Ensuring the Validity of Real-World Evidence Studies: How Much Can You Check the Data Before You Start?

ISPOR 2025, Montreal, Canada

Date: May 16th, 2025

Moderator: **Melvin (Skip) Olson, PhD;** Olson Strategies GmbH, Allschwil, Switzerland

Presenters: **Helene Karcher, PhD**; *Life Sciences, Philip Morris Products S.A, Neuchâtel, Switzerland*

Jennifer Christian, PhD; *Target RWE, Chapel Hill, North Carolina, USA*

Mary Beth Ritchey, PhD; *Rutgers University, New Brunswick, New Jersey, USA*

Agenda

Skip Olson, ScD
moderator

Sample Size Estimation

- Problem statement
- Typical minimal sample size estimation of eligible patients conducted

Helene Karcher, PhD
panelist

Scientific and operational aspects of data checks

- Separate data checks for exposure and outcomes
- Operational challenges

Jennifer Christian, PharmD, PhD, FISPE
panelist

A staged approach to conducting CER

- Framework and stages to checking data
- Role of a clean room committee to guide the checks

Mary Beth Ritchey, PhD
panelist

Data Exploration in RWE

- Uncertainties related to data explorations in RWE when used to support decision-making
- Examples

Comments and questions are welcome during the workshop

OLSON STRATEGIES
VALUE YOUR EVIDENCE

Ensuring the Validity of Real-World Evidence Studies: How Much Can You Check the Data Before You Start?

Skip Olson
Founder, Olson Strategies GmbH

How much can you check?

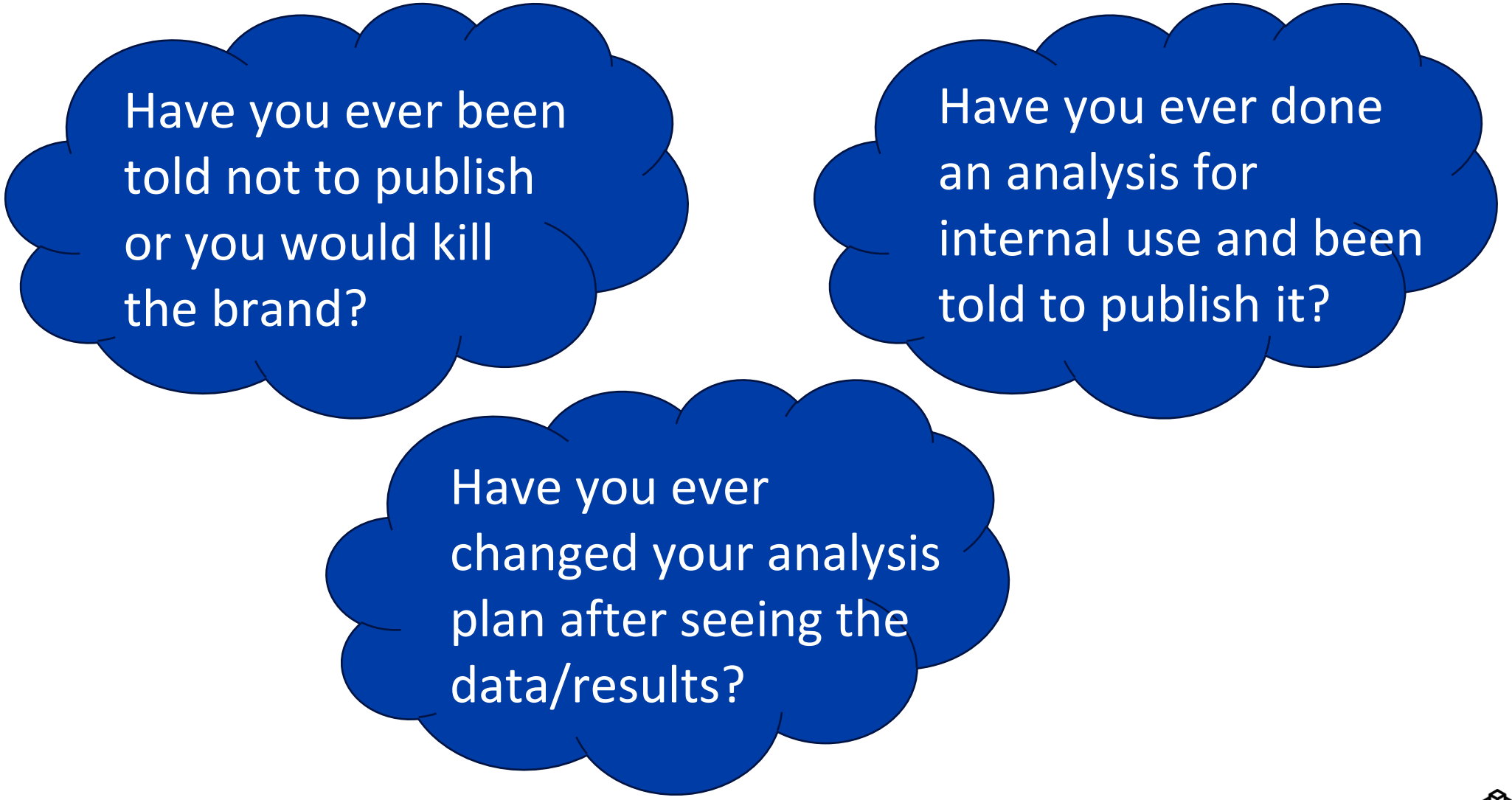

How many patients are there with this disease in this data source?


Is there an in-between solution?


How many events with Drug X versus SoC?

We will explore these questions in this Workshop!

The grey zone is large



Posting RWD study protocols

- ClinicalTrials.gov is designed for interventional trials but also supports observational studies, including RWD/RWE.
- Guidelines are evolving, especially as RWE gains importance in regulatory contexts.

When to post	Not required	Benefits
• Prospective, intervention • Supports regulatory • Journals • Required by law	• Purely retrospective • Exploratory with no interventions	• Transparency and credibility • Prevents publication bias • Facilitates reproducibility

Thank You!

Skip Olson
skipinbasel@gmail.com

Scientific and operational aspects of data checks

Helene Karcher, PhD

Life Sciences, Philip Morris Products S.A., Neuchâtel, Switzerland

Purpose of data checks before starting a RWE study for causal inference

1

Assess the Feasibility of the Study

Are there enough patients?

Are there enough events?

How are you defining events? (e.g., ICD-10)

2

Orient the Main Analysis

Establish appropriate methods and definitions in order to evaluate exposures and outcomes

