

Objectives

- Describe and implement new real-world data (RWD) quality frameworks or initiatives to improve the **relevance**, **reliability**, and **external validity** of oncology EHR-derived databases for research.
- Facilitate availability and use of robust and reliable RWD and real-world evidence (RWE) in oncology.

Methods

- The RWD quality dimensions in the FDA QCARD¹ initiative were implemented using Ontada's ON.Genuity RWD platform, which integrates EHR data from ~500 US community oncology clinics with claims and external mortality data.

Results

- Common challenges to RWD were mitigated by integrating data from multiple sources (**Table 1**).
- The RWD quality dimensions defined in the FDA QCARD¹ initiative were implemented (**Figure 1-3**).
- Assessment of relevance revealed consistent high availability of 250+ standardized variables across 20+ clinical domains for 500K+ patients over the past 10 years. Data completeness was improved by ON.Notate abstracted data and natural language processing (NLP) (**Figure 4**).
- Mortality data demonstrated high consistency with the external National Death Index (NDI) data source ($p=0.9$, with identical median overall survival (OS), **Figure 5**).

Table 1. Data Sources in Ontada's ON.Genuity RWD

Data source	Structured iKnowMed EHR Data	ON.Notate chart-abstraction	NLP	Non-EHR Data
Example Variables	Gender, race, ethnicity, vital signs, labs, medications, performance status, stage	Oral medication stop dates, surgery start dates and outcomes, disease progression status, histology and hospitalizations	Claims data (amount paid, payer type), supplemental mortality data	
Processing Method	Ingestion, standardization, reconciliation	Technology-enabled human abstraction	NLP model building and data validation	Linked data at patient level
Highlights				
<ul style="list-style-type: none"> 250+ variables were standardized following mCODE² initiative. EHR data were linked based on unique and accurate ID within iKnowMed. Unstructured data were available for all patients. 				

Figure 1. Relevance

Availability (data elements)	Feasibility (representative patients)
<ul style="list-style-type: none"> 250+ standardized variables across 50+ cancers Oncology-specific EHR data 	<ul style="list-style-type: none"> Largest network of community oncologists in the US 1.5M+ patients from ~500 clinics

