Clinically Relevant Edit Checks Improve Real-World Data Quality

DiLullo S, Espirito J, Haydon W, Lewis C, O'Brien M, Paulus JK, Rembert D, Spillson C, Tedesco B, Spark S, Stepp S, Su Z, Uranga J, The US Oncology Network, Ontada, Boston, MA, USA.

RWD158

Background

Ensuring real-world data (RWD) quality involves systematically reviewing the data via a combination of traditional edit checks and clinically relevant edit checks. Together, these complementary approaches improve data consistency, clinical plausibility, and overall reliability of RWD for research and regulatory use.

Definitions:

- **Edit Check:** A programmed or manual review rule used to identify missing, inconsistent, or clinically implausible data within real-world data sources to improve accuracy and reliability.

- **Traditional Standard Edit Check (TSEC)**: logic based, date sequencing.
  Ex: "Biomarker collection date is after the result date".

  Ex: "Stage at diagnosis is missing".

- **Clinically relevant edit check (CREC)**: additional plausibility checks on clinical variables.
  Ex: Resectability status captured as unresectable or not documented, with surgery also captured. (Resectability status is inconsistently documented in source data, and requires additional quality control via edit checks)

  Ex: Medication discrepancies: Total daily doses did not align with the tablet strengths, quantities, frequencies, durations, and cycles captured.

Methods

Study Design

Three Retrospective Studies

Chart abstracted data from 2024-2025 used to assess data accuracy

Quality Control

CRECs added to TSECs

CRECs applied using clinical expertise for specific tumor types and study requirements

Assessment

Queries and Corrections

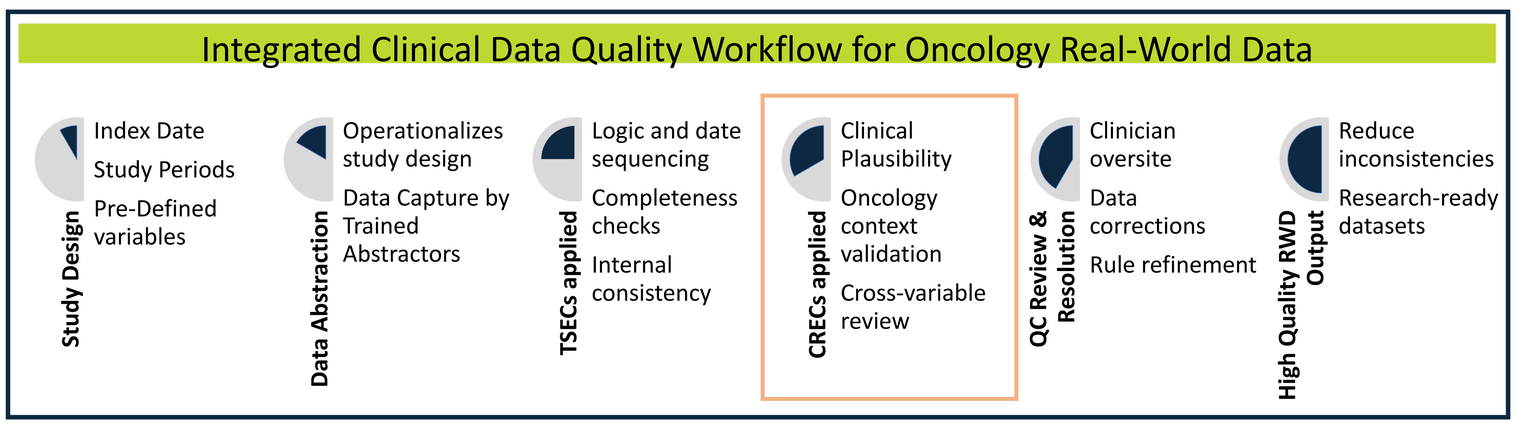
Number of edit checks resulting in data modifications (queries) calculated

Data Analysis

TSECs and CRECs

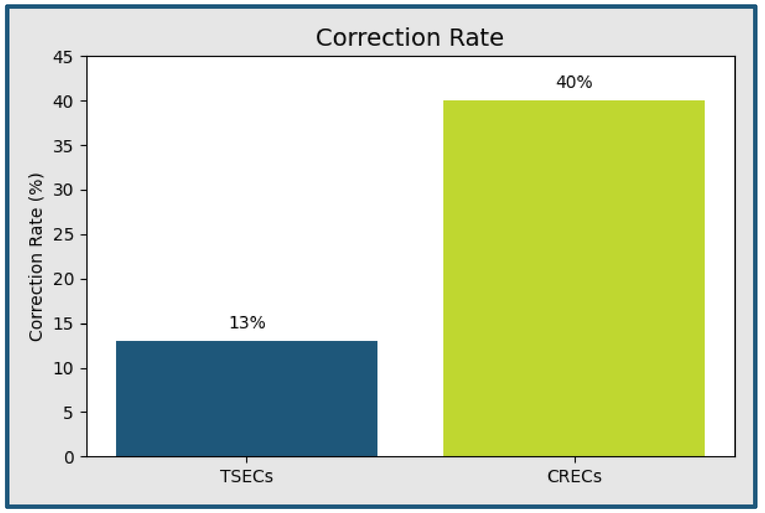
Descriptive Statistics applied to analyze correction rates for TSECs and CRECs

Data Quality Process



Real world research requires a structured and reliable quality control process to ensure a research-ready dataset. Each step in the process further strengthens the data quality and requires a multi-disciplinary team approach involving research scientists, biostatisticians, oncology nurse data quality leads, and data abstractors.

Results



Data were drawn from three retrospective studies, totaling 1100 charts. For this analysis, we focused on the first round of edit checks, applied to a subset of 272 charts (25% of total) across all three studies.

- TSECs applied to a subset of 272 charts (25%) across 35 variables.
  TSECs generated 502 edit checks, resulting in 65 queries.
  TSEC corrections impacted 13% of the data reviewed.
- CRECs applied to the same 272 charts (25%) across 40 variables.
  CRECs generated 616 edit checks, resulting in 249 queries.
  CREC corrections impacted 40% of the data reviewed.

Conclusions

Implementation of CRECs identified over three times more corrections than TSECs alone. These findings demonstrate incorporating clinical context into data validation meaningfully improves the accuracy and reliability of oncology real-world data. While TSECs provide a critical first pass in data quality assurance, CRECs address nuanced clinical plausibility issues that directly influence the downstream analyses. Future work should expand CRECs across tumor types and establish new oncology standard edit checks to enable scalable, quality RWD for research and regulatory decision-making.

Considerations

- CRECs require clinical perspective to generate and increase the review time needed during the quality control process.
- CRECs are difficult to generalize, as they are typically study specific.
- TSECs and CRECs assessed for different variables, so direct comparisons cannot be performed.

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