and potential benefits or risks of a medical product derived from analysis of RWD.

Real-World Study (RWS)

is a research or study to generate, collect or utilize RWD to provide RWE, including but not limited to, randomized trials, including large simple trials, pragmatic trials, and observational studies (prospective and/or retrospective).



Real-World Data (RWD) ¹

are data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.

- Electronic Health Records (EHRs)
- Claims and Billing data
- Product and Disease Registries
- Patient-generated Data (mobile devices, wearable devices, social media)
- Literature
- Prior Clinical Trials
- Clinical Labs

Real-World Data (RWD) Characteristics

Databases (Type)		Therapeutic Areas	
Region	All	Oncology	
US	• Truven MarketScan (Claim); • HealthVerity (HV) (EMR and claim); • Patient Like Me (PLM) (Patient-reported data); • NHANES (National Health and Nutrition Examination Survey) (Survey)	• Symphony (Claim); • Flatiron (EMR); • CancerLinQ (EMR); • SEER (Registry)	
	UK	• CPRD (EMR)	• Simulacrum (Registry)
Pan-Europe	• EHR4CR Insite Platform (Europe hospital network) (EMR)		

Table. Example Sources of RWD

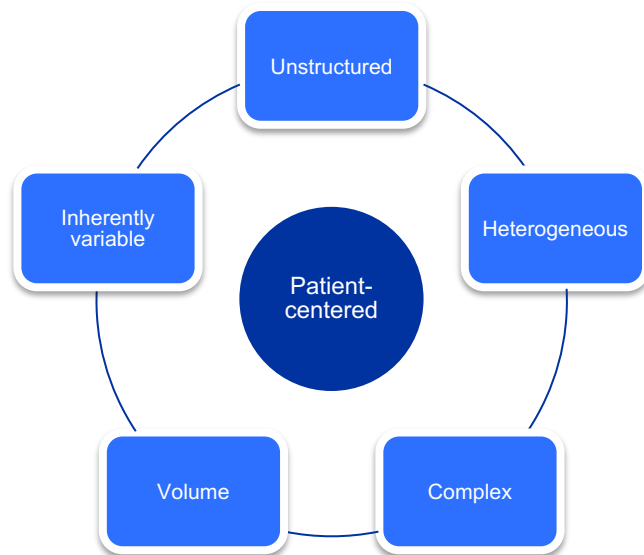
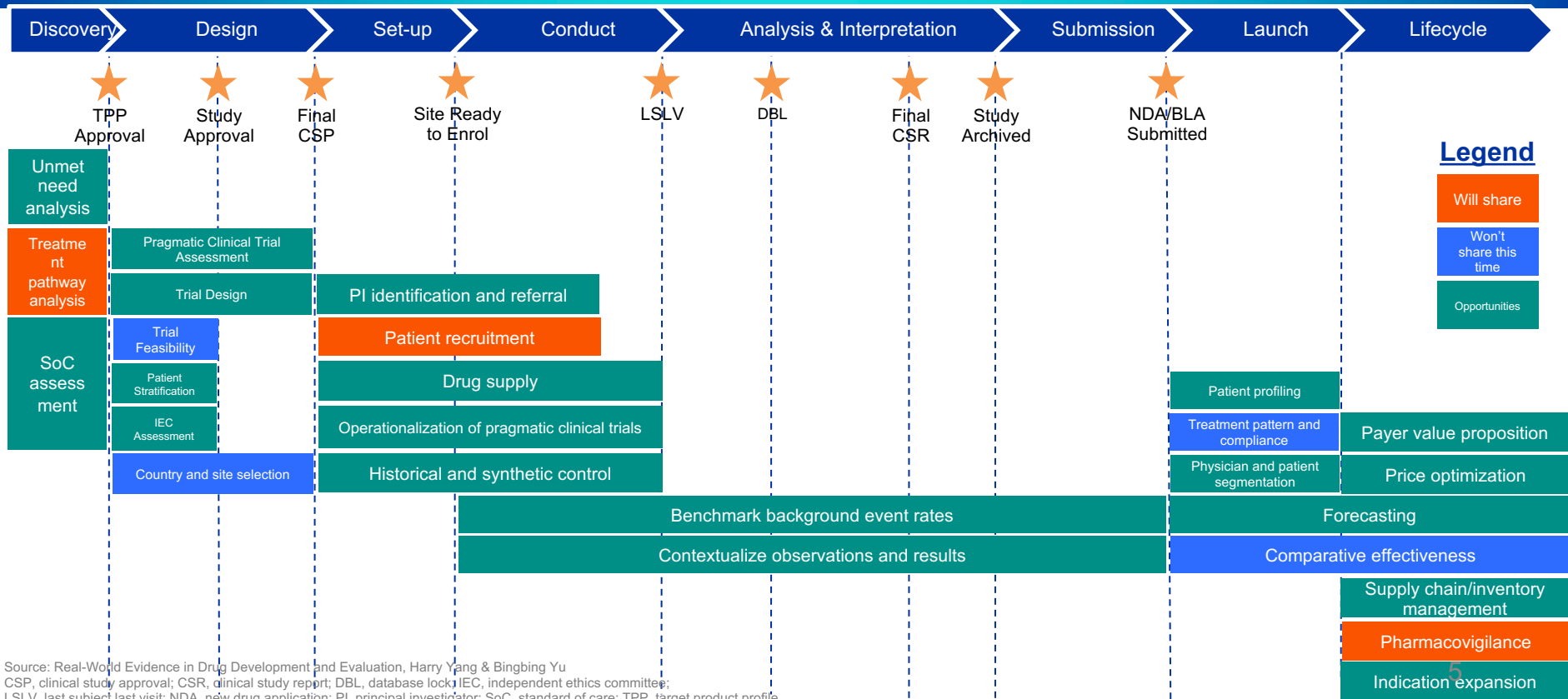


