Oncovalue: Using Real-World Data for Real-World based Health Technology Assessment

Cancer Economics WG Session
Kevin A. Tittel, MSc
PhD Candidate – Netherlands Cancer Institute [NKI]

prof. dr. W.H. van Harten
prof. dr. V.P. Retèl
dr. N. Franzen

ONCOVALUE

Objective: To further unlock the full potential of real-world hospital data, generated in European cancer centres, for efficient use in health technology assessments [HTAs] of novel cancer treatments.

HUS Helsinki University Hospital

REGION Rigshospitalet

Netherlands Cancer Institute – Antoni van Leeuwenhoek

Rijnstate

Istituto Romagnolo per lo Studio dei Tumori Dino Amadori

IPO Porto – Instituto Português de Oncologia do Porto FG, EPE

Siemens Healthineers

BC Platforms

IQVIA™

elevate

TTOPSTART

CiaoTech

Real-World Hospital Data

Available?

Structured?

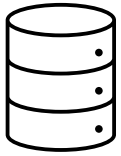
High-quality?

Standardised?

Interoperable?

Linkage?

Real-World Hospital Data



Available?

Structured?

High-quality?

Hospitalisation due to adverse events

Individual drug exposure

Adverse events

Dose reduction

Concurrent therapy

Reason for treatment discontinuation

Patient-reported outcomes

Real-world costs

Medical history

Reasoning behind clinical decisions

Previous anticancer treatment

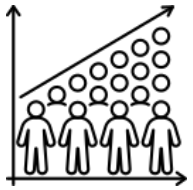
Treatment for adverse events

Comorbidities

Standardised?

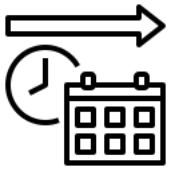
Interoperable?

Linkage?



Generalise to the normal population

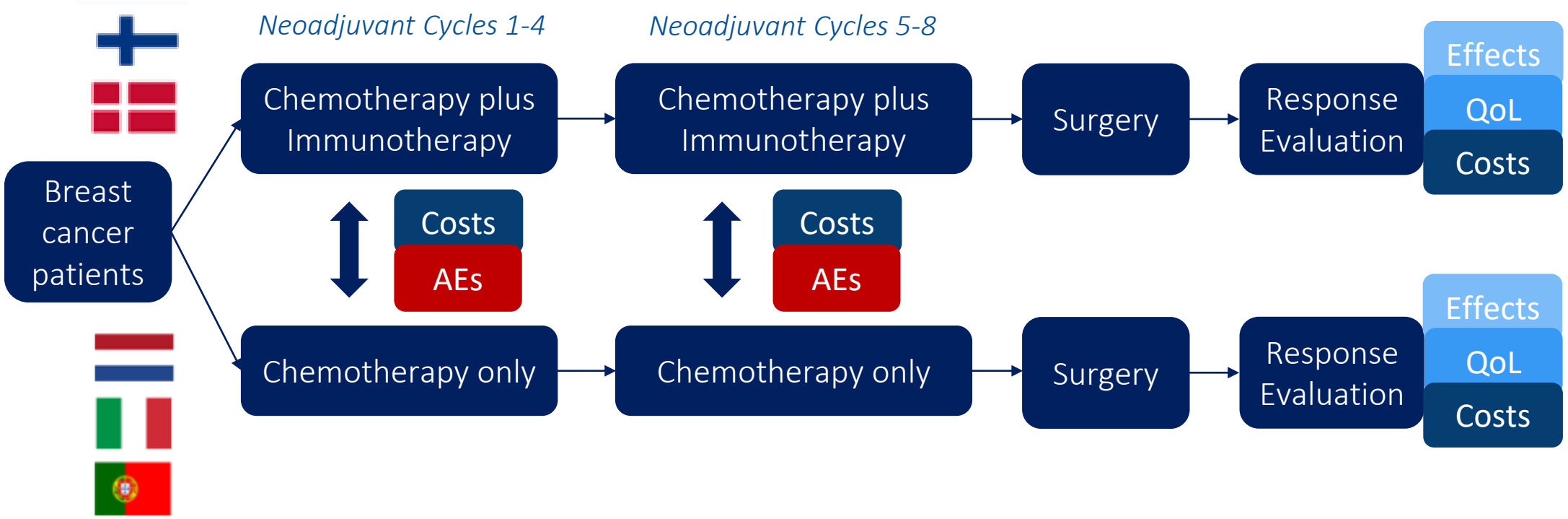
Results from clinical practice



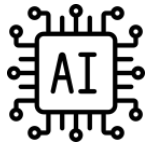
Longer follow-up period

Health Technology Assessment

Illustrative example



ONCOVALUE output



- NLP-tool to structure free text
- AI-tool to quantify tumour growth

- Federated analysis infrastructure

- Mapping data to the OMOP Common Data Model

- Hospital RWD-based HTA framework

Main questions

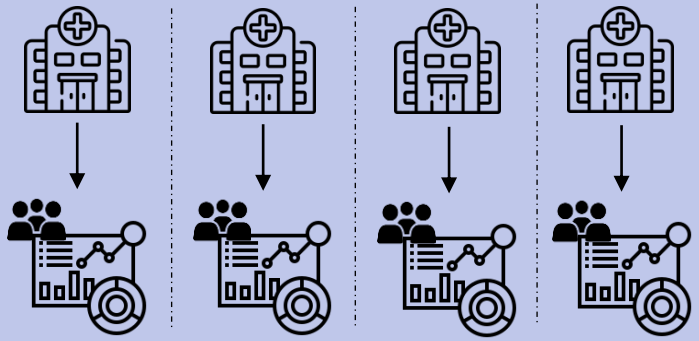
1. What is the current status of real-world hospital data integration in Europe?
2. What challenges and issues arise when working with real-world hospital data?
3. What improvements are needed to advance the use of hospital data for HTAs?

