MULTI-NATIONAL FEASIBILITY ASSESSMENT OF DERIVING STANDARD CANCER INDICATORS THROUGH OHDSI'S OMOP-CDM

Bruno Lima [1,2], Michael Schnell [3], Tapio Niemi [4], Maaike van Swieten [5], Harlinde De Schutter [6], Siri Laronningen [7], Espen Enerly [8], Jean Luc Bulliard [4], Evelyne Fournier [9], Peter Prinsen [5], Claudine Backes [1, 2]

1 - Registre National du Cancer du Luxembourg (RNC); 2 - Cancer Epidemiology and Prevention Group (EPICAN), Department of Precision Health (DoPH), Luxembourg Institute of Health (LIH); 3 - Data Integration and Analysis (DIA), Luxembourg Institute of Health (LIH); 4 - University Center for Primary Care and Public Health (Unisanté); 5 - Netherlands Comprehensive Cancer Organization, Utrecht, The Netherlands; 6 - Belgian Cancer Registry, Brussels, Belgium; 7 - Department of Registration, Cancer Registry of Norway, Norwegian Institute of Public Health; 8 - 6Department of Research, Cancer Registry of Norway, Norwegian Institute of Public Health; 9 - Geneva Tumor Registry, University of Geneva, Switzerland

BACKGROUND

Six Population-Based Cancer Registries (PBCRs) from five countries are assessing the feasibility of mapping their data structures to an harmonized health data format - the Observational Health Data Sciences and Informatics (OHDSI) Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) - to enhance interoperability, comparability and secondary use of health data. The OHDSI framework defines a standardized relational schema and controlled vocabularies, enabling the integration and secure sharing of heterogeneous health data.

AIM: To evaluate whether breast cancer indicators generated by six PBCRs after transformation into the OMOP-CDM are comparable to their previously published results based on the ICD-O-3 coding system.