Data checks allow you to establish credibility for the study design

RWE data checks in studies intended for causal inference

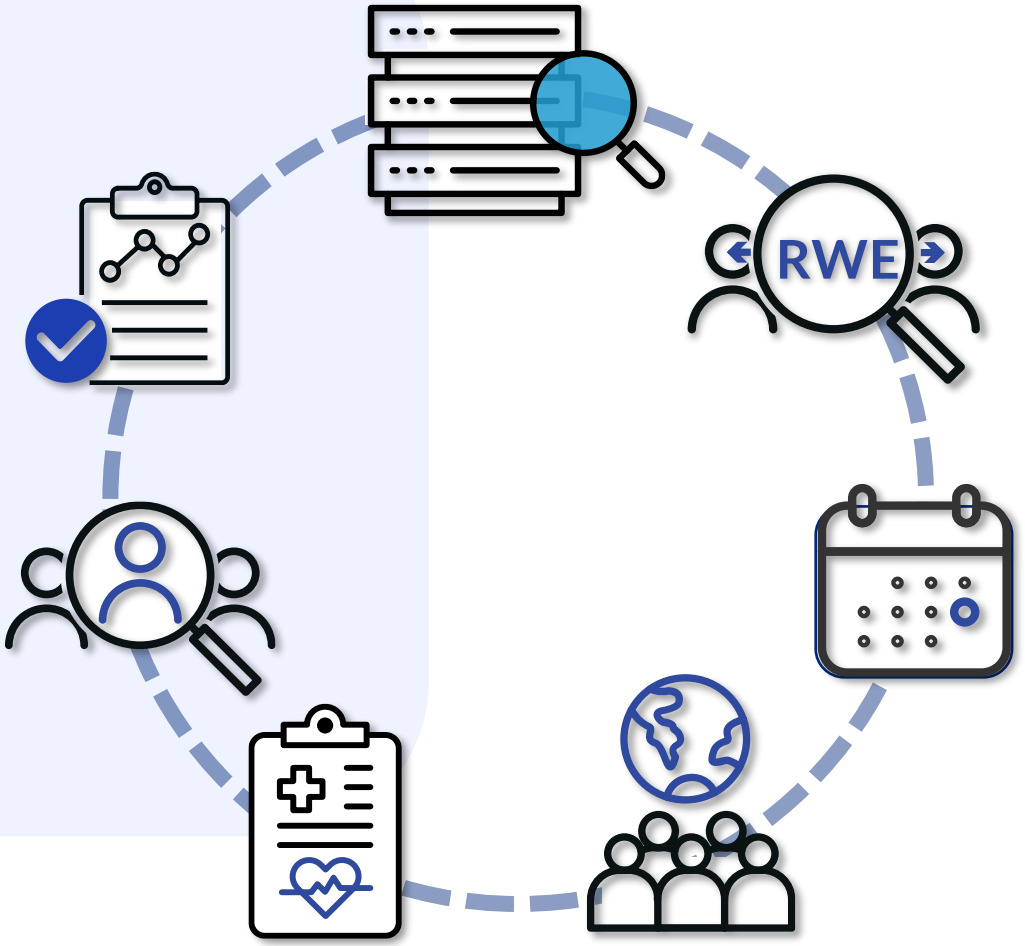
Scientific aspects

- **Population characteristics**
- **Exposure data**
  - Prevalence
  - Patterns
- **Outcomes data**

Operational hurdles

- Data availability
- Ethic committee submissions needed

Data checks on exposure and outcomes data should be separate and blinded



Example of RWE study with data checks: Comparing COPD outcomes between smokers (CIG) and users of an alternative product (HTP)

Purpose

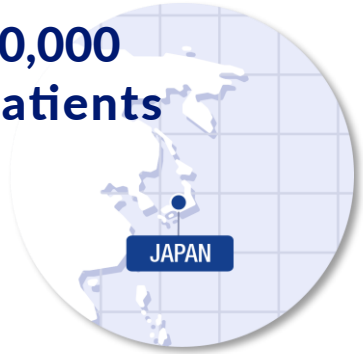
Evaluate impact of switching from CIG to HTP on time to first post-index COPD-related hospitalization or all-cause mortality

Main objective

Compare time from Index Event* to first post-index COPD-related hospitalization or all-cause mortality between exclusive HTP users and exclusive CIG smokers

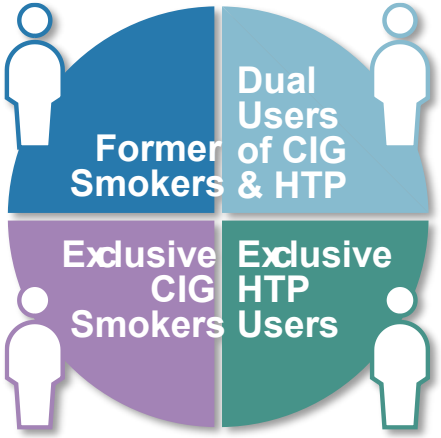
Study Population

30,000 Patients



- Record of Index Event
- Age ≥40 at time of Index Event
- Has healthcare encounter records in the database of our hospital network partners.

Smoker Type
(alive and deceased)



Data Collection

- EMR database of health and related data from a large hospital network
- Questionnaires: tobacco exposure, mortality
- Consumer databases

*Index Event = First COPD-related hospitalization within established time frame; EMR = Electronic Medical Record

Reference: Chronic Obstructive Pulmonary Disease Outcomes Among Individuals in Japan Who Switched to Heated Tobacco Products Compared to Those Who Continued Smoking or Formerly Smoked Combustible Cigarettes: Protocol for a Real-World Retrospective Study - PubMed

Data checks for exposure and outcomes are performed on separate datasets

1 Exposure Data Check

- **Data Source: exposure questionnaires**
- **Data Check:** investigate persistence on HTP to find a balance in exposure group definition between:
  - Duration of HTP exposure
  - Enough patients in HTP group
  - Start of HTP exposure with respect to index event

2 Outcomes Data Check

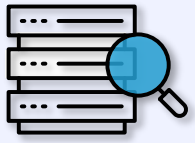
- **Data Source:** electronic medical database from a network of hospitals across Japan
- **Data Check:** outcomes
  - **Time to event analysis:** Are there enough events ?
  - Median length of time to first post-index event
- Available covariates (check missingness, distribution)
  - Will they be present to adjust populations in the main analysis?
- Other design considerations
  - Example: are they covariates that are associated with high prevalences of HTP usage?

