RWD172

Enhancing Real-World Data (RWD) Quality: Implementation of Regulatory Data Quality Framework within a Large US Oncology Electronic Health Record (EHR)-Derived Database

Zhaohui Su, PhD; Jessica K. Paulus, ScD; Kat Aguilar, MPH, PhD(c); Aniketh Talwai, MBA; Saamir Pasha, MPH; Jay Hugalavalli; Jennifer Frytak, PhD; Amy K. O'Sullivan, PhD; Janet L. Espírito, PharmD; Ontada, Boston, MA, USA

PT3

Objectives

- The FDA-led QCARD Initiative and European Medicines Agency (EMA)'s Data Quality Framework (DQF) provide structured guidance to advance RWD quality.
- This study aimed to present the approaches and results of applying QCARD and DQF to the ON.Genuity platform, which integrates iKnowMed EHR and supplementary data sources.

Methods

- QCARD and DQF principles were applied by standardizing structured and unstructured data following Fast Healthcare Interoperability Resources (FHIR) and Minimal Common Oncology Data Elements (mCODE).
- Metadata and lineage documentation were used to support traceability and transparency.
- External validation was performed using the US National Death Index (NDI).
- Automated checks for reliability and fit-for-use were implemented in a cohort of multiple myeloma (MM) patients diagnosed between 2018 and 2024.

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	• 250+ variables were standardized following mCODE[2] 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)
• 250+ standardized variables across ~70 cancers • Oncology-specific EHR data	• Largest network of community oncologists in the US • 4M+ patients from ~700 clinics

Presented at ISPOR 2026, May 17-20, 2026, Philadelphia, PA

Results

- Improved data completeness with data from multiple sources (**Table 1, Figure 4**).
- The RWD quality dimensions defined in the FDA QCARD[1] initiative and EMA DQF[2] were implemented (**Figure 1-3**).
- Assessment of relevance revealed consistent high availability of 250+ standardized variables across 20+ clinical domains for 4M+ patients over the past 10 years.
- 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**).
- Standard edit checks identified data discrepancies affecting <1% of patients.

Figure 2. Reliability

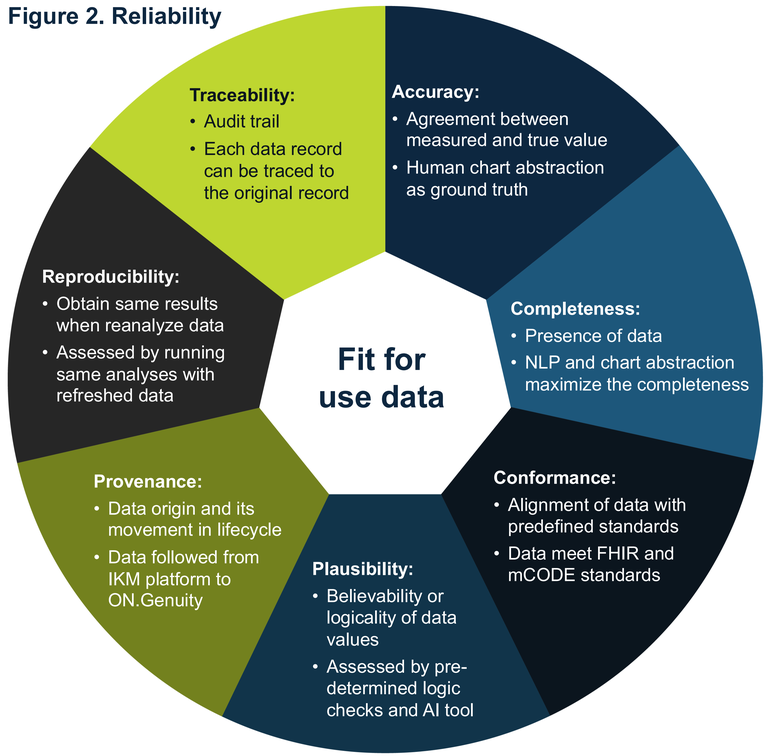


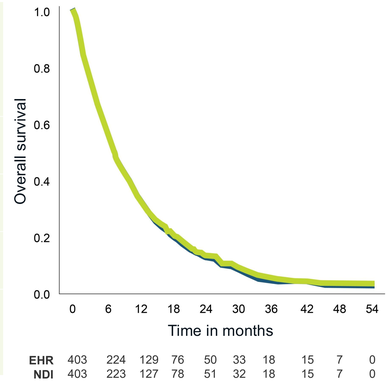
Figure 3. External Validity

Generalizability (inference to a broader population)	**Replicability** (confirm findings in different populations)	**Transparency** (clear communication and description)
• Generalizable to community oncology setting in the US • Exercise caution when generalizing outside community oncology setting	• Validation against National Death Index mortality data • Compare with published clinical trial and National Cancer Institute data	• Detailed documentation of data derivations, study design, methods and analyses

Figure 4. Incremental Data Completeness by Source

Data Domain	Complete Rate	QCARD/EMA DQF Assessment
Demographic	100%	✅ Excellent
Medications	92%	✅ Excellent
Lab Data	98%	✅ Excellent
Performance Status	80%	✅ Good

Figure 5. External Validity of OS for Metastatic Pancreatic Cancer



Discussion

To our knowledge, we are the first organization to implement the FDA QCARD and EMA DQF initiative on fit-for-use RWD in oncology.

Ontada's ON.Genuity EHR database was standardized using FHIR and mCODE[3], 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

- Implementation of the QCARD and DQF within the ON.Genuity platform introduces a process for generating high-quality RWD in oncology.
- This research contributes to streamlining regulatory submissions, promoting the use of RWD in in evidence-based decision-making.
- 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. Data Quality Framework for EU Medicines Regulation. European Medicines Agency; first published October 10, 2022; last updated December 12, 2023.
3. 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.