NETHERLANDS CANCER INSTITUTE
ANTONI VAN LEEUWENHOEK

ONCOVALUE feasibility study

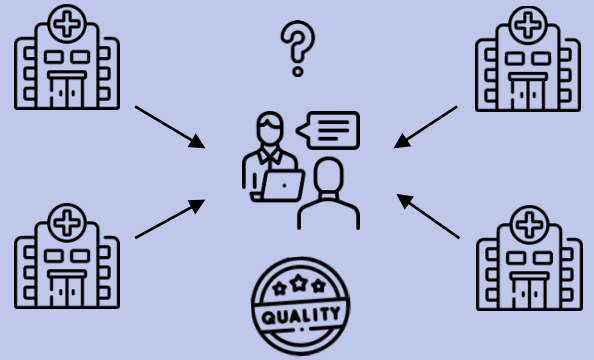
Objective: To investigate the quality and suitability of real-world hospital data to conduct an HTA of neoadjuvant therapies for early breast cancer patients treated between 2015 and 2023.

Retrospective Cohort Study



Monocenter study
separately in each clinical center

Semi-Structured Interviews



Multicenter study
interviews with each clinical center

Retrospective cohort study data

1. Baseline Characteristics

Demographics

Clinical (General)

- Performance status
- Anthropometic measures
- Comorbidities
- Medication use

Cancer-specific

- Risk factors
- Hereditary predisposition
- Previous anticancer treatment
- Tumour (sub)type
- Pathology-specific criteria
- Staging at diagnosis
- Biomarker analyses
- Metastases

2. Comparators

Neoadjuvant systemic therapy

- Types of medication given
- Dosage amount/form
- Length/number of cycles
- Start/end dates

Surgery

- Type of breast cancer surgery
- Date of surgery

Radiotherapy

- Type of radiotherapy
- Start/end dates

General considerations

- Treatment discontinuation

3. Outcomes & Endpoints

Tumour-related and survival

- Partial/complete response
- Stable/progressed disease
- Local/regional/metastatic recurrence
- Death

Treatment-related and safety

- Adverse events

Patient-reported outcomes

- Quality of Life questionnaires

Costs

- Diagnostic tests
- Follow-up related checkups
- Procedures
- Medications
- Hospitalization

Data categories and domains selected according to Supplementary Table S8 of ESMO Guidance for Reporting Oncology real-World evidence (GROW), 2023.

Semi-structured interviews

Participating centers

HUS Helsinki University Hospital

REGION Rigshospitalet

Netherlands Cancer Institute – Antoni van Leeuwenhoek

Rijnstate

Istituto Romagnolo per lo Studio dei Tumori Dino Amadori

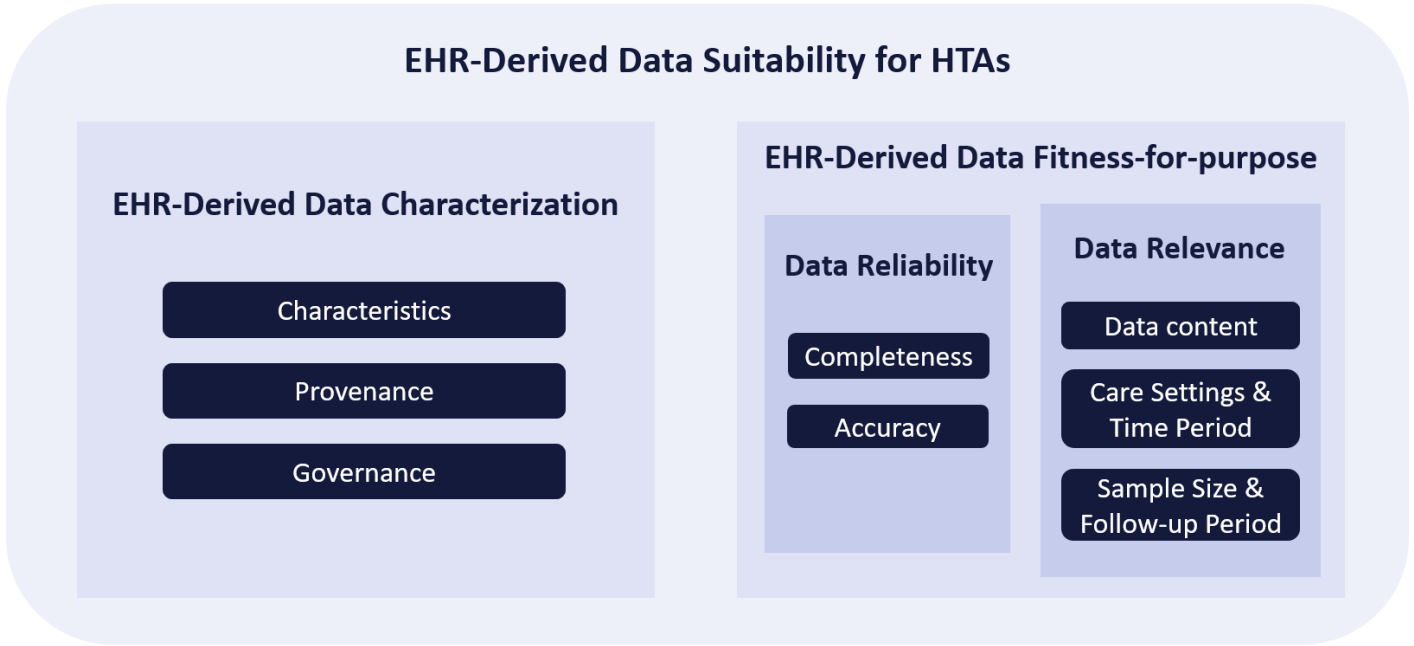
IPO Porto – Instituto Português de Oncologia do Porto FG, EPE