**Data checks for exposure and outcomes data should be separate and blinded*

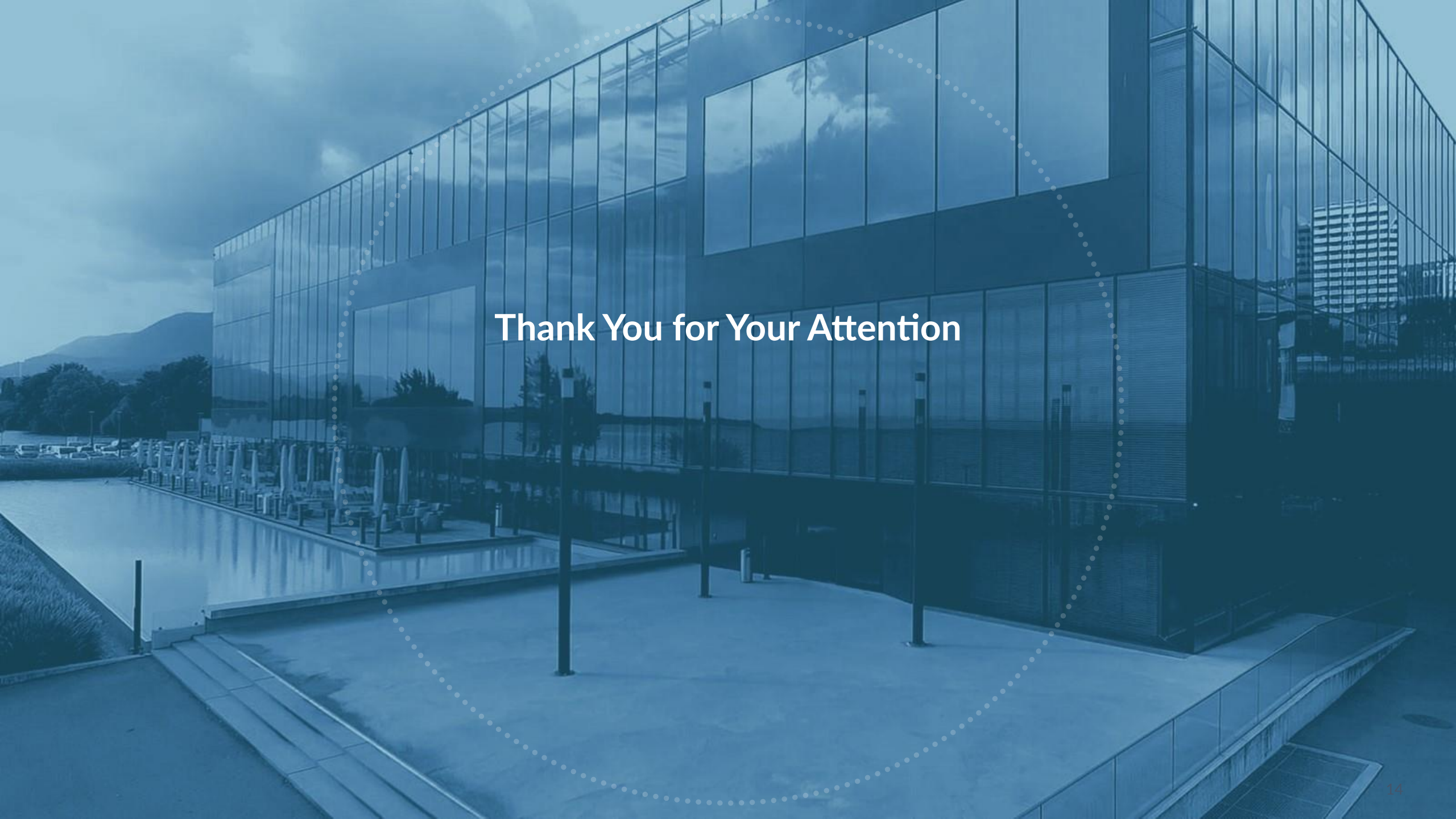
Operational hurdles

- **Willingness of Data Partners to Share Data for Analysis Before Study Begins**
  - Some partners want to do the data checks themselves but do not have the statistical capabilities -> possibility depends on how extensive the data checks are
  - Some partners require a commitment for a full study to conduct the data checks (contract in place, commitment to a certain budget) -> also dependent on how extensive the data checks are
  - Some partners require Ethics Committee approval to conduct data checks

- **Time**
  - Some data checks could take too long to implement (see above) to be compatible with the timeline required for the study to have business impact



- **Data Standardization**


Thank You for Your Attention

A staged approach to conducting comparative effectiveness studies

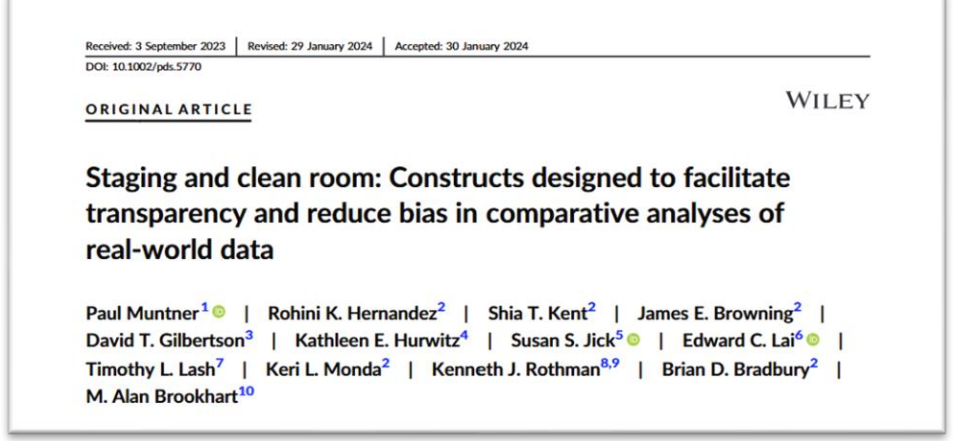
May 2025

Jennifer Christian, PharmD, PhD, FISPE
Chief Scientific Officer

TARGET RWE
BETTER EVIDENCE • BETTER HEALTH

Staging and clean room

Constructs designed to facilitate transparency and reduce bias in comparative RWD analyses



Received: 3 September 2023 | Revised: 29 January 2024 | Accepted: 30 January 2024
DOI: 10.1002/pds.5770

ORIGINAL ARTICLE

WILEY

Staging and clean room: Constructs designed to facilitate transparency and reduce bias in comparative analyses of real-world data

Paul Muntner[1] | Rohini K. Hernandez[2] | Shia T. Kent[2] | James E. Browning[2] | David T. Gilbertson[3] | Kathleen E. Hurwitz[4] | Susan S. Jick[5] | Edward C. Lai[6] | Timothy L. Lash[7] | Keri L. Monda[2] | Kenneth J. Rothman[8,9] | Brian D. Bradbury[2] | M. Alan Brookhart[10]

Stage 1a: Study Development

- Feasibility assessment
- Protocol development
- Develop analysis plan
- Establish roles/responsibilities
- Register protocol

Stage 1b: Build cohort

- Set up clean room(s)
- Apply inclusion/exclusion criteria
- Assess sample size/precision

Check Point 1: Assess sample size, number of events, precision of estimates

Stage 2: Assess Covariate balance

- Assess patient characteristics
- Propensity score/ treatment weighting

Check Point 2: Assess comparability of treatment arms based on propensity scores

Stage 3: Assess residual confounding

- Negative control outcome analyses

Check Point 3: Assess comparability of treatment arms based on analysis of residual bias

Stage 4: Conduct comparative analyses†

Publish results

◆ Step involved in the staging framework conducted by investigators

⬢ Check Point: Time when a decision is made to stop, continue, or revise the study

Results are shared with a review team (after Stages 1 through 3) or disseminated (after Stage 4)

† Only after all check points have been passed should the comparative analyses be conducted.

TARGET RWE

Example Study

RW External control arm for a single arm trial

Zanidatamab for the treatment of HER2-amplified Biliary Tract Cancer

- Biliary tract cancer is a rare, aggressive cancer with few effective treatment options
  - Includes bile duct cancer, gallbladder cancer, and cholangiocarcinoma
  - US incidence is 0.35 to 2 per 100K annually
  - Represents <1% adult cancers
  - 5-year overall survival = 3% to 19%
    - Median survival with first line therapy → 11.7 to 12.8 months

- Treatments include surgery, chemotherapy, radiation therapy, targeted therapies and immunotherapy. No targeted agents currently approved in 1L setting and most patients are not eligible for available 2L targeted agents.

- HER2 (or its gene) mutated, amplified, or overexpressed in 4%-31% of biliary tract cancers

- Zanidatamab is a novel bispecific antibody, meaning it binds to two different regions on the HER2 protein simultaneously (dimerization domain and extracellular juxtamembrane domain) leading to more potent anti-tumor effect

