Beyond Missingness: Systematizing Methods for Comprehensive Data Fitness Assessment in Clinical Research

OHDSI Community Call

June 30, 2026

Hanieh Razzaghi, PhD, MPH & Charles Bailey, MD, PhD

PEDSnet
a pediatric learning health system

The Problem

Analysis relies on the day saying what you mean

The data say what they mean.

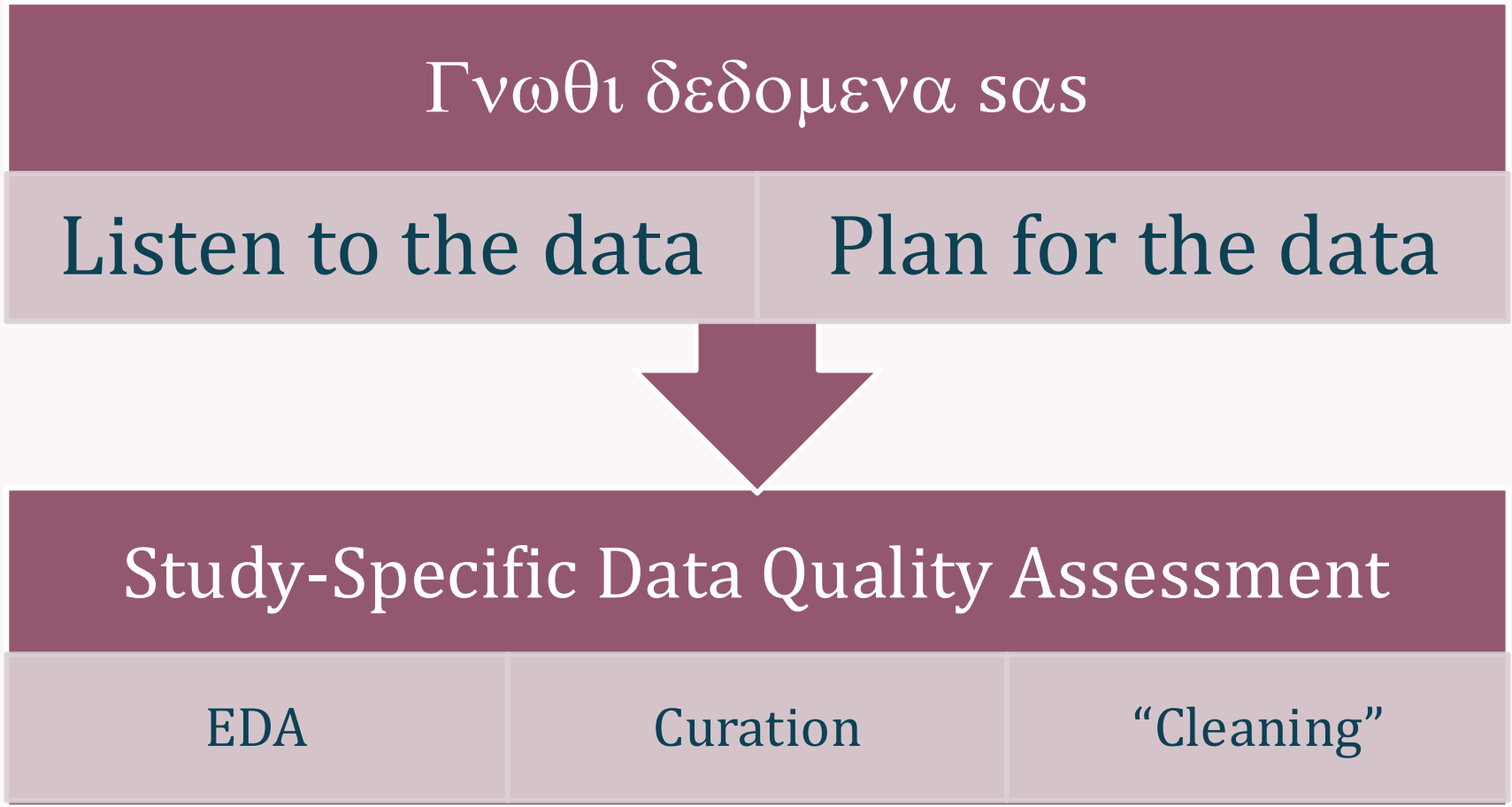
You can't change the data.

You can't see the data.

You can't recollect the data.

PEDSnet
a pediatric learning health system

What to do?