Methods

The ISPOR EHR Data Suitability Framework and SUITED Checklist[1].

Targeted literature review for data integration concepts and definitions.

Flowcharts of the hospital data architecture and pipeline across centres.



[1] Using Data from Electronic Health Records for Health Technology Assessment: An ISPOR Emerging Good Practices Task Force Report. *Draft version November 2023.*

Preliminary results

Clinical information systems

Clinical information systems of ONCOVALUE centres differ considerably in terms of architecture and maturity.

- Most centres (*n* = 5) employ data warehouses and repositories containing structured and modelled data.
- Some centres *(n* = 4) (plan to) build data lake(house)s where all (un)structured data are integrated with improved quality and data linkage.
- Some centres (*n* = 4) build an OMOP database and (*n* = 1) federated data infrastructure.

Heterogeneity in system architecture and maturity affects the availability, quality and fitness-for-purpose of data.

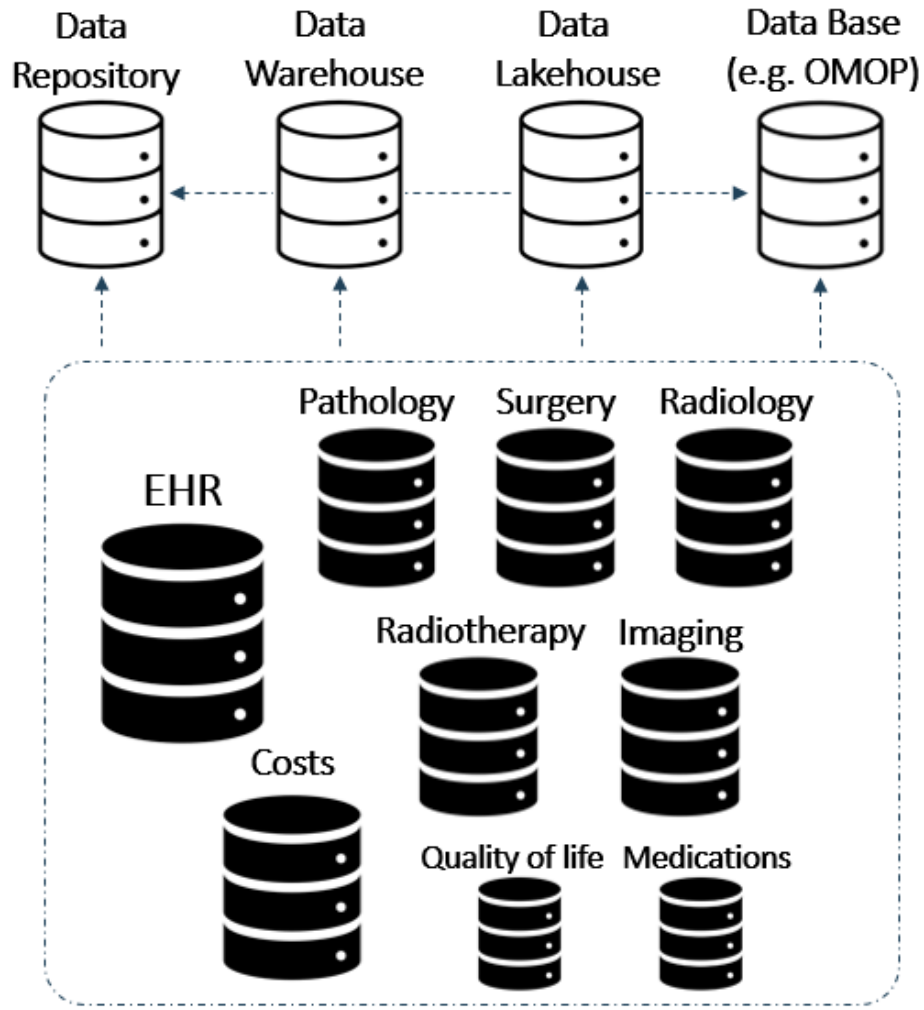
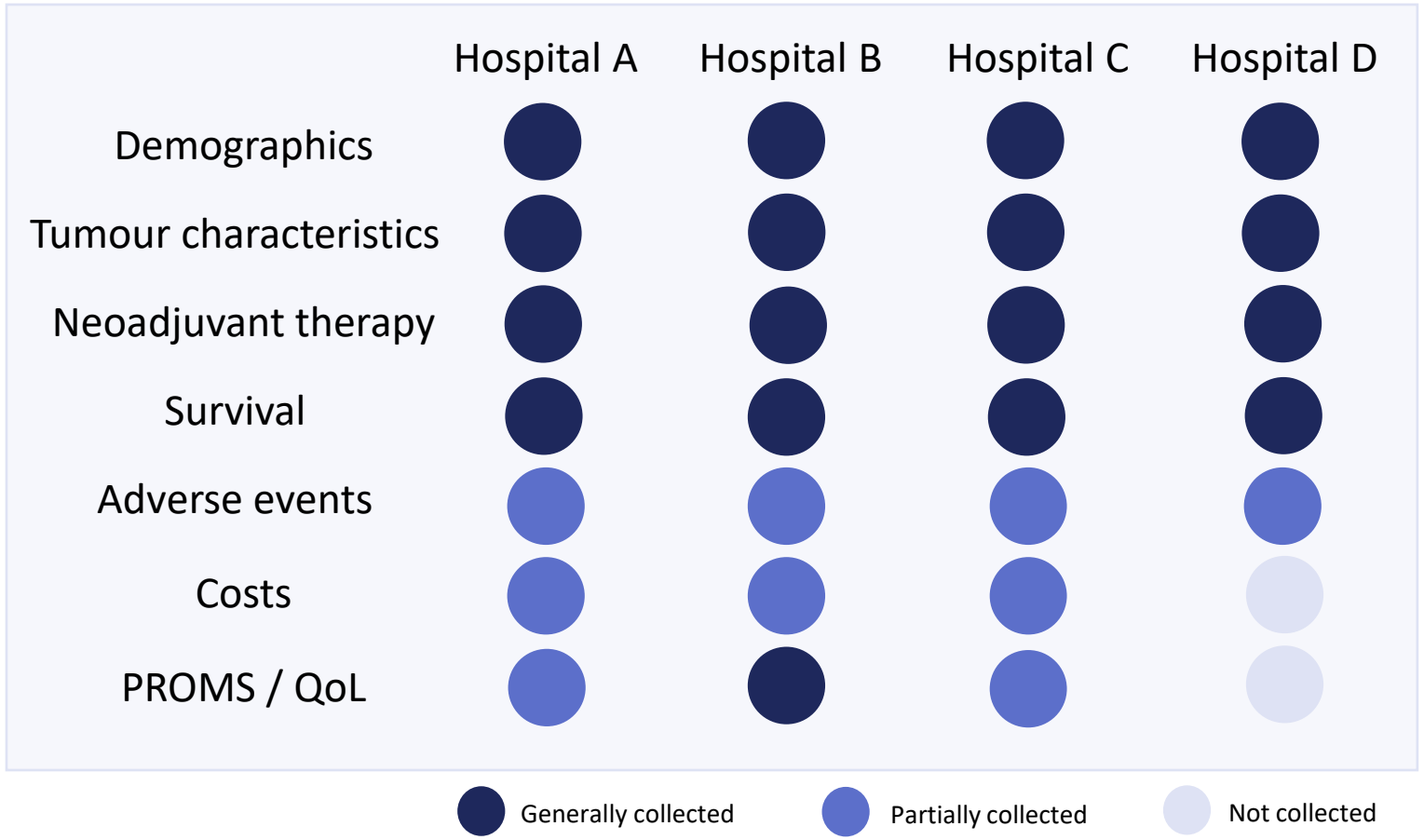


Figure. Example of clinical information system architecture

Preliminary results

Data availability

Table. Breast cancer data availability for a selection of data blocks.



Data availability varies over time, patient groups, and data domains.

- Lack of data for earlier years and for patients coming from or moving to other care centres.
- Only proxies or partial collection for certain variables (e.g. severe adverse events).
- Quality of life and real-world costs not collected or accessible in multiple centres.

Multiple centres are reforming their data collection (e.g. adverse event forms, QoL questionnaires).