Figure 2. Reliability

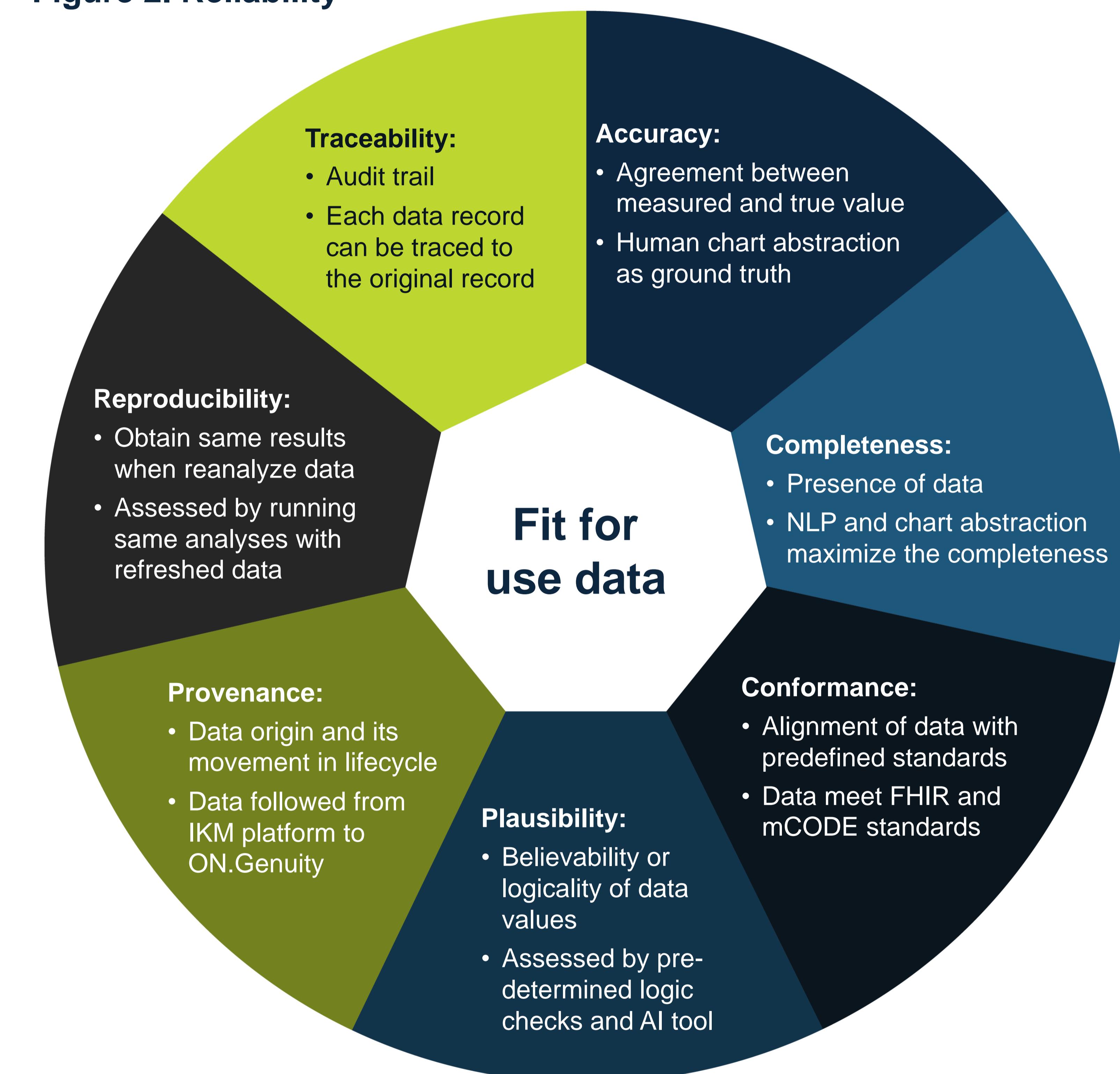
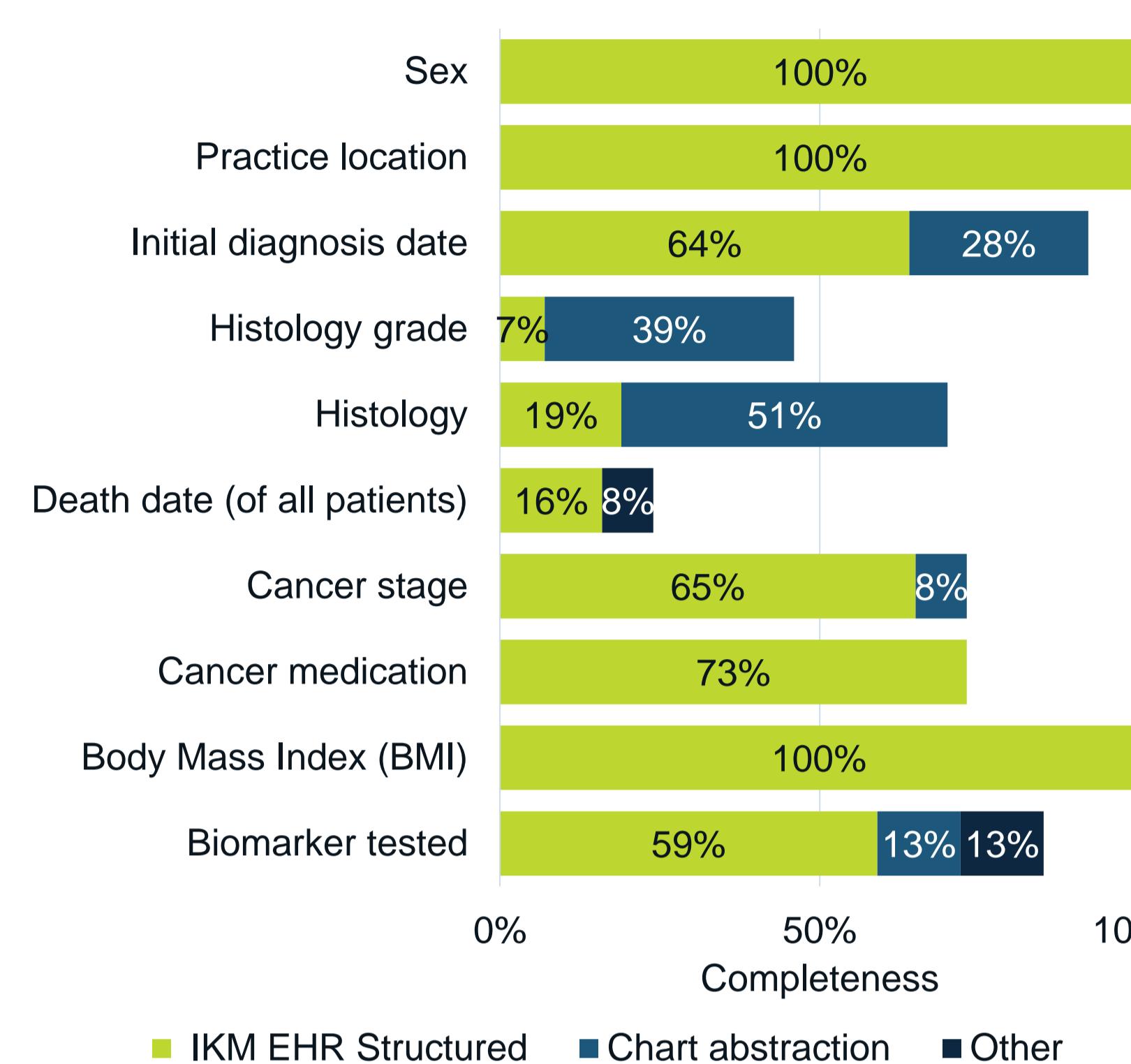