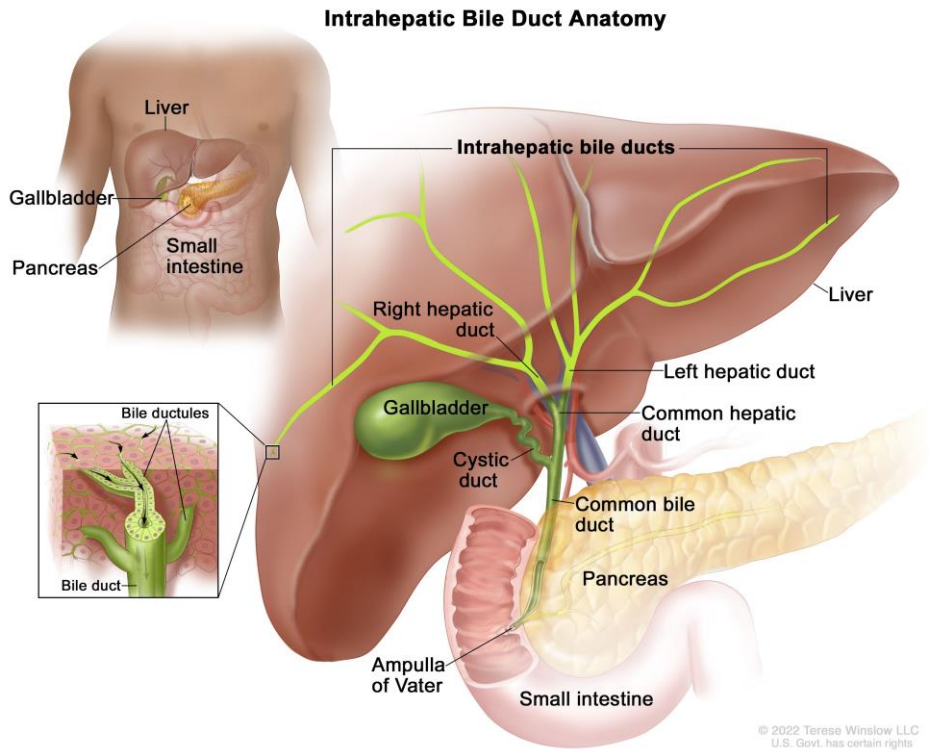
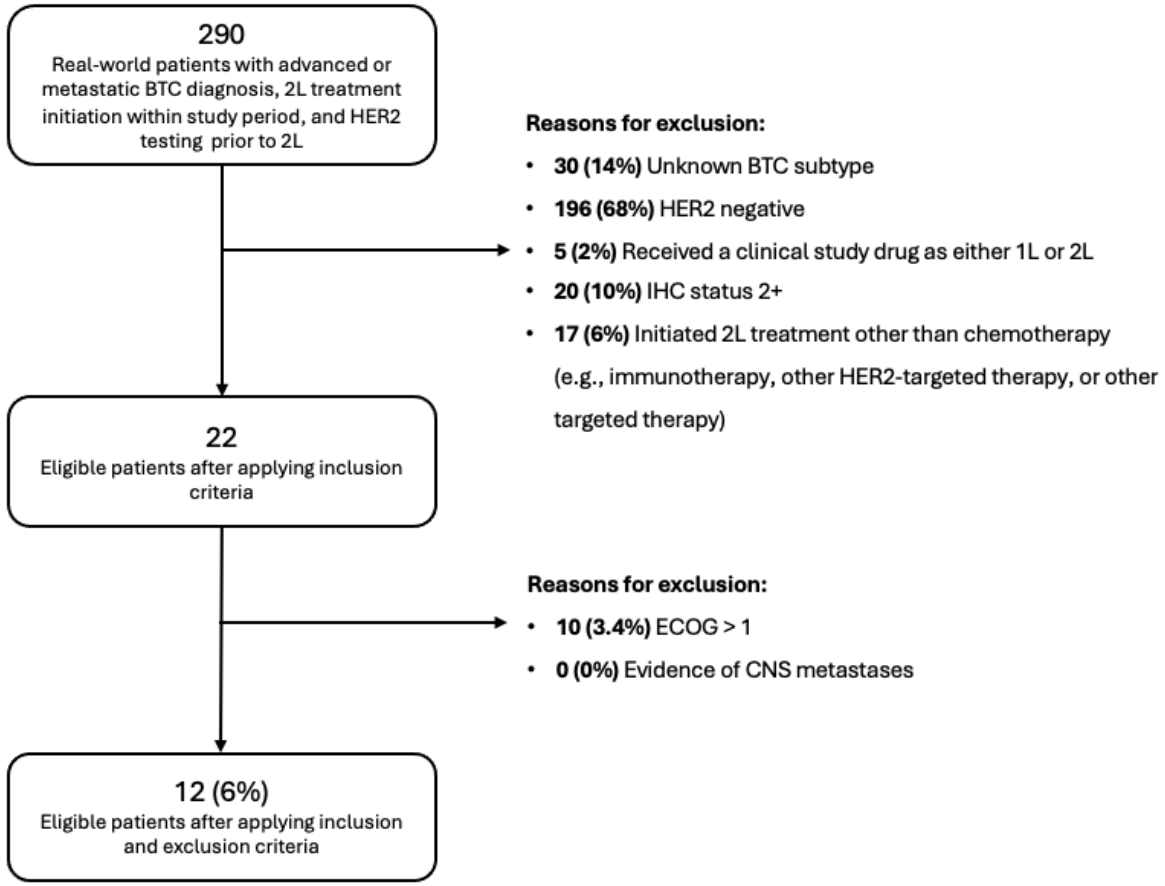


TABLE 1 Example of roles for team members conducting a comparative analysis with the staging and clean room constructs.

	Analysis stage			
	Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
Investigators				
Programmer	Establish clean room place data sets in clean room	Add covariates to data set	Add negative control outcomes to data set	Add outcomes to data set
Analyst	Apply inclusion/exclusion criteria	Assess covariate balance prepare propensity score, inverse probability treatment weights conduct trimming of data as needed	Conduct negative control outcomes analyses	Conduct comparative analyses
Analytic advisor (lead investigator)	Develop study aims Identify appropriate data set draft and publicly post protocol review estimates of sample size, number of outcomes and/or statistical precision	Review data to assess covariate balance	Review data to assess if residual confounding is present	Confirm analyses were done following the study protocol prepare final study report
Review team	Determine if adequate statistical precision is available	Determine if the treatment groups are comparable	Determine if residual confounding is present	None

CHECKPOINT 1: Adequate Study Size/Statistical Precision



Study flow diagram of Oncology EHR external control arm.

- The *a priori* precision analyses had been conducted using estimated sample sizes of 160, 180, and 200, finding adequate precision for primary objective
- But after receiving the data, there were much higher rates of exclusion
  - Mostly due to lack of IHC3+ test results pre-initiation of 2L (78%, n=216)
- Updated precision analyses using the final sample size of n=12 found an estimated 95% CI of **(0.13, 0.72)** for the HR comparing OS

Abbreviations: BTC: biliary tract cancer; IHC: immunohistochemistry; CI: confidence interval; HR: hazard ratio

CHECKPOINT 2: Covariate Balance

- Due to known challenge of sample size, another step taken before main analysis:
  - We used a systematic literature review and input from KOLs to create a pool of important prognostic factors for BTC
  - With KOL feedback, we ranked prognostic factors by importance for inclusion in SMR-weight model (i.e., we ranked patient characteristics by how likely they were to confound the treatment effect on overall survival)
    - Created an *a priori* plan for dealing with model non-convergence (collapse categories, remove lower ranked variables, etc.)

- We ended up having to drop ”group stage at diagnosis” from SMR-weight model due to ~75% missingness
  - Per our study protocol, we only considered multiple imputation for covariates with <40% missingness

TARGET RWE

Abbreviations: KOL: Key Opinion Leader; IHC: immunohistochemistry

CHECKPOINT 2: Covariates listed apriori by priority

Priority	Variable	Values available in the data	Implementation 1	Implementation 2 (Collapsed/simplified)
High	Disease subtype	Categorical variable: • GBC • iCCA • eCCA	Implemented as two binary indicators to accommodate three possible states of disease subtype: • GBC • iCCA • eCCA	No change.
Medium	Group stage at initial diagnosis	Categorical variable: • Stage I • Stage II • Stage III • Stage IV	Implemented as three binary indicators to accommodate four possible stage values: • Stage I • Stage II • Stage III • Stage IV	Implemented as one binary variable, where stages are collapsed into I and II-IV: • Stage I • Stage II, III or IV
High	Liver disease	Two binary variables: • Mild chronic liver disease – yes or no • Moderate to severe chronic liver disease – yes or no	Implemented as two binary indicators to accommodate three possible states: • No liver disease • Mild chronic liver disease • Moderate to severe chronic liver disease	Implemented as one binary variable, where no liver disease and mild chronic liver disease are collapsed into 'no or mild chronic liver disease: • No or mild chronic liver disease • Moderate to severe chronic liver disease
Medium	Sex	Categorical variable: • Female • Male	Implemented one binary indicator to accommodate two possible states: • Female • Male	No change.
Low	Age	Continuous variable (in years).	Implemented as one continuous variable indicating age in years at the time of 2L initiation using a spline function.	No change.