Preliminary results

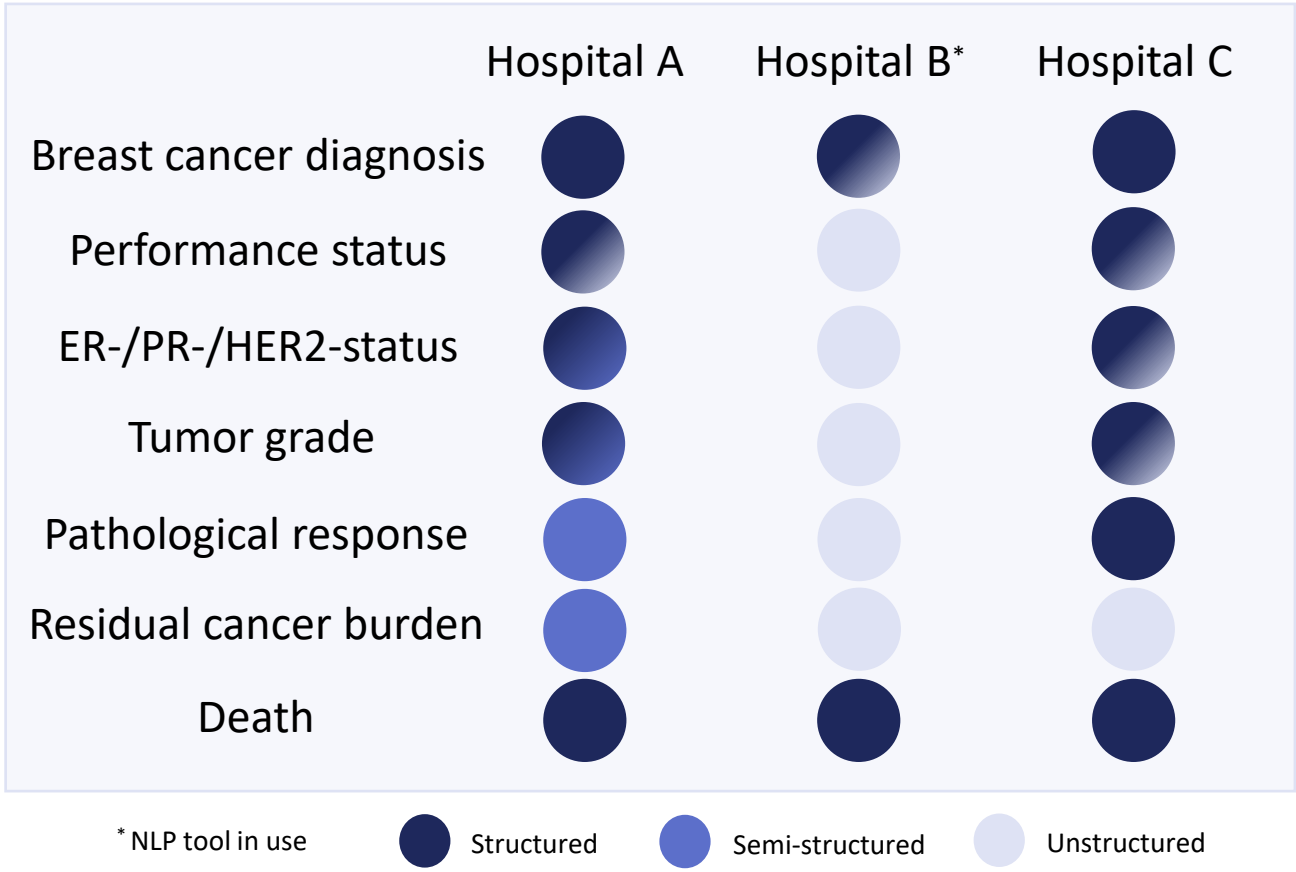
Data structure

Often multiple sources (e.g. EHR fields, forms, reports, notes) are needed to derive one variable of interest.

- Various ways in which clinicians denote variable outcomes complicate data processing.
- Data managers and researchers do manual processing using complex logic and interpretation.

Next generation NLP tools seek to derive data from heterogeneous clinician notes and reports.

Table. Breast cancer data structure for a selection of variables.