PEDSnet
a pediatric learning health system

Network (Structural) DQA

Tends to focus on structure

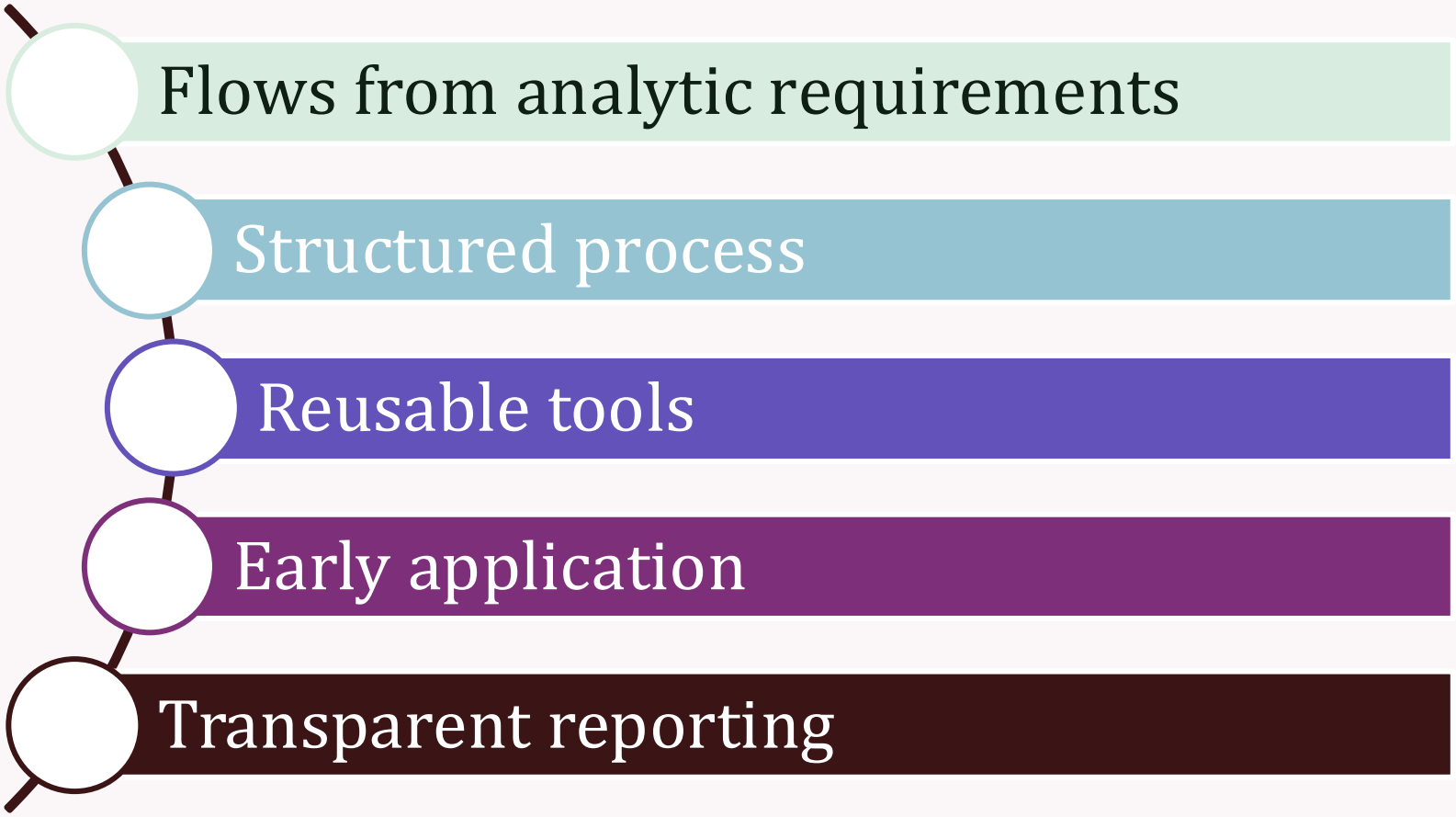
- Missingness
- Data model conformance
- Value sets or ranges

