

Lift Off: From Real World Data Quality Guidelines to Actionable Assessment

Mui Van Zandt¹, Elizabeth Eldridge¹, Kelly Brotherhood¹, Christina Mack¹

1. IQVIA, Durham, North Carolina



Background

As efforts increase to assess regulatory acceptability of Real-World Data (RWD) submissions, the number of guidelines around RWD quality considerations have grown^{1,2}. Globally, these align on emphasizing the need for transparency in RWD quality assessments, focused on the “what” and giving opportunity for “how”. The Kahn data quality framework³ harmonizes assessment methods across standards, defining three assessment categories (conformance, completeness, plausibility) within two contexts (verification, validation). Creation of standardized tools can operationalize RWD guidance in a transparent, reproducible manner.

Objectives

1. Translate the Kahn framework into actionable business rules and thresholds defined in accordance with regulatory guidance, applicable to assessing quality of any Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) data asset
2. Deploy metrics through an interactive OMOP-compatible dashboard, enabling determination of data’s fit-for-purpose for generation of Real-World Evidence (RWE)
3. Illustrate how researchers can use this data quality dashboard, coupled with the open science analytics platform Atlas, to determine the suitability of a data asset for use in generating RWE

Methods

We developed a data quality dashboard operationalizing the Kahn framework, defining business rules, constraints and thresholds measuring conformance, completeness, and plausibility of data measures at the field, table, and concept levels. The dashboard’s assessment of patient ethnicity measures within an Electronic Medical Record (EMR) OMOP data asset is provided as an example. This dashboard was developed in partnership with the Observational Health Data Science & Informatics (OHDSI) community. An overview of the data quality dashboard is presented in *Figure 1*. Findings presented in the dashboard at the data asset level were further characterized at the patient cohort level using Atlas, to illustrate the capabilities of these tools to assess RWD’s fit-for-purpose when coupled together.

Table 1: Results from the data quality dashboard assessing an ambulatory EMR OMOP data asset

N = 3,312 data quality checks for conformance, completeness and plausibility in two contexts: verification and validation

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1620	219	1839	88%	276	11	287	96%	1896	230	2126	89%
Conformance	681	0	681	100%	104	0	104	100%	785	0	785	100%
Completeness	382	4	386	99%	15	0	15	100%	397	4	401	99%
Total	2683	223	2906	92%	395	11	406	97%	3078	234	3312	93%

Disclosures: MVZ, EE, KB and CM are all employees at IQVIA

Results

- Eleven data quality checks passed within the set thresholds
- 1 data quality check failed due to mismatched field-level values between tables where high concordance was expected
- This information can be used to investigate source data, where needed