METHODS

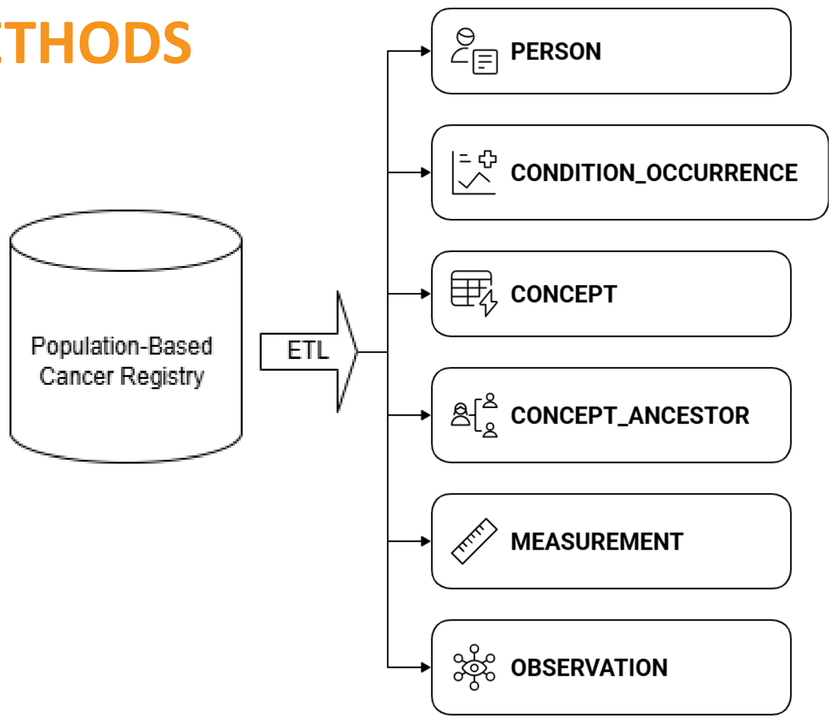

Figure 1. ETL Process, from PBCR to OMOP-CDM

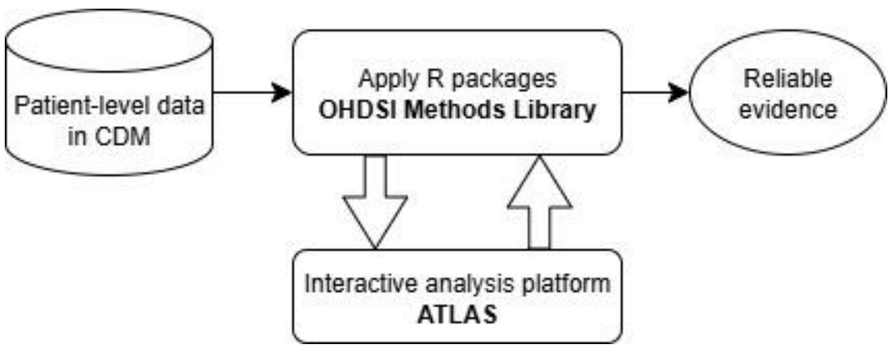

Figure 2. Data analysis implementation in the CDM

RESULTS

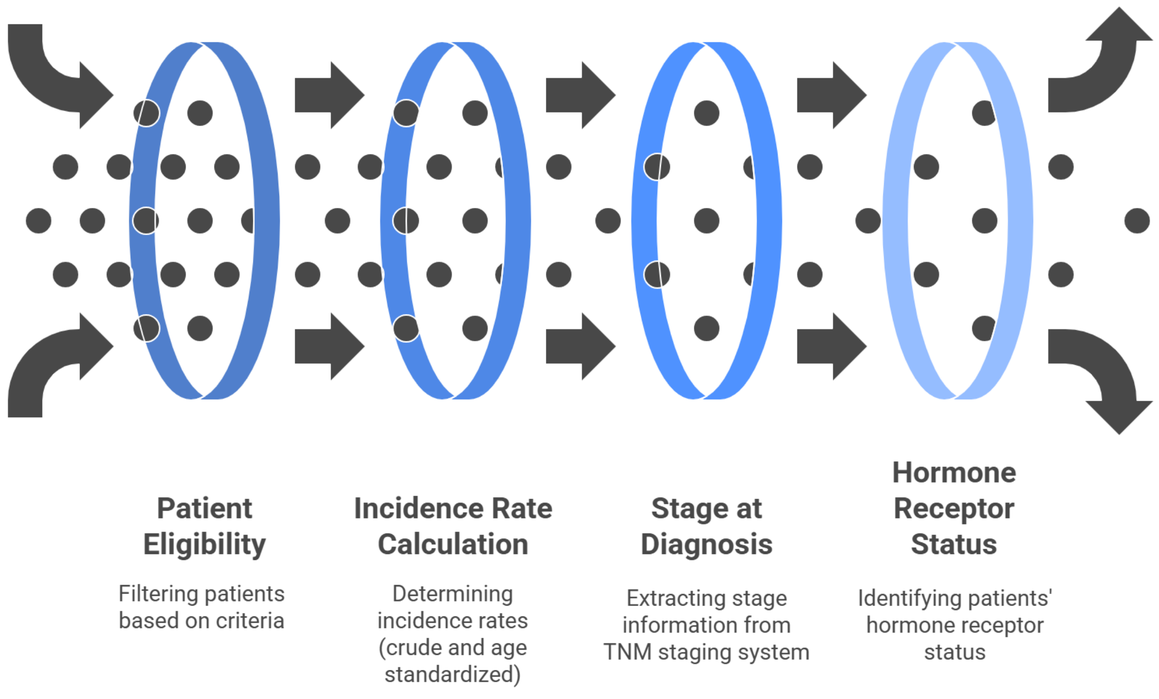

Figure 3. Data flow: cohort definition and calculated indicators

OHDSI's analytical tools and the R programming language were used to define a cohort of breast cancer patients. Using the {CohortMethod} package, generic connections to OMOP-CDM databases were established and an SQL script that could be run on the various PBCR databases involved in this study was created.

The patient cohort was defined as all cases of «Malignant tumor of breast» diagnosed in females in 2019. In addition to identifying the cohort, the common SQL script also identified stage at diagnosis (0 - IV) for each patient.

DISCUSSION

Sharing standardized, executable code across OMOP-CDM–compliant PBCRs improves reproducibility and comparability of results. However, differences in mapping conventions (e.g., hormone receptor status) persist, limiting the use of a single unified script across all registries. Establishing common cohort definitions and shared scripts is essential to enable federated analyses. This approach supports the goals of the European Health Data Space (EHDS) by promoting interoperability, privacy, and the secondary use of health data across borders. Further studies are needed to evaluate the feasibility, scalability, and sustainability of OMOP-CDM in PBCRs and its role in pan-European federated health data networks.

4-7 November 2025
Ege Palas Business Hotel
Izmir / Türkiye

Bridging Continents and Data:
Cancer Registries for Cancer Control

IACR 2025
İZMİR · TÜRKİYE

IACR · International Association of Cancer Registries