Covariates by priority and possible implementations in the propensity score model

Abbreviations: KOL: Key Opinion Leader; IHC: immunohistochemistry

CHECKPOINT 2: Covariate Balance Using SMR Weighting

STAGE: **2** IHC STATUS: **IHC3+ Only** 2L TREATMENT: **Chemotherapy Only**

	Unweighted Zanidatamab	Unweighted ECA	Unweighted SMD	Weighted Zanidatamab	Weighted ECA	Weighted SMD
ELIGIBLE PATIENTS						
N	62	12		62.0	62.1	
AGE AT INITIATION OF 2L TREATMENT						
Mean (SD)	62.65 (9.31)	65.92 (8.65)	0.360	62.65 (9.31)	63.28 (8.38)	0.072
DISEASE SUBTYPE (COLLAPSED)						
Gallbladder cancer	33 (53%)	9 (75%)	0.470	33 (53%)	31 (50%)	0.056
Intrahepatic or extrahepatic cholangiocarcinoma	29 (47%)	3 (25%)		29 (47%)	31 (50%)	
HISTORY OF CHRONIC LIVER DISEASE						
No	51 (82%)	10 (83%)	0.030	51 (82%)	55 (89%)	0.203
Yes	11 (18%)	2 (17%)		11 (18%)	7 (11%)	
SEX						
Female	34 (55%)	8 (67%)	0.240	34 (55%)	34 (55%)	0.001
Male	28 (45%)	4 (33%)		28 (45%)	28 (45%)	

- Weighting a small sample to a larger one can create problems in balancing across all characteristics
- Example: history of chronic liver disease was **less balanced** across trial and ECA cohorts after SMR-weighting

Abbreviations: IHC: immunohistochemistry

CHECKPOINT 2: Covariate Balance Using SMR Weighting

STAGE: **2** IHC STATUS: **IHC3+ Only** 2L TREATMENT: **Chemotherapy Only**

	Unweighted Zanidatamab	Unweighted ECA	Unweighted SMD	Weighted Zanidatamab	Weighted ECA	Weighted SMD
ELIGIBLE PATIENTS						
N	62	12		62.0	62.1	
AGE AT INITIATION OF 2L TREATMENT						
Mean (SD)	62.65 (9.31)	65.92 (8.65)	0.360	62.65 (9.31)	63.28 (8.38)	0.072
DISEASE SUBTYPE (COLLAPSED)						
Gallbladder cancer	33 (53%)	9 (75%)	0.470	33 (53%)	31 (50%)	0.056
Intrahepatic or extrahepatic cholangiocarcinoma	29 (47%)	3 (25%)		29 (47%)	31 (50%)	
HISTORY OF CHRONIC LIVER DISEASE						
No	51 (82%)	10 (83%)	0.030	51 (82%)	55 (89%)	0.203
Yes	11 (18%)	2 (17%)		11 (18%)	7 (11%)	
SEX						
Female	34 (55%)	8 (67%)	0.240	34 (55%)	34 (55%)	0.001
Male	28 (45%)	4 (33%)		28 (45%)	28 (45%)	

- Study team consulted with Clean Room Committee (CRC)
- Two analytic options:

1. Include history of chronic liver disease as a variable in the Cox proportional hazards model that estimates treatment effect on OS
2. Accept the SMD of 0.2 to maintain ability to estimate a marginal, rather than conditional, treatment effect

➢ CRC recommended #2, but also recommended exploring residual confounding with quantitative bias analysis

Abbreviations: CRC: Clean Room Committee; SMD: standardized mean difference; IHC: immunohistochemistry

CHECKPOINT 3: Assessment of Potential Impact of Residual Confounding

STAGE: **2** IHC STATUS: **IHC3+ Only** 2L TREATMENT: **Chemo Only**

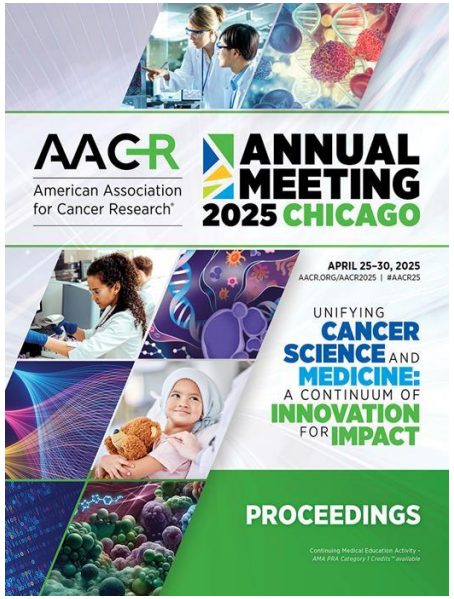
	N	Mean	Min	5th quartile	25th quartile	50th quartile	75th quartile	95th quartile	Max
ECA	12	5.18	2.09	2.48	3.34	3.57	5.08	11.84	18.84
Zanidatamab	62	1	1	1	1	1	1	1	1

- After applying inverse probability of treatment weights, we examined the weight distribution for outliers
- One patient had an extreme (>18) weight
  - We conducted a sensitivity analysis, removing this patient from the primary analysis
- Another patient had died soon after the censoring window
  - We conducted a sensitivity analyses extending the censoring window from 90 days after last EHR activity to 120 days after last EHR activity

Abbreviations: IHC: immunohistochemistry; EHR: electronic health record

Next steps

- ASCO abstract presenting the main comparative effectiveness results for overall survival
- The biggest limitation of current study is the small sample size in the real-world cohort
  - ➢ To address limitation, we will be pooling data from sources in US, France, and Spain



[Abstract 5954: Real-world (RW) second-line (2L) treatment (tx) patterns and clinical outcomes in patients (pts) with HER2-overexpressing biliary tract cancer (BTC) | Cancer Research | American Association for Cancer Research]

2025 ASCO ANNUAL MEETING

Abstract 4101. Survival outcomes for zanidatamab-hrii compared to chemotherapy in previously treated HER2-positive (IHC3+) biliary tract cancer (BTC): HERIZON-BTC-01 vs a real-world (RW) external control arm (ECA).

Conclusions

- We are in the era of RWE – with increased access to rich, clinical RWD and advanced epidemiology methods
- Principled approaches to conducting comparative effectiveness and safety studies are needed, which include transparent methods for addressing potential limitations in the data
- Numerous challenges can arise well after an approved protocol – a staged approach and a Clean Room Committee can provide a transparent way of reducing the risk of bias, increasing reliability of RWE, and improved trust in the process

Thank you to the team!