Plausibility focuses on never events

- Impossible dates
- Physiologic impossibilities

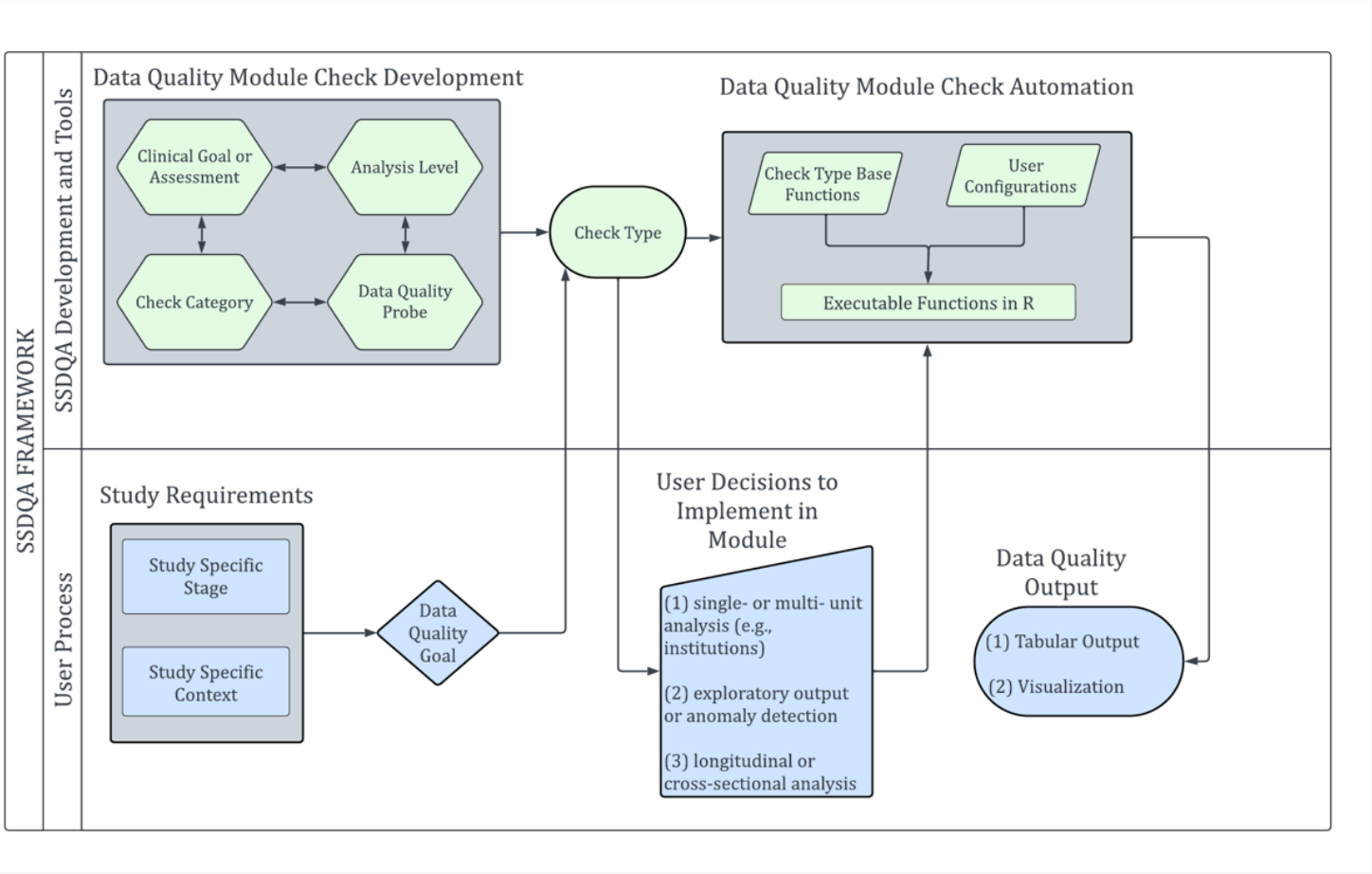
PEDSnet
a pediatric learning health system