Figure. RWD Characteristics



Case Sharing

Case #1



Study
1L HCC Treatment Pattern



RWD
MarketScan (Claims, US)



Cooperation
Product Valuation and
Strategy Team (PVS)



Business Impact
Support better pipeline
planning

Case #2



Study
B-Cell Lymphoma
Patient Recruitment



RWD
Symphony (Claims, US)



Cooperation
Clinical Operation (CO)



Business Impact
Support find target
population

Case #3



Study
CLL Cardiovascular AEs
(CVAEs) Investigation



RWD
Flatiron (EMR, US)



Cooperation
Health Economics &
Outcomes Research (HEOR)



Business Impact
Help clinicians optimize
treatment decisions

Case #1 1L HCC Treatment Pattern

Case #1



Study

1L HCC Treatment Pattern



RWD

MarketScan (Claims, US)



Cooperation

Product Valuation and
Strategy Team (PVS)



Business Impact

Support better pipeline
planning

- **Background**

Map the **treatment** pattern for **first-line HCC** patients to support product valuation and strategy at the early discovery stage

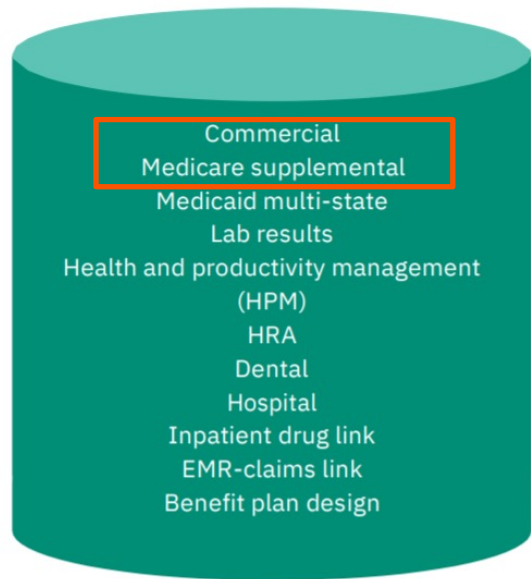


Case #1 1L HCC Treatment Pattern

- **Database**
MarketScan (Claims)

The MarketScan databases are a family of administrative claims databases that contain data on **inpatient** and **outpatient** claims, outpatient **prescription** claims, clinical **utilization** records, and healthcare **expenditures**.

Data sets



Tables

Healthcare Service	Table (Dataset)	Table Name (Dataset Name)
Services	Inpatient Admissions (I)	ccaeiyyv; mdcriyyv
	Inpatient Services (S)	ccaesyyv; mdcrsyyv
	Outpatient Services (O)	ccaeoyyv; mdcroyyv
	Facility Headers (F)	ccaefyyv; mdcrfyyv
	Outpatient Pharmaceutical Claims (D)	ccaedyyv; mdcrdyyv
	Enrollment	ccaeayyv; mdcrayyv
Enrollment	Annual Enrollment File (A)	ccaeayyv; mdcrayyv
	Enrollment Detail (T)	ccaetyyv; mdcrtyyv