- **TargetRWE Team Study Members:** Kathleen Hurwitz, Kayla Hendrickson, Catherine Wiener
- **Study collaborators:** Richard Kim, Xiaozhou Fan, Javier Sabater, Wayne Su, Phillip Garfin, Joan Zape, Mark A. Ozog, John Bridgewater, Juan W. Valle, Farshid Dayyani
- **Clean Room Committee:** M. Alan Brookhart, Michael Fried, Jennifer Christian

- **Staging and Clean Room Framework**: Paul Muntner, Rohini Hernandez, Shia Kent, James Browning, David Gilbertson, Kathleen Hurwitz, Susan Jick, Edward Lai, Tim Lash, Keri Monda, Ken Rothman, Brian Bradbury, Alan Brookhart

- **Advancing Principled Pharmacoepidemiologic Methods:** Rohini Hernandez, Cathy Critchlow, Nancy Dreyer, Tim Lash, Robert Reynolds, Henrik Sorensen, Jff Lange, Nicolle Gatto, Rachel Sobel, Edward Chia-Cheng Lai, Marieke Schoonen, Jeff Brown, Jennifer Christian, M. Alan Brookhart, Brian Bradbury

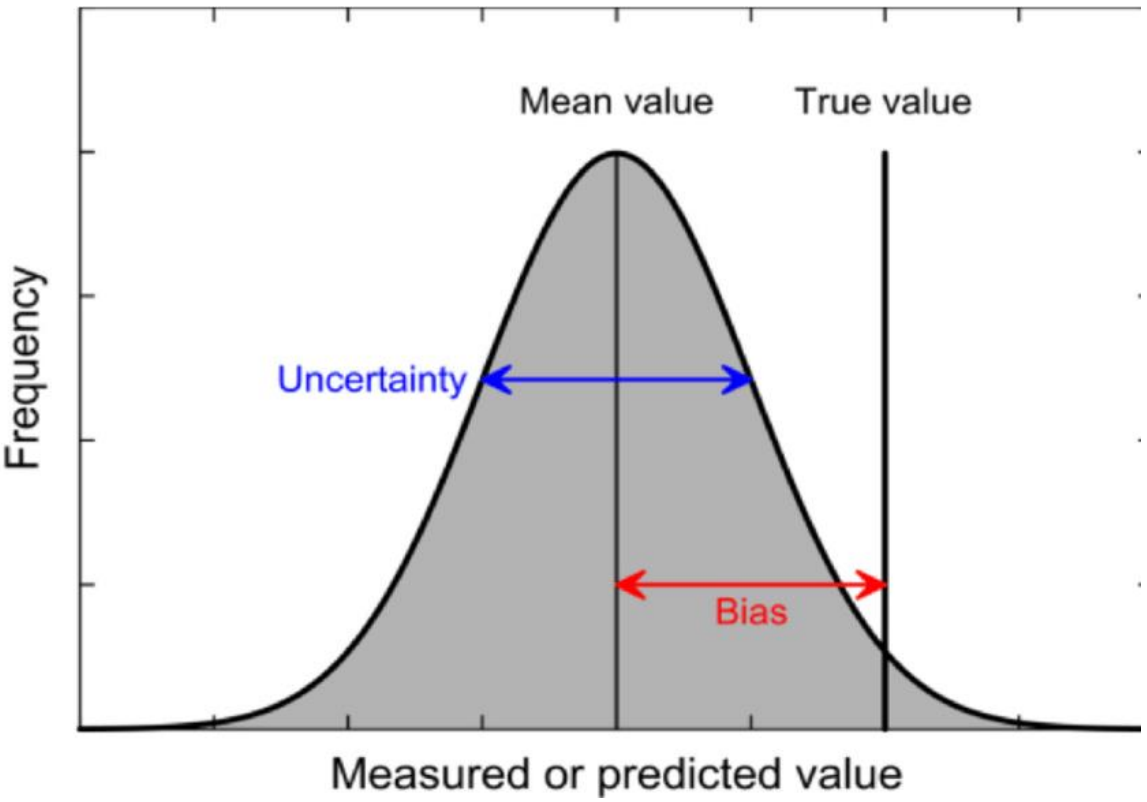
Data Explorations and Deciding on a RWD Source

What questions are we solving for?

Mary Beth Ritchey, PhD, FISPE

Med Tech Epi, LLC; Rutgers University PETS & HOPE

What is the issue?



Regulatory (or other Stakeholder) Review

Are the proposed study design and analysis likely to address the research question?

Uncertainty

Variability in a prediction or measurement

Due to random error, heterogeneity, or other factors that cause variability

Expressed as interval around a point estimate

Bias

Systematic error that skews data

Due to systematic variation in measurement or study design, conduct, or analysis

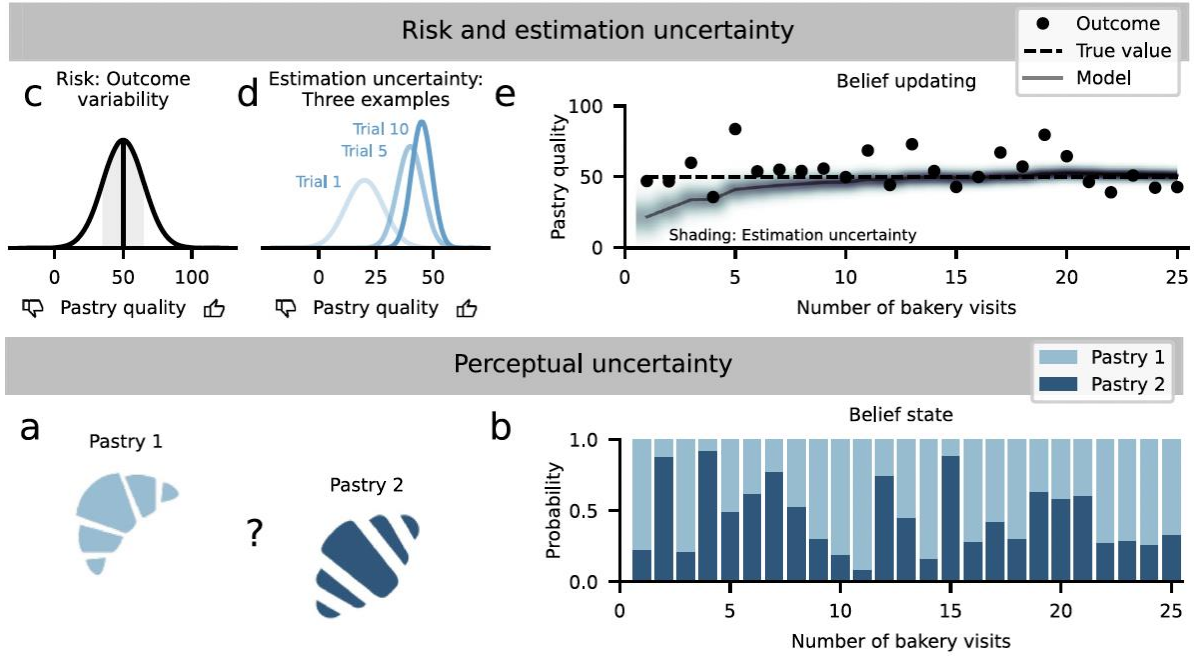
Can be addressed through study design and analysis

Uncertainty

"There can be uncertainty around the type, magnitude, duration, frequency, and other aspects of benefits and risks to patients"
– US FDA Medical Device Guidance on Uncertainty

This is true of all variables in all studies –
need enough data to have stable estimates

With RWD, there is less control over data collected for the study and regulatory reviewers have less comfort with the data



Sources: US FDA. Consideration of Uncertainty in Making benefit-risk determinations in medical device premarket approvals, de novo classifications, and humanitarian device exemptions. 2019. Bruckner. Understanding learning through uncertainty and bias. Commun Psychol. 2025 Feb 13;3(1):24.

Bias

"Operational definitions are usually imperfect… misclassification… may bias the association between exposure and outcome…"

– US FDA Drug/Biologic Guidance on Assessing EHR as RWD

Impact of misclassification is influenced by:

- Degree of misclassification
- Differential vs non-differential misclassification
- Dependent vs independent misclassification
- Directional bias of the association between the treatment and the outcome

Table 1 Hypothetical example of studies in which individuals exposed to one of two drugs are each compared with non-users, or compared with each other in the presence of nondifferential exposure misclassification.