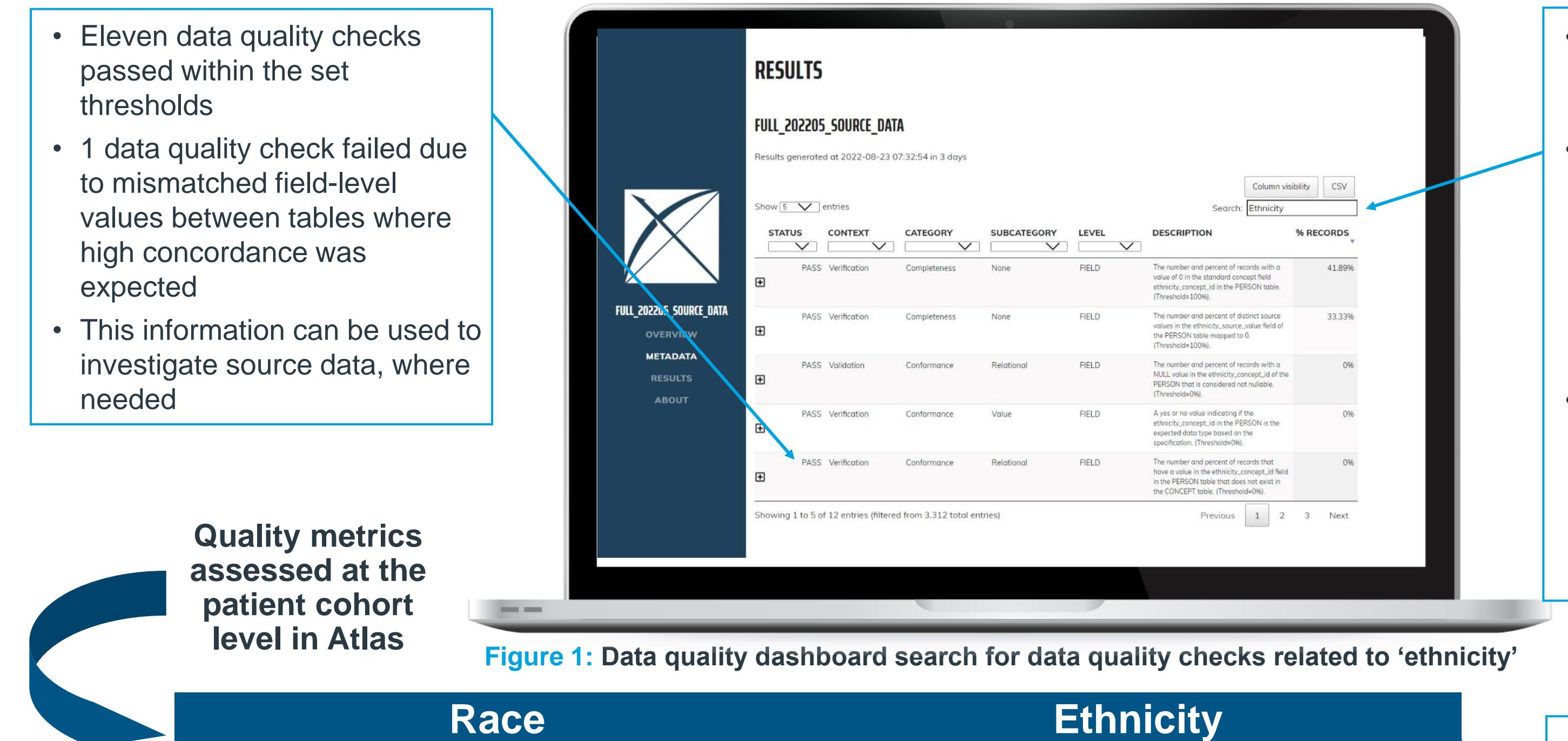


Figure 1: Data quality dashboard search for data quality checks related to ‘ethnicity’

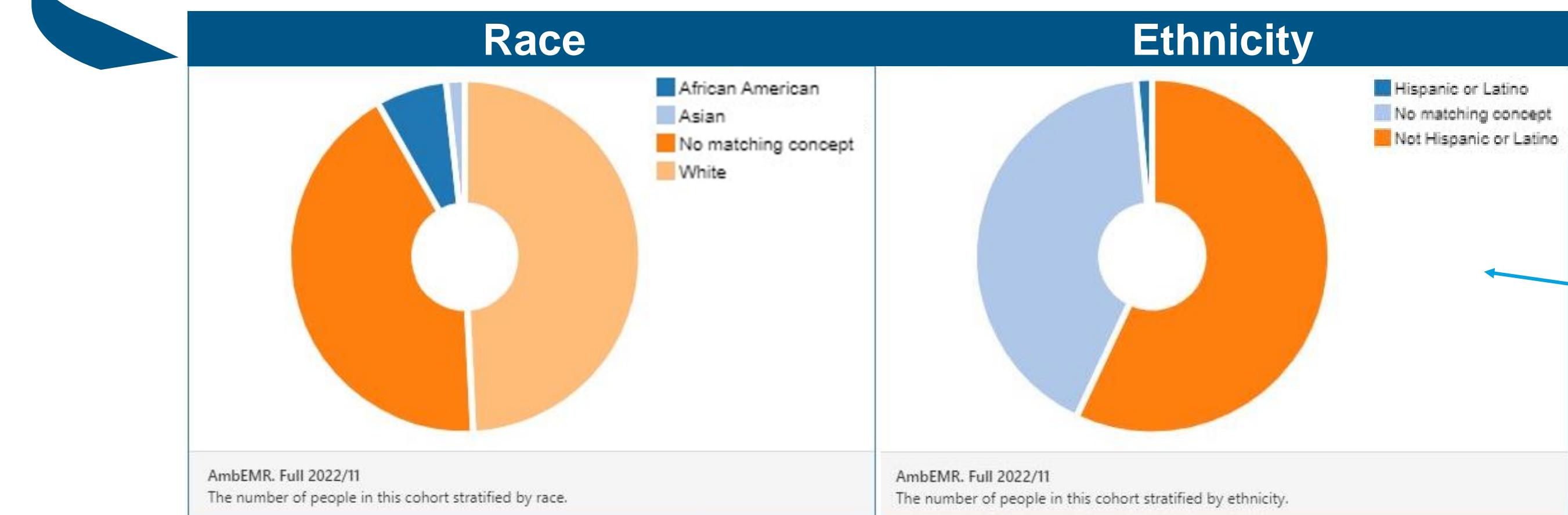


Figure 2: Breakdown of cohort by race and ethnicity in Atlas

Conclusions

- Implementing RWD quality guidelines within a data quality dashboard enables researchers and regulators to select, evaluate and communicate the quality of their data in a rapid, consistent manner with a shared vocabulary
- The data quality dashboard, coupled with the Atlas tool, enables real-time results facilitating transparent assessment of fit-for-purpose data

References

1. Girman CJ, Ritchey ME, Lo Re V: Real-world data: Assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products. *Pharmacoepidemiology and Drug Safety* 2022; 31:717-720.
2. National Institute for Health Care Evaluation: NICE real-world evidence framework. 2022.
3. Kahn MG, Callahan TJ, Barnard J, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. *EGEMS (Wash DC)*. 2016;4(1):1244. Published 2016 Sep 11. doi:10.13063/2327-9214.1244

Follow to see the OHDSI Data Quality Dashboard GitHub page