Data Connection

- Hardcopy
- sas7bdat.gz

Data Access Platform

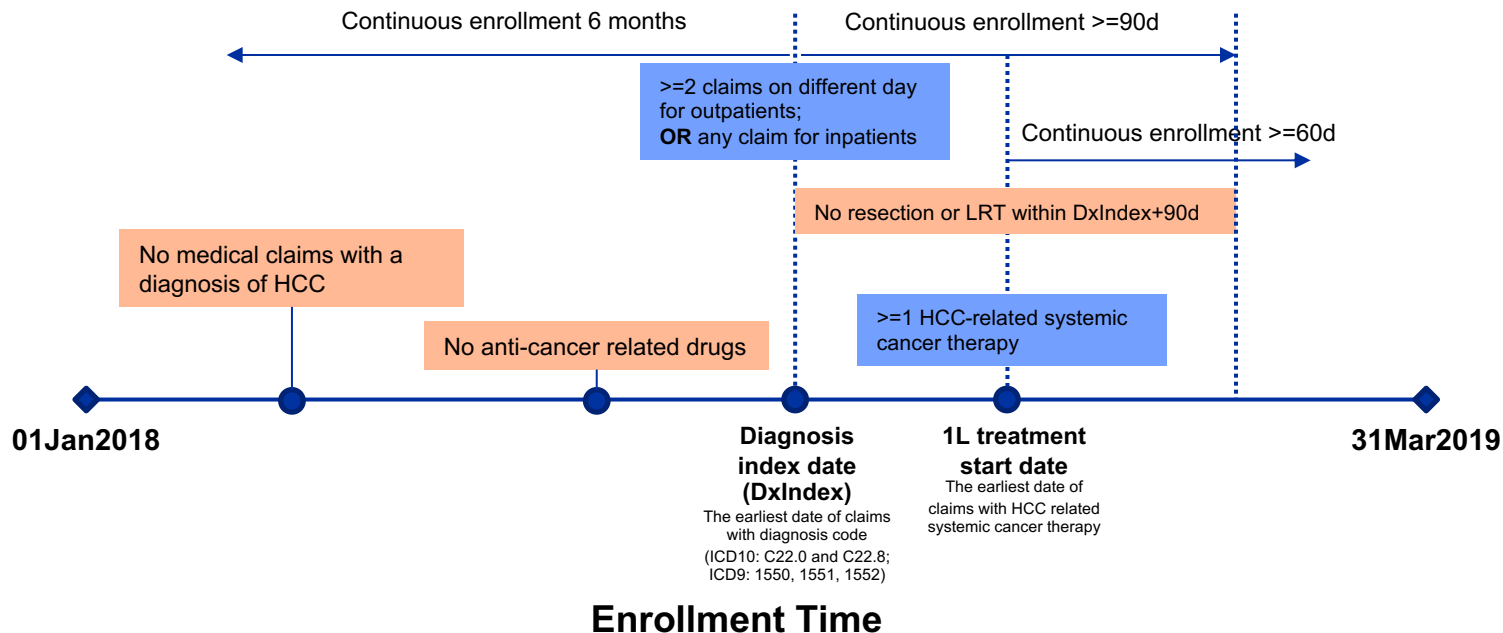
- Microsoft Azure Databricks

Data Access & Analysis Language

- Python/R/SQL

Case #1 1L HCC Treatment Pattern

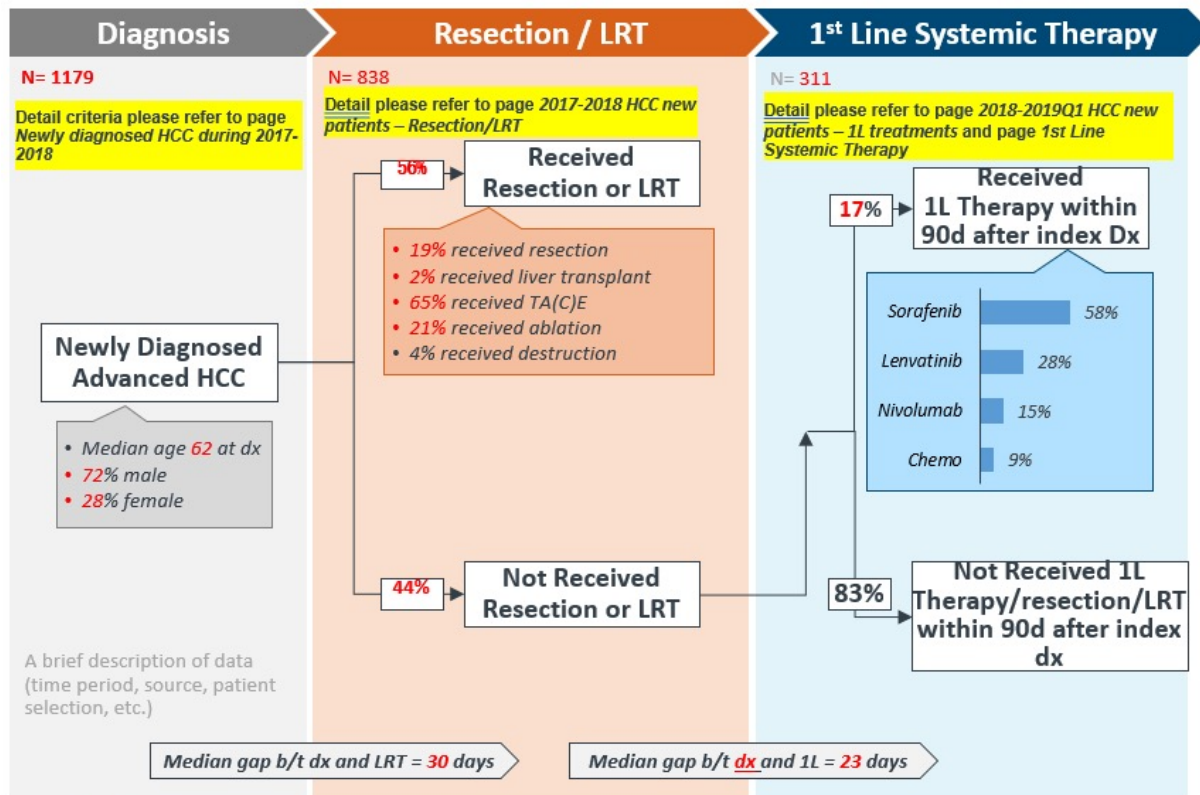
- 1 L HCC Patient Identification according to treatment pathway of MarketScan



Case #1 1L HCC Treatment Pattern

Results

HCC Treatment Flow



Notes: LRT = locoregional therapy

Case #2 B-Cell Lymphoma Patient Recruitment

Case #2



Study
B-Cell Lymphoma
Patient Recruitment



RWD
