Evaluating the Feasibility of Claims Databases for Emulating Interchangeability and Switching Trials: A Multi-Therapeutic Evaluation of Originator Biologics and Biosimilars

Patrick Pithua[1], Djeneba Audrey Djibo[2], Terese DeFor[3], Scott Myers[4], ***Catherine M. Lockhart[1]***

1. Biologics and Biosimilars Collective Intelligence Consortium, 2. CVS Healthspire Life Sciences Solutions, 3. HealthPartners Institute, 4. PearlDiver Technologies

BBCIC Biologics & Biosimilars Collective Intelligence Consortium

SA45

BACKGROUND

- There is an opportunity to utilize real-world data (RWD) to support U.S. Food and Drug Administration (FDA) approvals of biosimilars, and particularly to improve efficiency of assessment for interchangeable biosimilars, which is a unique regulatory designation in the U.S.
- For RWD to support regulatory approvals it must meet high data quality and completeness standards, ensuring the data's reliability and relevance for meeting rigorous regulatory requirements.
- Different RWD sources offer different characteristics such as data type, population size, and geographic distribution that can influence selection of the data most suited for a research question.
- It is important to understand nuances of different databases to inform and optimize study designs.

OBJECTIVE

This study aimed to evaluate the feasibility of using three distinct healthcare databases to capture the inclusion/exclusion criteria and outcome measures of clinical trials intended to support interchangeability by assessing switching between originator biologics and their biosimilars.

METHODS

Three databases were included in this analysis representing different settings as described in **Table 1**. These databases were selected to capture nuances from different data types that may inform appropriate database selection for observational research. We identified eight clinical trials examining switching between reference biologics and biosimilars from published literature and the ClinicalTrials.gov registry of protocols (**Table 2**). All variables representing trial inclusion/exclusion criteria, demographics, interventions, and outcomes were listed for each trial, and grouped into eight general categories: demographics, treatment, diagnostics, laboratory, vitals, behavior, assessment, and procedures **(Table 3**).

Databases were assessed based on characteristics such as data type and population size (**Table 1)** and availability of variables relevant to the trials included for evaluation. To measure the completeness and accuracy of each database we calculated the Observed-to-Expected (O/E) ratio, where the observed data represented the variables available in each database, compared to a fictitious ideal database containing all relevant variables. We also assessed other features relevant for database selection including data type, population size, and geographic distribution.

RESULTS

Table 1. Database Characteristics

	Database A	Database B	Database C
Data Source	National commercial health plan	Regional integrated delivery network	National claims database
Database Size	>44 million individuals	>4 million individuals	>170 million individuals
Date Range	Jan 2008 – Feb 2024	Jan 2000 – Apr 2024	January 2010 - April 2023
Average Follow-Up Time	2 years	5 years	9 years
Characteristics	Administrative claims and enrollment data from a commercial health plan	Administrative claims and electronic health records from a regional health system and insurance plan	Administrative claims data from all payer types: commercial, Medicaid, Medicare,* employer, cash
Geographic Region	All 50 states and territories	Midwestern U.S.	All 50 states

Table 2. Clinical Trials Included in Database Assessment

Reference	Product	Disease
1	Insulin glargine	Type 1 Diabetes Mellitus
2	Adalimumab	Chronic Plaque Psoriasis
3	Ranibizumab	Neovascular Age-Related Macular Degeneration
4	Infliximab	Rheumatoid Arthritis
5	Ustekinumab	Moderate to Severe Plaque Psoriasis
6	Rituximab	Rheumatoid Arthritis
7	Adalimumab	Chronic Plaque Psoriasis
8	Bevacizumab	Metastatic Colorectal Cancer

DATABASE ASSESSMENT

DATABASE A Commercial Health Plan

- Strengths: Large national health plan; Race/ethnicity for over 50% of records; Some laboratory and disease assessment data available (e.g., HbA1c); Relatively large population size; Possible EHR linkage
- Limitations: Incomplete laboratory data; Incomplete or unavailable behavioral data (e.g., smoking status) and limited patient-reported data

DATABASE B Integrated Network

- Strengths: Detailed clinical data as expected from an integrated delivery network, especially in laboratory and other disease measures (e.g., best corrected visual acuity, DAS28 and other patient reported outcomes)
- Limitations: Small population limited to the Midwestern U.S. region

DATABASE C Claims Database

- Strengths: Extremely large population from all 50 states with long follow-up time due to the ability to track patients between different health plans or insurance types
- Limitations: No medical record linkage so limited to data from billing claims; no laboratory data or other disease measures

DISCUSSION

- Each database exhibited different characteristics that may influence database selection. For example:
  - Database A includes race/ethnicity data for over 50% of individuals, but incomplete laboratory data.
  - Database B includes linkage to electronic medical records with robust laboratory and clinical data, but the population size is small and limited geographically.
  - Database C includes an extremely large population that is geographically diverse with over 9 years of average follow-up, but only includes information found in billing claims.
- All databases captured complete data on procedures and treatment identified using common coding systems including the International Classification of Diseases – 9th and 10th Revisions (ICD-9; ICD-10) or Healthcare Common Procedure Coding System (HCPCS). **Figure 1**.
- All databases included most demographic information such as age and sex, but some lacked race/ethnicity or other social determinants of health measures. **Figure 1.**