	Truth		Observed Data	
Drug A vs. Nonusers	Drug A_1	Drug A_0	Drug A_1	Drug A_0
Y=1	2,000	5,000	1,950	5,050
Y=0	8,000	45,000	9,050	43,950
Total	10,000	50,000	11,000	49,000
Risk	0.20	0.10	0.18	0.10
OR	2.25		1.88	
RR	2.00		1.72	
RD	0.10		0.07	
Drug B vs. Nonusers	Drug B_1	Drug B_0	Drug B_1	Drug B_0
Y=1	200	5,000	280	4,920
Y=0	800	45,000	1,620	44,180
Total	1,000	50,000	1,900	49,100
Risk	0.20	0.10	0.15	0.10
OR	2.25		1.55	
RR	2.00		1.47	
RD	0.10		0.05	
Drug A vs. Drug B	Drug A_1	Drug B_1	Drug A_1	Drug B_1
Y=1	2,000	200	1,950	280
Y=0	8,000	800	9,050	1,620
Total	10,000	1,000	11,000	1,900
Risk	0.20	0.20	0.18	0.15
OR	1.00		1.25	
RR	1.00		1.20	
RD	0.00		0.03	

Drug A vs. non-users: sensitivity 0.85, specificity 0.95. Drug B vs. non-users: sensitivity 0.90, specificity 0.98. All are non-differential with respect to disease (Y) status. No individuals exposed to Drug A are misclassified as exposed to Drug B or vice versa. OR: odds ratio; RR: risk ratio; RD: risk difference. Y=1: experienced the outcome of interest; Y=0: did not experience the outcome of interest.

Sources: US FDA. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. 2024. Funk. Misclassification in Administrative Claims Data: Quantifying the Impact on Treatment Effect Estimates 2014.

What info do we need to show that a data source can address a research question?

Knowledge of Data Source

Reason for selecting

Timeframe of data available

Background data on health care system(s)

Specific method of collecting diagnoses and treatments

Any practices which may influence feasibility of study or interpretation of findings (e.g. stepped therapy, prior authorizations, formulary restrictions)

Generalizability of study in this data source to target population

Data management: data accrual, curation, transformation into final dataset

Source: US FDA. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. 2024.

Continuity of Care and Coverage

Continuity of Care

Interactions with health care are available in data source

Continuity of Coverage

Timeframe of patient availability in data source

Source: US FDA. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. 2024.

Data Elements

Availability and missingness of key variables (initial feasibility)

Treatment and comparator, primary outcome, age and sex,

Conceptual definition of variables: clinical criteria to define condition or measurement of intervention

Operationalization of variables: detailed description of how each variable will be defined for the study

Source: US FDA. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. 2024.

Types of Validity

Face validity – does this measure *appear* to make sense?

Construct validity – does this measure the *intended* concept?

Content validity – does the measure cover *all aspects* of the intended concept?

Criterion validity – Do the results *accurately* measure what they are designed to measure?

Validation and Data Elements

Exposure – content validity

- Active comparator, new user design requires specific code/date
- Billing for medicines lends credibility, if code is present

Population – depends

- Rare disease, new therapeutic area – criterion validity
- Common condition, known therapeutic area – construct validity

Key covariates

- Usually construct validity

Outcome – criterion validity

- Quantitative information about misclassification provides needed information to conduct sensitivity analysis
- Conduct formal validation study

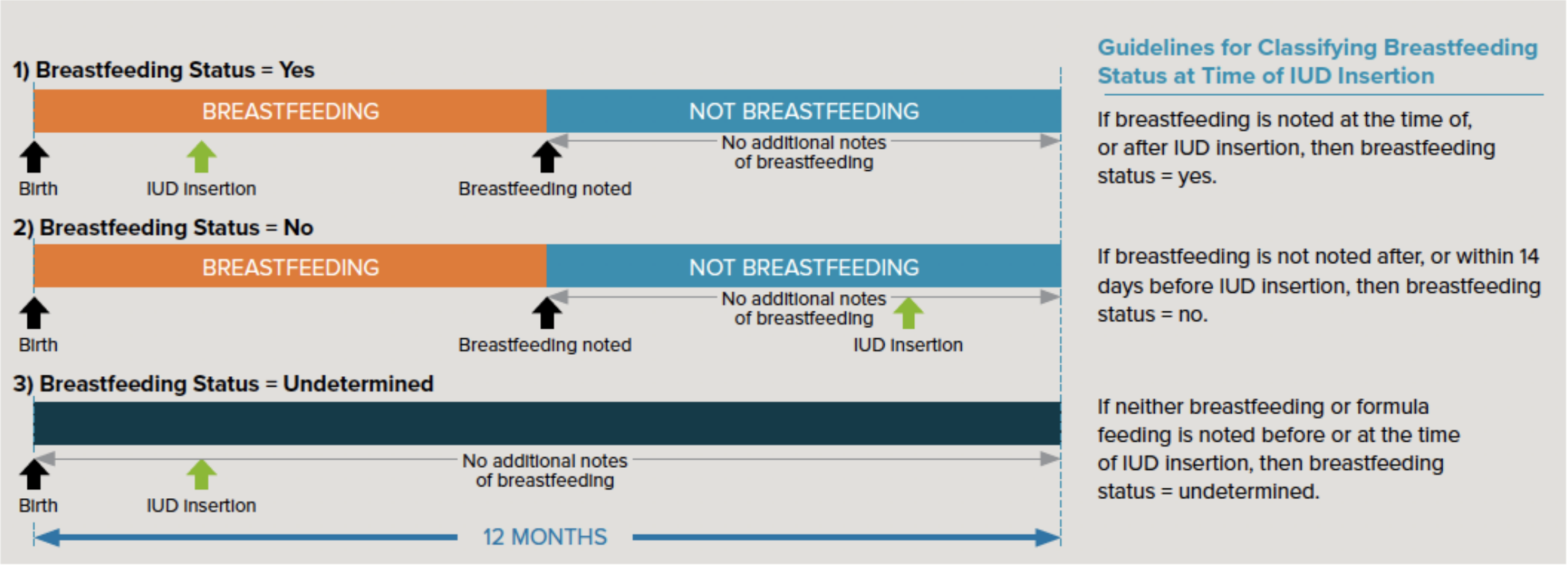
How to Conduct Validation Study

1. Select the appropriate health outcome
2. Determine the reference standard against which to validate the algorithm
3. Develop the case-identifying algorithm
4. Select persons for validation
5. Collect relevant data to confirm the health outcome of interest
6. Assess the algorithm performance

Source: Weinstein EJ, Ritchey ME, Lo Re V III. Core concepts in pharmacoepidemiology: Validation of health outcomes of interest within real-world healthcare databases. PDS 2022;1-8. doi:10.1002/pds.5537