Figure 3. External Validity

Generalizability (inference to a broader population)	Replicability (confirm findings in different populations)	Transparency (clear communication and description)
<ul style="list-style-type: none"> Generalizable to community oncology setting in the US Exercise caution when generalizing outside community oncology setting 	<ul style="list-style-type: none"> Validation against National Death Index mortality data Compare with published clinical trial and National Cancer Institute data 	<ul style="list-style-type: none"> Detailed documentation of data derivations, study design, methods and analyses

Figure 4. Incremental Data Completeness by Source



Discussion

To our knowledge, we are the first organization to implement the FDA QCARD initiative on fit-for-use real-world data in oncology.

Ontada's ON.Genuity EHR database was standardized using FHIR and mCODE², with increased relevance, reliability and external validity by linking to external structured data at a patient level, human chart abstraction, and leveraging NLP technology.

The quality and fit-for-use of RWD should be carefully evaluated for each study, including a detailed assessment of relevance, reliability and external validity in relation to the research objectives.

Conclusions

- The newly published QCARD initiative on data quality represents a significant advancement in EHR-based oncology research, ensuring use of relevant, reliable and externally valid data.
- The implementation of the QCARD initiative has the potential to substantially enhance RWE generated from RWD and is expected to improve informed decision-making in clinical practice and policy.
- Our research serves as a model for other therapeutic areas and underscores the importance of rigorous data quality standards in real-world research.

References

1. Rivera DR, Eckert JC, Rodriguez-Watson C, et al. The Oncology QCARD Initiative: Fostering efficient evaluation of initial real-world data proposals. *Pharmacoepidemiol Drug Saf*. 2024 Nov;33(11):e5818.
2. Osterman TJ, Terry M, Miller RS. Improving Cancer Data Interoperability: The Promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. *JCO Clin Cancer Inform*. 2020 Oct;4:993-1001.

Figure 5. External Validity of OS for Metastatic Pancreatic Cancer