Figure 1. Percent of Variables Captured Across Categories (O/E Ratio)

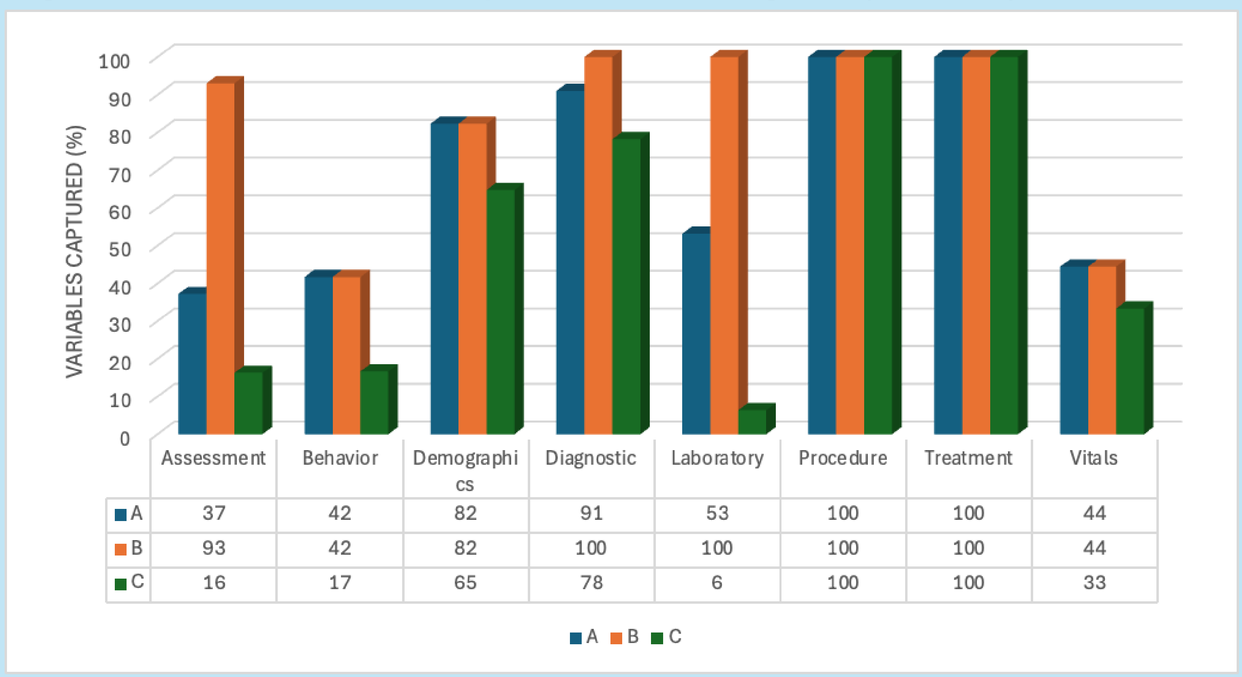

Table 3. Trial Variable Categories

Variable Category	Example Variables
Demographics	Age; Sex; Race and Ethnicity; Geographic Region
Diagnostics	Any disease diagnosis or condition
Procedures	Clinically administered medications; surgery; medically attended events
Treatment	Drug therapy – current or historical use
Laboratory	HbA1c, C-reactive protein, absolute neutrophil count; neutralizing antibody positivity
Vital Signs	Weight; Body Mass Index (BMI); QT Interval measurement
Behavior	Smoking status; birth control methods; pregnancy/lactation; adherence; duration of response
Assessment	Quality of life; ECOG status; best corrected visual acuity

CONCLUSION

The choice of a database should align with specific research needs, considering each database's strengths and limitations. While claims data offer valuable real-world insight, understanding data accuracy, completeness, and traceability is crucial for advancing its use for regulatory purposes.

REFERENCES

1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7065063/
2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9464749/
3. https://www.aaojournal.org/article/S0161-6420(21)00325-0/fulltext
4. https://clinicaltrials.gov/ct2/show/NCT02990806?term=interchangeability&draw=6&rank=47
5. https://clinicaltrials.gov/ct2/show/NCT04761627?term=interchangeability&draw=3&rank=12
6. https://clinicaltrials.gov/ct2/show/NCT01759030?term=interchangeability&draw=6&rank=45
7. https://clinicaltrials.gov/ct2/show/NCT05637515?term=interchangeability&draw=2&rank=3
8. https://clinicaltrials.gov/ct2/show/NCT05378867?term=interchangeability&draw=2&rank=4

ACKNOWLEDGEMENTS

This project is fully supported by a cooperative agreement (Award #1U01FD007757) with the Food and Drug Administration (FDA) of the U.S. Department of Health and Human Services (HHS). The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government.

Presented at ISPOR Europe 2024, November 17-24, 2024, Barcelona, Spain