Symphony (Claims, US)



Cooperation
Clinical Operation (CO)



Business Impact
Support find target
population

- Background
 - Identify potential **new sites** and **patients** to support recruitment and help increase enrollment for a study in patients with **previously** treated **B-Cell Lymphoma** at the study execution stage



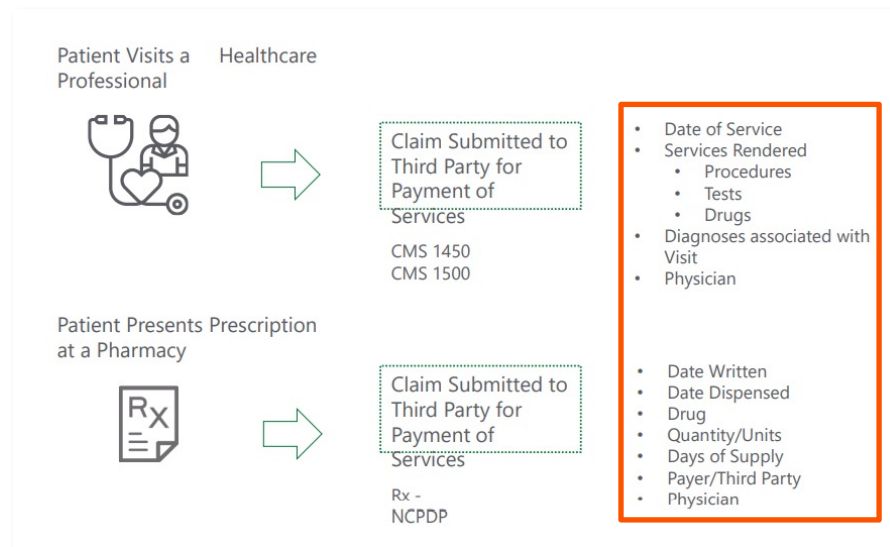
Case #2 B-Cell Lymphoma Patient Recruitment

- Database

Symphony

Symphony Data is a leading health care and insurance administrative outsourcing and information management services provider