Preliminary results

Data processing steps

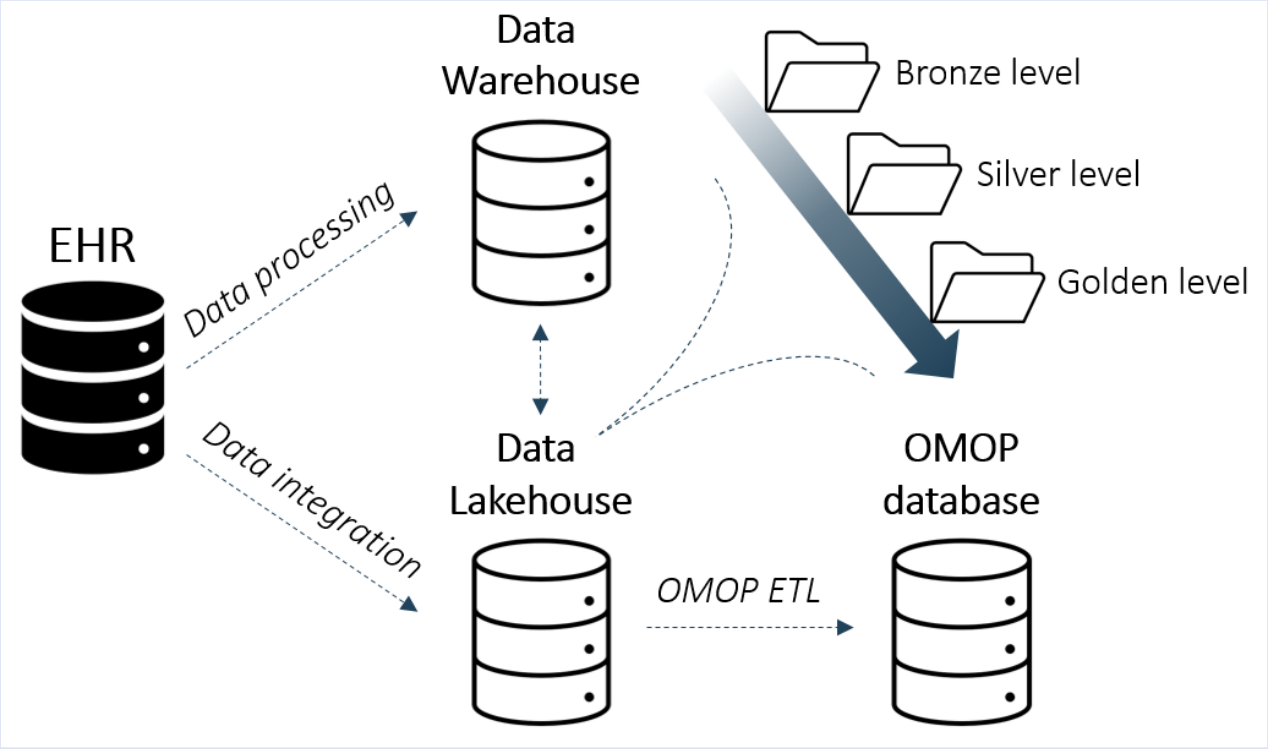


Figure. Example of data manipulations in ONCOVALUE centres.

From data collection by clinicians to data analysis by researchers, many processing steps manipulate the data.

- All centres still do manual processing, specifically for variables in need of interpretation (e.g. TNM status).
- The majority of data is modelled and integrated using automatic packages and SQL-based ETL processes.
- AI-tools (e.g. CTcue) are in use to process and retrieve data.

The implemented processing steps affect the availability, quality and fitness-for-purpose of the data.

Preliminary results

Data quality checks

From data entry to data analysis, all centres implement varying quality assurance procedures.

- Common quality checks at data entry:
  - Invalid data warnings
  - Incomplete entry warnings
  - Training
  - Blinded review
  - Regular quality reporting
- Extensive quality checks along data integration and OMOP CDM processes.
- Data lake(house)s consist of varying levels with different degrees of quality.

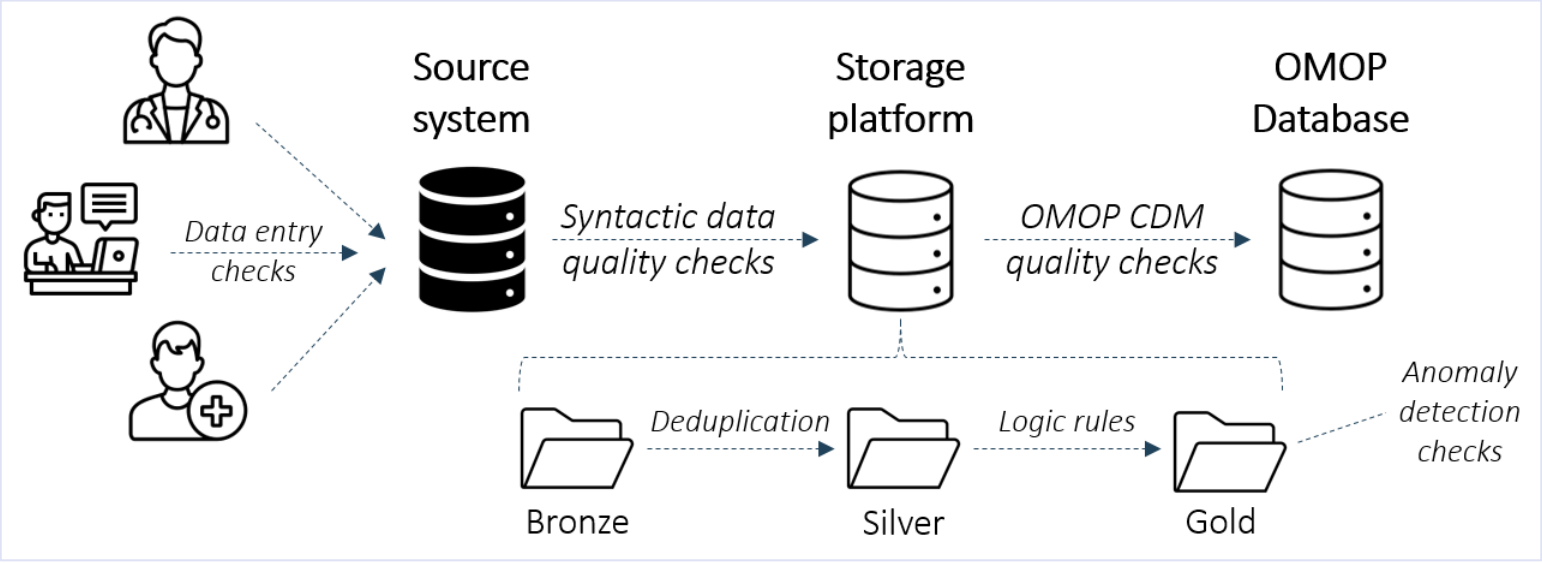


Figure. Example of data quality checks in ONCOVALUE centres.

