T1D Exchange: Review of Common Errors during EMR Data Mapping and Transformation to the T1D Exchange QI Data Specification

Anton Wirsch, Ann Mungmode, Saketh Rompicherla, Emma Ospelt, Nudrat Noor, Joyce Lee, Marina Basina, Jesse Cases, Osagie Ebekozien

Background

The T1D Exchange Quality Improvement Collaborative (T1DX-QI) is a network of 50 endocrinology data-sharing centers collaborating with the goal of improving care for people with type 1 diabetes. T1DX-QI is a Learning Health System where participating centers transfer de-identified electronic medical records (EMR) after data mapping, transformation, and validation for population health improvement.

T1DX-QI centers map existing data fields in their respective EMR to a unified T1DX-QI data specification (T1DX-DS). T1DX-DS contains 120 variables across seven files including the Patients, Providers, Encounters, Observations, Conditions, Medications and Diabetes Files. Errors in the mapping and transformation process lead to a significant delay in the onboarding process.

Methods

The authors reviewed the initial files shared from eleven centers and feedback documentation from the T1DX-QI team following the initial data mapping and transformation process. These errors were categorized by T1DX-DS file and common error types.

Results

- All clinics had issues with Value Case, Value Coded, and Value Incorrect (Figure 1 left)
- Error types are not evenly distributed among clinics
- All clinics had issues with Encounter, Observation, and Patient files (Figure 1 right)
- 73% of the clinics had errors in the Diabetes section (Figure 1 right)

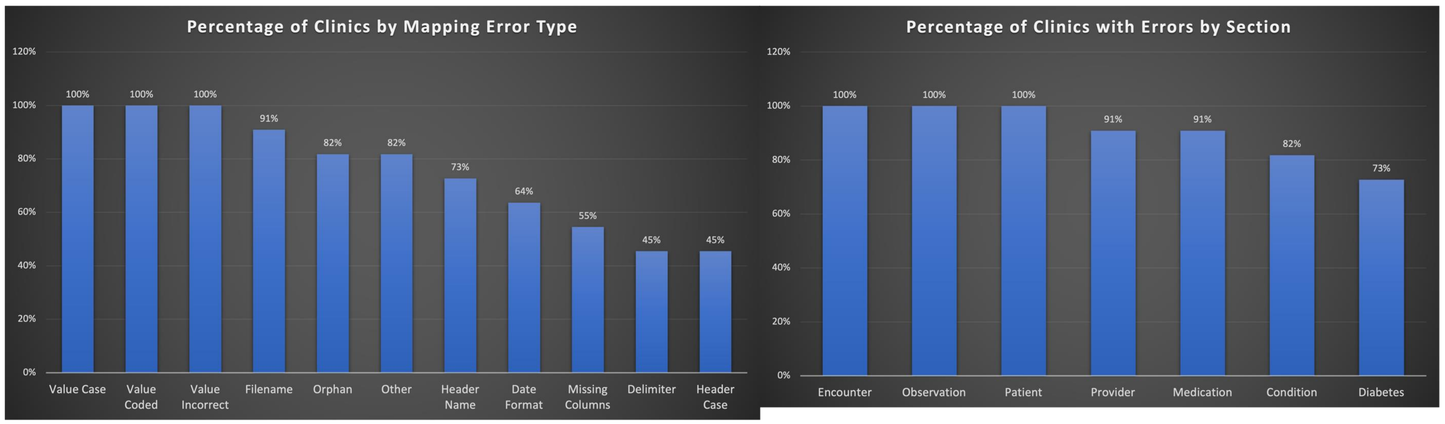

Figure 1 Type Percentage of Clinic by Error Type and Percentage of Clinics with Errors by Section

Discussion

Our findings show that data quality is a common issue in EMR systems.

- Data quality issues, present in 91.7% of the articles reviewed[1]
- Data include inconsistent or missing diagnostic coding and risk factor designation, 'dirty data' (misspelled words, inconsistent word strings, free text strings instead of structured data), missing 'meta-data'[2]

Clinics being aware of common data mapping issues at the start of the data mapping process should reduce or eliminate these types of errors and thus reduce the time to complete data mapping. To assist with this, we have created a checklist with the data quality checks that are conducted during the data mapping process. This will be provided to clinics at the start of the data mapping process to bring awareness of common errors.

The data mapping process is also an opportunity for clinics to uncover potential data quality issues in their EMR system and address them.

Conclusion

Our review identified common data mapping and transformation errors to the T1DX-DS. The errors could be due to several factors that include the clarity of the data specifications, local clinic capacity, nature of the data, and technical challenges.
T1DX-QI data team is testing multiple improvement interventions to reduce these common errors in the future and further streamline the data mapping process.

1. Edmondson M, Reimer A. Challenges Frequently Encountered in the Secondary Use of Electronic Medical Record Data for Research
2. Greiver M, Barnsley J, Aliarzadeh B. Using a data entry clerk to improve data quality in primary care electronic medical records: a pilot study. Inform Primary Care.2011 ; 19 : 241 – 250

This research was supported in part by
The Leona M. and Harry B. Helmsley Charitable Trust