Data Model



Data Connection

- Server

Data Access Platform

- Rstudio/Dbeaver

Data Access & Analysis Language

- R/SQL

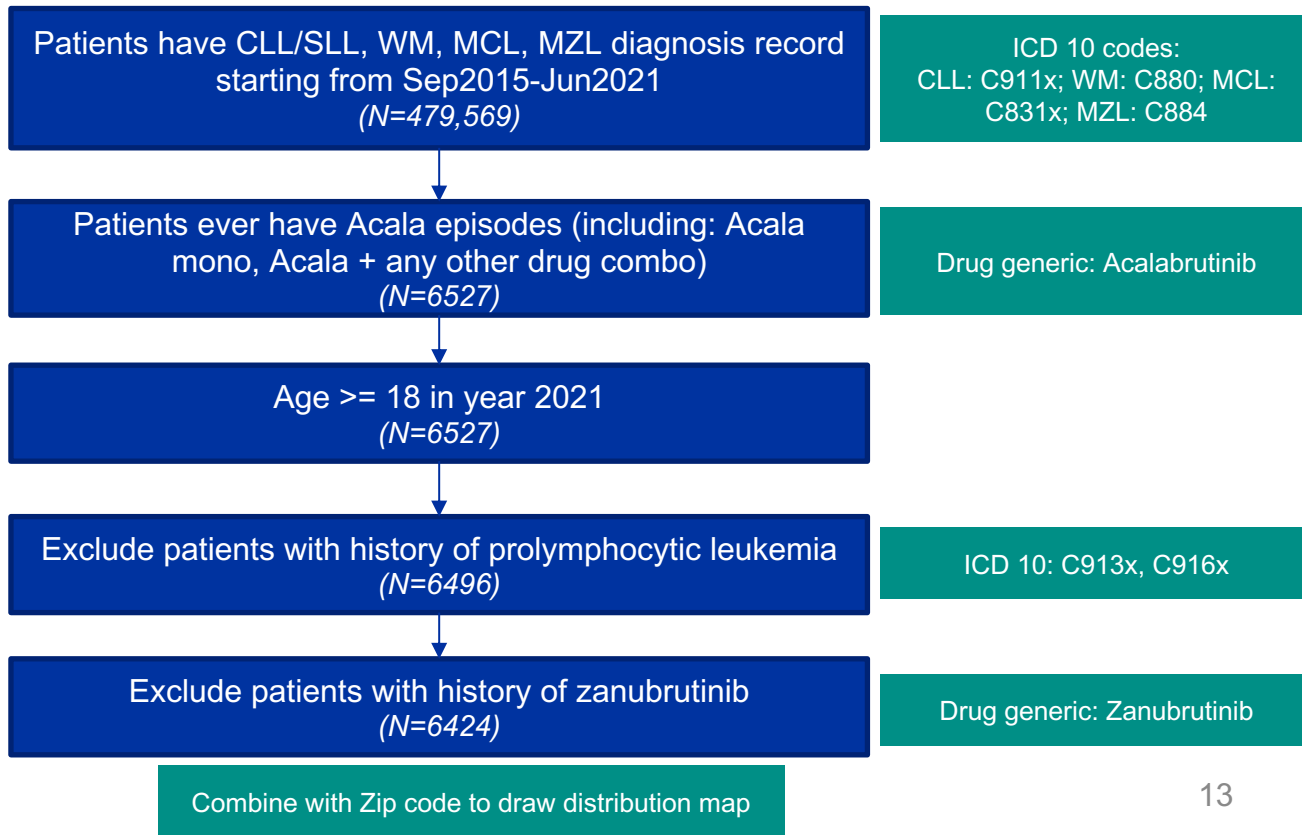
Data Visualization

- PowerBI



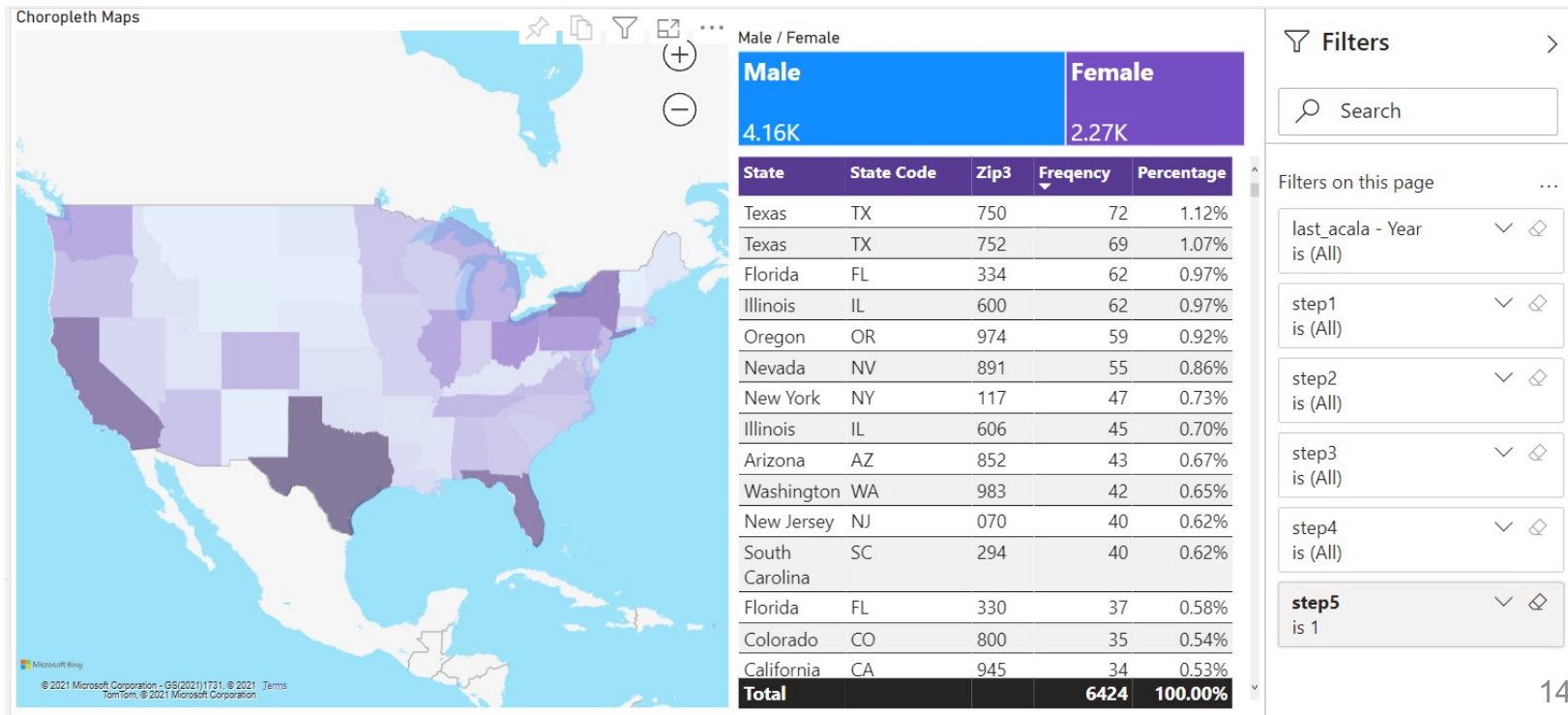
Case #2 B-Cell Lymphoma Patient Recruitment

- Potential target patient identification process
- According to protocol (I/E criteria)



Case #2 B-Cell Lymphoma Patient Recruitment

- Potential target patient distribution map



Case #3 CLL Cardiovascular AEs (CVAEs) Investigation

Case #3



Study