Preliminary results

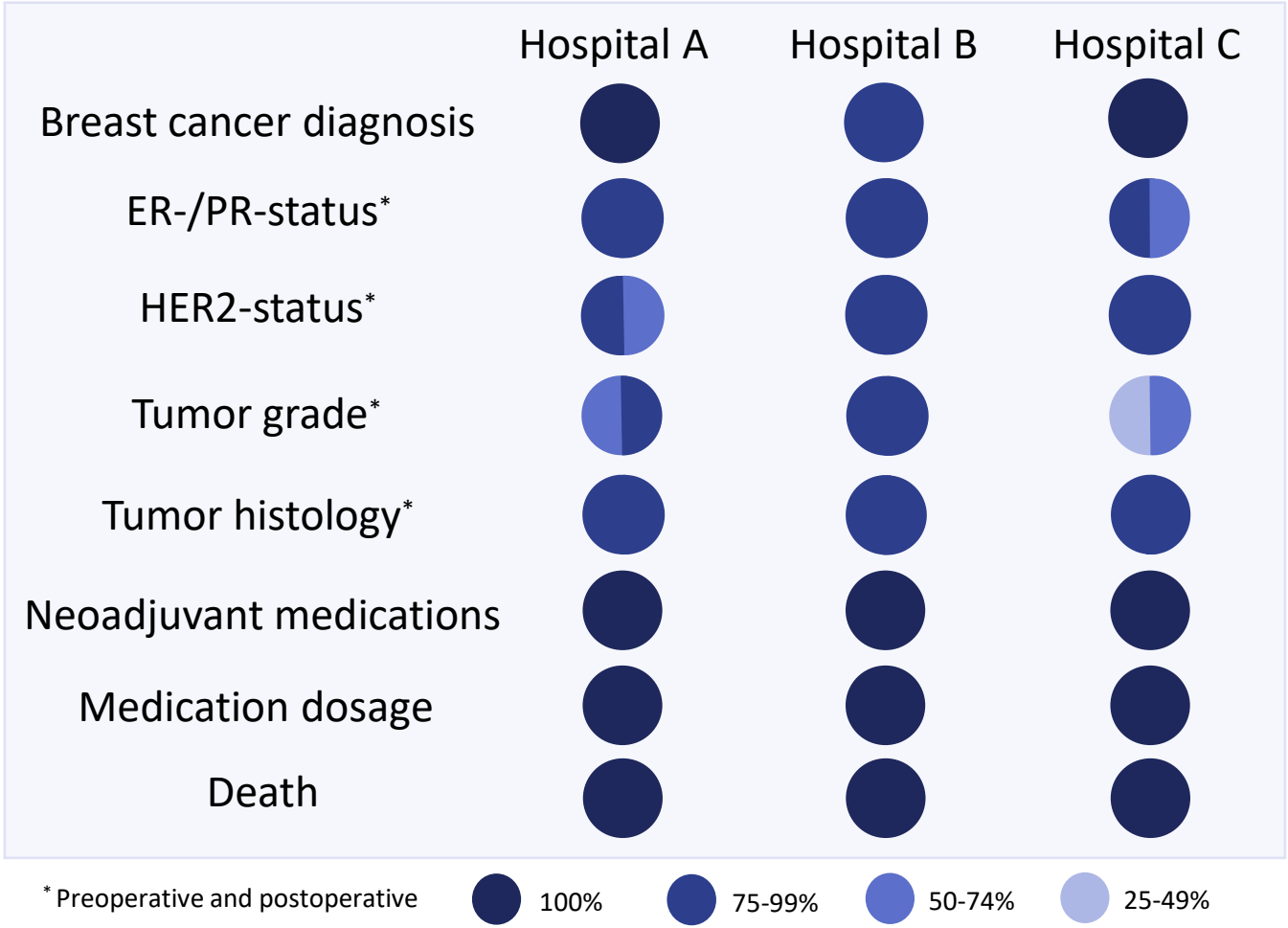
Data quality

Data completeness differs per data domain and data source structure.

- Data quality will further be assessed by
  - Data conformance
  - Data plausibility
  - Data uniqueness
  - Data persistence
  - Data accuracy
- Methods are needed to account for bias and uncertainties in the data.

Data quality reporting is paramount to generating trustworthy real-world evidence for HTA purposes.

Table. Breast cancer data completeness for a selection of variables.



Preliminary results

OMOP CDM mapping

Mapping data to OMOP and the Oncology Module extension has started in 4 centres.

Complex albeit important for reliable and efficient multi-centre real-world data studies.