Examples

Example: Algorithm for Breastfeeding

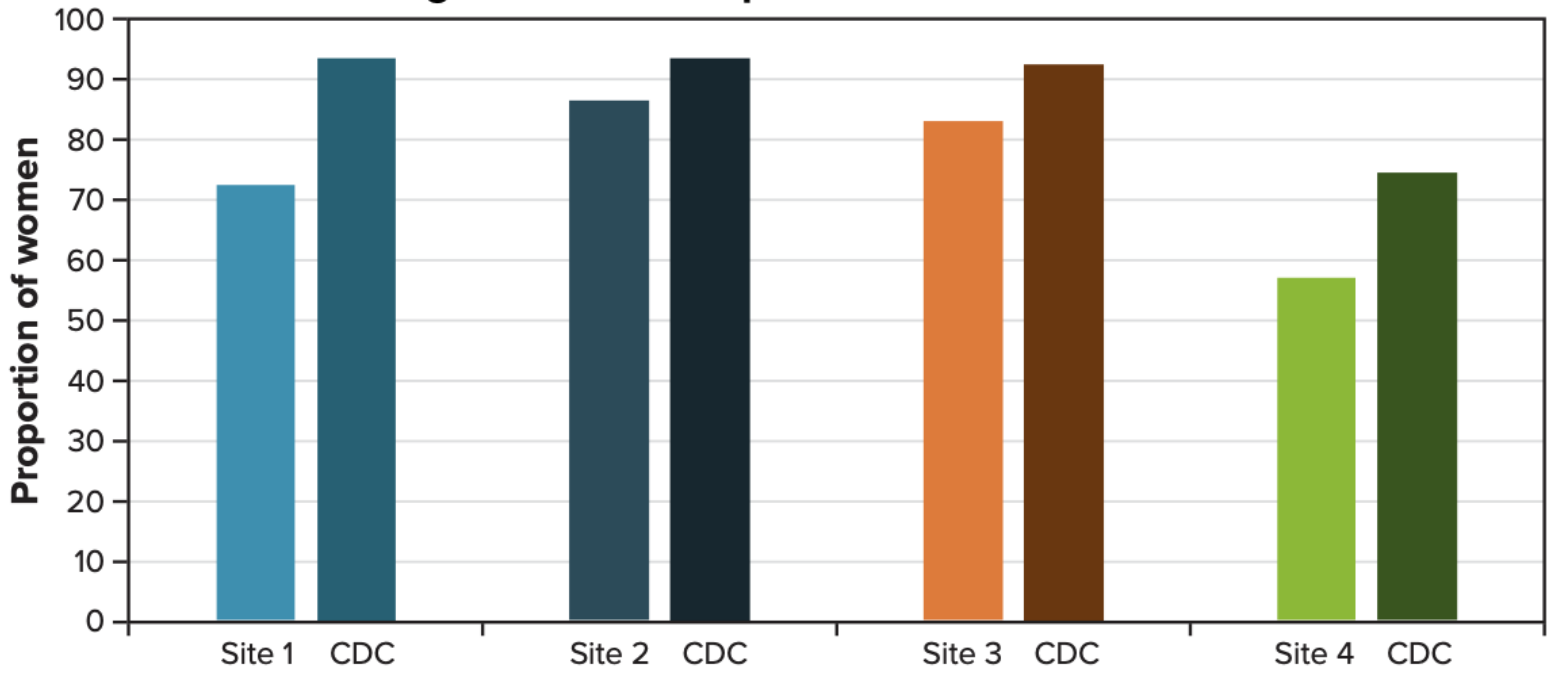


Source: Anthony MA, et al. Feasibility of Assessing Breastfeeding Status in Electronic Health Records. Pharmacoepidemiol Drug Saf. 2018;27(S2):122. Anthony MA, et al. Identification and validation of uterine perforation, intrauterine device expulsion, and breastfeeding in four health care systems with electronic health records. Clin Epi 2019;11:635–643.

Feasible to Assess Breastfeeding

Proportions at each site classified as breastfeeding were generally aligned with state-specific breastfeeding data

Figure 5. Proportion Breastfeeding at Study Sites Versus Proportion Ever Breastfeeding in CDC State-Specific Data[1]



[1]Percentage of women breastfeeding, by state in the National Immunization Survey, 2011 births

Source: CDC. Breastfeeding Report Card: United States/2014. Available at: https://www.cdc.gov/breastfeeding/pdf/2014breastfeedingreportcard.pdf. Accessed June 5, 2018.

Source: Anthony MA, et al. Feasibility of Assessing Breastfeeding Status in Electronic Health Records. Pharmacoepidemiol Drug Saf. 2018;27(S2):122. Anthony MA, et al. Identification and validation of uterine perforation, intrauterine device expulsion, and breastfeeding in four health care systems with electronic health records. Clin Epi 2019;11:635–643.

Feasible to Assess Breastfeeding

There were substantial differences in how breastfeeding information was collected and stored across the databases (e.g., structured questionnaire, clinician notes)

Breastfeeding was determined to be feasible for use in these data for this research question

It is important to evaluate breastfeeding data within the specific data source prior to initiating a study

Example: Peri-prosthetic Joint Infection (PJI)

Infrequent, though severe, adverse event which occurs after joint replacement

PJI can be difficult to identify

Definitive diagnosis requires preoperative blood and synovial fluid tests followed by intraoperative and post-operative examinations of the joint and surrounding tissue

Hip and knee PJI have been considered similarly through development of consensus diagnostic criteria

Source: Abdelaziz H, et al. The 2018 International Consensus Meeting Minor Criteria for Chronic Hip and Knee Periprosthetic Joint Infection: Validation from a Single Center. JArthroplasty 2020. (35):2200-2203.

2018 International Consensus Meeting on PJI – “Minor Criteria”

Serum
- C-reactive protein > 10 mg/L

Synovial fluid
- White blood cell count ≥ 3000 cells/µL
- Polymorphonuclear leukocytes ≥ 70%
- Leukocyte esterase ≥ + +
- Alpha-defensin > 1.0

Intraoperative
- Positive culture ≥ 1
- Positive histology
- Intraoperative purulence

≥6 of ”minor”

Compared against ≥1 ”major criteria”:
- Sinus tract communication with joint
- At least 2 positive cultures for same pathogen

Source: Abdelaziz H, et al. The 2018 International Consensus Meeting Minor Criteria for Chronic Hip and Knee Periprosthetic Joint Infection: Validation from a Single Center. JArthroplasty 2020. (35):2200-2203.

Validation of PJI in Claims

Canada – 4 Tertiary Care Hospitals – Hip and Knee

Compared PJI diagnosis/ procedure codes to chart review using the Musculoskeletal Infection Society criteria (assessed all hip or knee arthroplasty readmissions) in 2010-2016

ICD-10 diagnoses and CCI procedures

US – Veterans Health Administration – Knee

Compared PJI diagnosis/ procedure codes to chart review using the 2013 International Consensus definition (assessed subset of those with PJI codes) in 2000-2020

ICD-9 and ICD-10 diagnoses, CPT procedures

Sources: Kandel CE et al. InfCont&HospEpi 2021(42):325-330; Weinstein et al. PDS 2021(30):1184-1191.

PJI Algorithms

Canada – 4 Tertiary Care Hospitals

Diagnosis codes alone

Diagnosis + procedure

13 algorithms total

US – Veterans Health Administration

Diagnosis + procedure

Diagnosis + procedure + x-ray

Diagnosis + procedure + x-ray + arthrocentesis, arthrotomy of knee, blood culture, or other microbiologic procedure

6 algorithms total

Sources: Kandel CE et al. InfCont&HospEpi 2021(42):325-330; Weinstein et al. PDS 2021(30):1184-1191.

Validation of PJI Algorithm in Claims

	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Canada Hip+Knee Dx only	88% (85%-92%)	100% (100%-100%)	78% (74%-82%)	100% (100%-100%)
Canada Hip+Knee Dx + Proc	92% (88%-94%)	100% (100%-100%)	**81% (77%-85%)**	100% (100%-100%)
US Knee ICD-10 Dx + Proc			60% (48%-71%)	
US Knee (Best) ICD-10 DX + Proc + x-ray + ath/micro			**85% (75%-92%)**	

Source: Abdelaziz H, et al. The 2018 International Consensus Meeting Minor Criteria for Chronic Hip and Knee Periprosthetic Joint Infection: Validation from a Single Center. JArthroplasty 2020. (35):2200-2203.

Can we address uncertainty and bias?

Uncertainty

Knowledge of how data come to be in data source

Enough patients available

Bias

Critical variables (or suitable proxies) are available in data source

Data elements sufficiently define concepts

Limited misclassification

Comparability of treatment and comparator. Demonstration that exposure-outcome assessment is not being done before study is conducted. Results robust to design and analytic choices. Results robust to unmeasured confounding and bias.

Source: US FDA. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. 2024.

Thank you!

Contact: marybeth@medtechepi.com