CLL Cardiovascular AEs (CVAEs) Investigation



RWD

Flatiron (EMR, US)



Cooperation

Health Economics & Outcomes Research (HEOR)

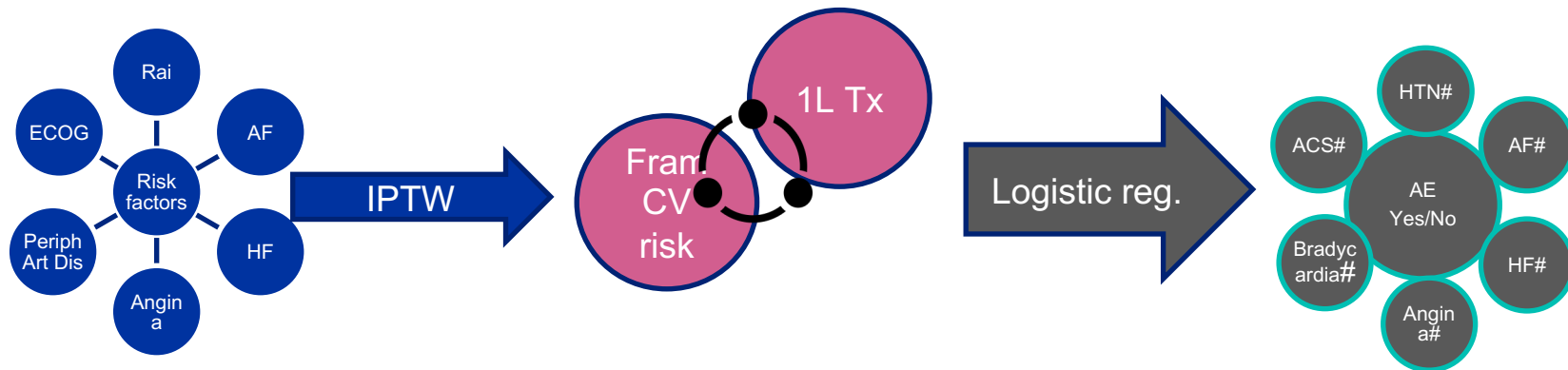


Business Impact

Help clinicians optimize treatment decisions

- Background
 - The cardiotoxicity profile of BTK inhibitor has been noted in previous trials and real-world studies and poses a major limitation to its use as a treatment progression strategy.
 - Dive deep into the relationship between the BTK inhibitor and CVAEs to help clinicians optimize treatment decisions to achieve improved outcomes.