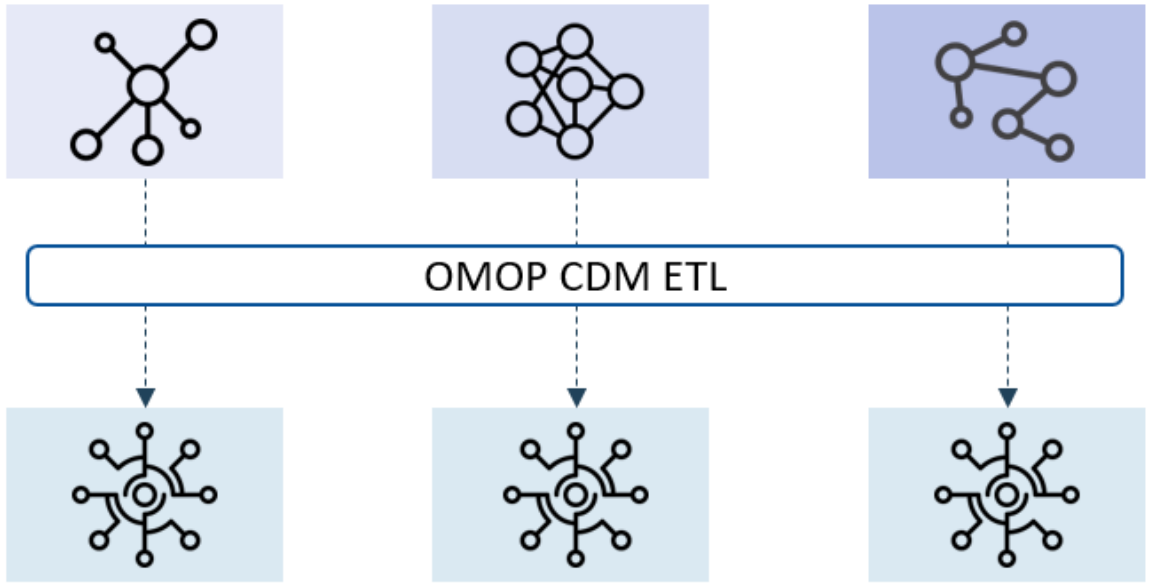


Figure. Data modelling to the OMOP CDM.

Preliminary results

Federated data analytics

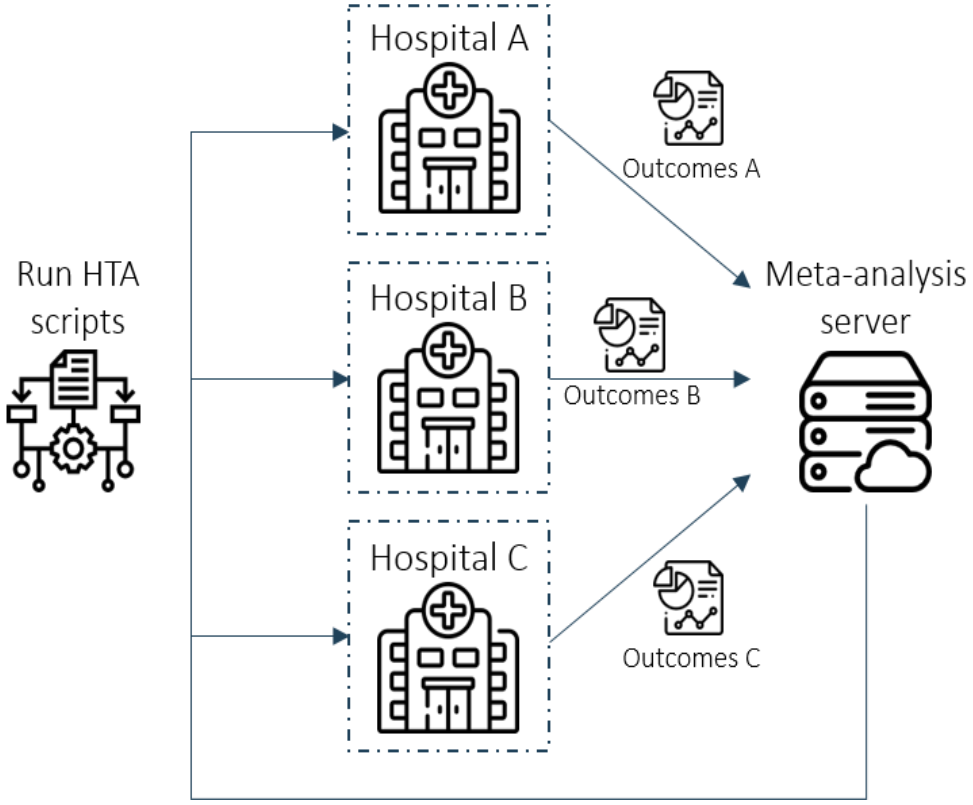


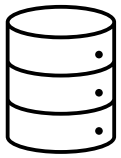
Figure. Federated data analytics infrastructure.

Federated infrastructure is tested and implemented in a first centre.

- Enables multi-centre analyses without sharing or centralizing patient-level to preserve patient privacy.
- Centres only share aggregate results to a meta-analyses server.

Relevant for multi-centre hospital RWD-based future-proof HTA

Discussion



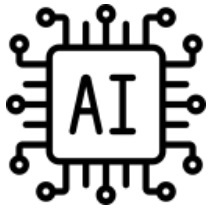
Available?

Structured?

High-quality?

Transparent data documentation is needed to understand the quality and suitability of real-world hospital data for HTAs.

Hospital system architecture and maturity



Data processing and extraction using next generation AI

Standardising data to a Common Data Model (e.g. OMOP)

Standardised?

Interoperable?

Linkage?

Thank you for listening!

Cancer Economics WG Session
Kevin A. Tittel, MSc
PhD Candidate – Netherlands Cancer Institute [NKI]

prof. dr. W.H. van Harten
prof. dr. V.P. Retèl
dr. N. Franzen