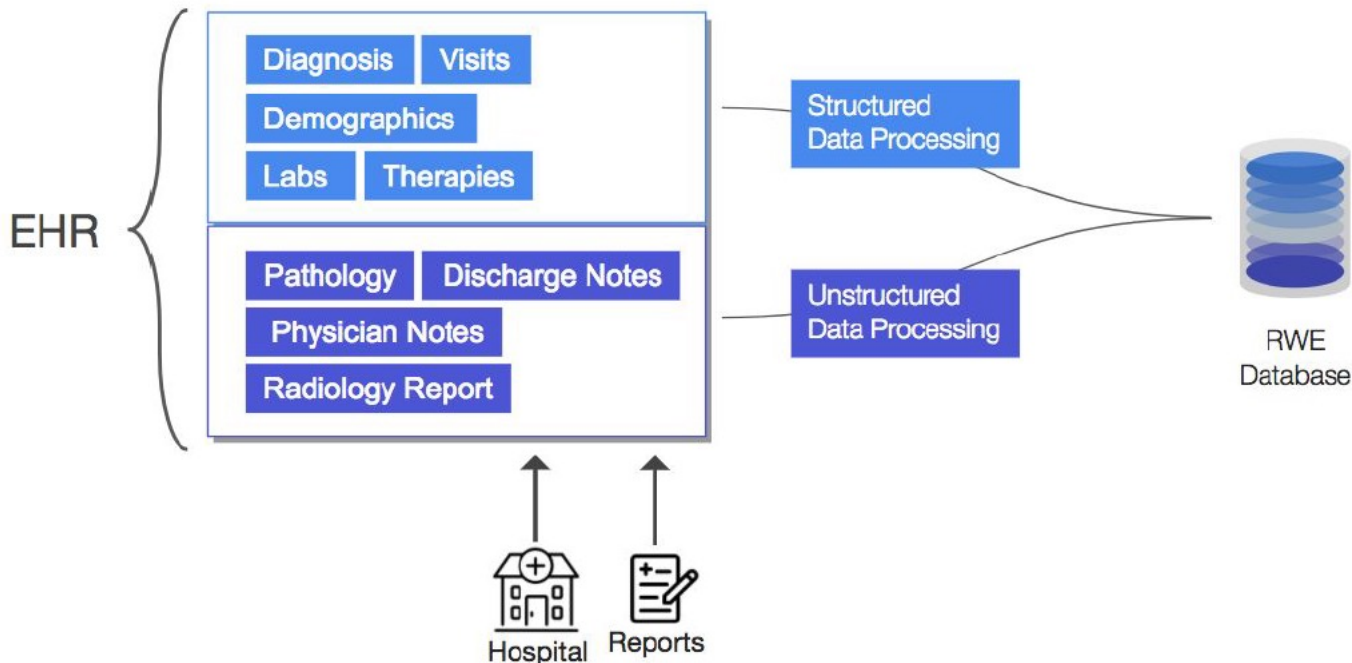
Case #3 CLL Cardiovascular AEs (CVAEs) Investigation



	Description
Main objectives of updated analysis:	Objective 1: To evaluation the association between CV risk factors and CV AEs Objective 2: To evaluate relationship between therapy selection and CV AEs after adjusting for baseline CV risk
Dependent Variables: (Binary outcome)	- Any CV AE - New or worsened hypertension - New or worsened AF/Atrial Flutter
Sample size:	Total N=515 - BTK inhibitor (n=191) - Intensive: primarily bendamustine plus anti-CD20 therapy and fludarabine, cyclophosphamide plus anti-CD20 therapy (n=195) - Non-intensive: (primarily anti-CD20therapy alone and chlorambucil plus anti-CD20 therapy) (n=129)

• Flatiron

Flatiron is an oncology platform that aggregates patient data from electronic health records (EHR), lab and medical claims and matches it with patient trial data.



Overview of Flatiron's EHR Data Processing



Case #3 CLL Cardiovascular AEs (CVAEs) Investigation

Overview of Flatiron Real-World Data Products

Core Registries + Supplemental		Clinico-Genomic Databases (CGDBs)
✓ Structured / unstructured tumor-specific clinical data variables ✓ Various supplemental variables can be added on	Data Model	✓ Tumor-specific Flatiron clinical data linked to deep genomic data provided by Foundation Medicine sequencing data
✓ Core: Monthly delivery with 30 day recency ✓ Supplemental: Monthly or quarterly	Delivery Cadence	✓ Biannual refreshes (<i>quarterly refreshes starting Q4 2019</i>), 3 month recency
✓ Core: Diagnosis codes, line of therapy info, deceased status, demographics, labs ✓ Supplemental: Sites of mets, dose modification, reasons for discontinuation, treatment spans	Example Data Elements	✓ Flatiron clinical data + FMI genomic data, including alterations (variant type / class, functional status) and complex signatures (TMB, MSI, LOH, somatic / germline zygosity)

Core registries (current) and CGDBs (by Q1 2020) can power “spotlight” projects

“Spotlights”

Data Model	Delivery Cadence	Example Data Elements
✓ Custom data variables for a defined study subpopulation	✓ One time or recurring delivery with 3-4 month recency	✓ Adverse events, real-world progression, real-world tumor response



Case #3 CLL Cardiovascular AEs (CVAEs) Investigation

Chronic Lymphocytic Leukemia Data Model

From structured data

- Demographics (year of birth, gender, race, ethnicity, state)
- Diagnosis (ICD-9 codes and ICD-10 codes and dates)
- Visits (date and type)
- Labs & Vitals (test date, result date, test name, result with units, and normal ranges)
- Medication Administration (date, drug, dose and units)
- Medication Orders (date, drug, dose and units, route)
- Performance Status (ECOG, date)
- Insurance (payer category, start date, end date)

+

From unstructured data

- Confirmation of CLL diagnosis
- Date of diagnosis
- Disease type (CLL, SLL, CLL/SLL)
- Rai stage at diagnosis
- Lymphadenopathy at diagnosis
- Hepatosplenomegaly at diagnosis
- Bone marrow biopsy (date of first documented bone marrow biopsy)
- Testing at diagnosis for first available after diagnosis for the following*:
 - Immunophenotyping by flow cytometry or IHC
 - FISH cytogenetic testing
 - Classical cytogenetic testing
 - IgHV mutation testing
- Disease transformation
- Date transformation was first identified
- Next-generation sequencing (NGS) testing status (flag only) and date of NGS testing
- Abstracted oral therapies

+

Derived data elements

Line of therapy

- Regimen name
- Line number
- Start date
- End date

Mortality

- Date of death (month and year)

- Results
 - The study confirmed that higher baseline CV risk and treatment with BTK inhibitor were independently
 - Clinical consideration may be warranted when selecting BTK inhibitor treatment for CLL patients with higher CV

- Database comparison

Database	PROS	CONS
MarketScan (Claims, US)	- Capture continuum of care: inpatient, outpatient, pharmacy data diagnosis, procedures, surgeries - Patient health plan enrollment details - Metropolitan Statistical Area (Location) in the raw database	6-9 months data lag - No info about race, disease stage, biomarker, lab result, ECOG, etc. - Need to define line of therapy
Symphony (Claims, US)	- Short (1-2 months) data lag, inpatient, outpatient, pharmacy data; - Practitioner's 5-digit postal zip code	- No info about race, disease stage, biomarker, lab result, ECOG, etc. - Need to define line of therapy
Flatiron (EMR, US)	- Has info about biomarker, race, ECOG - Defined line of therapy 1 month data lag	- Licensed by indication - Just have state level info.

- Key data field definition comparison

Key data field definition	Claims (MarketScan, Symphony)	EMR (Flatiron)	Clinical Trial Data
Indication	Need to define by diagnosis code	Pre-defined	- Pre-defined in protocol - Collected in EDC
Line of therapy (LOT)	Need to define by users	- Can provide business rule of LOT - Can provide derived variables	- Pre-defined - Collected in EDC
Drug/Therapy	Rx (Drug Table), Procedure (Inpatient & Outpatient)	- MedicationOrder and MedicationAdmin (structured) - DrugEpisode (derived from unstructured document)	Collected in EDC
Adverse Event	Need to define by diagnosis code	Can request dataset based on unstructured data	Collected in EDC or Safety database
Overall Survival	Derived from discharge status	Mortality table (structured/unstructured/external)	Collected in EDC

Thank You!

Explore a new patient-centric development paradigm together



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
📞 USA +1 609 514 5105
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://facebook.com/phusebook)
▶ [/phusetube](https://youtube.com/phusetube)
🌐 [/company/phuse](https://linkedin.com/company/phuse)