Study-Specific DQA



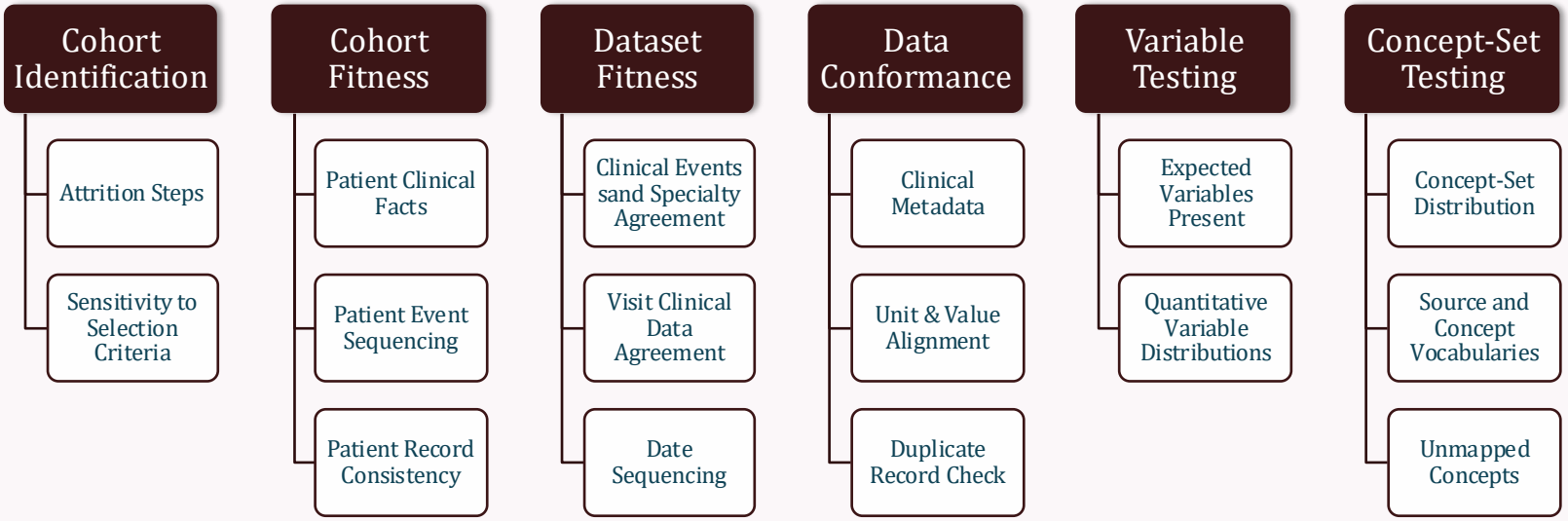
PEDSnet
a pediatric learning health system

Conceptual Framework



PEDSnet
a pediatric learning health system

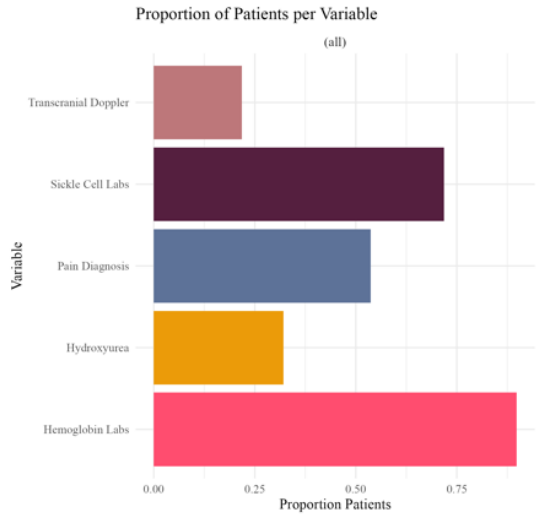
SQUBA: Study-Specific DQ Modules



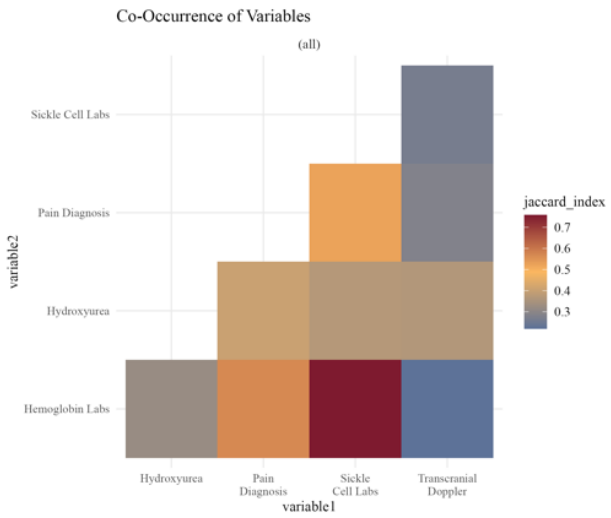
PEDSnet
a pediatric learning health system

Applying SSDQA: EVP check

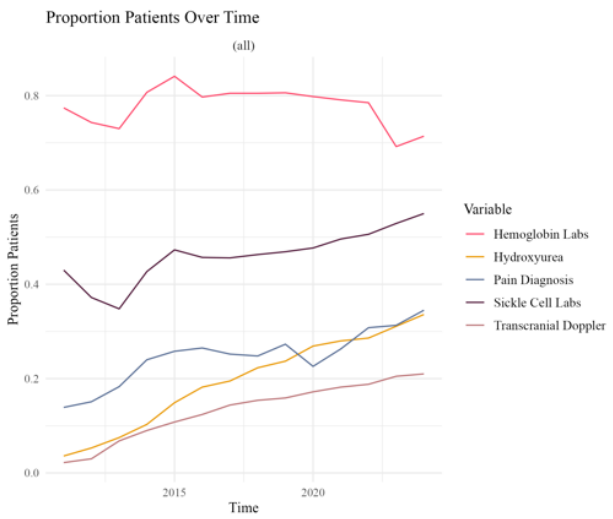
A



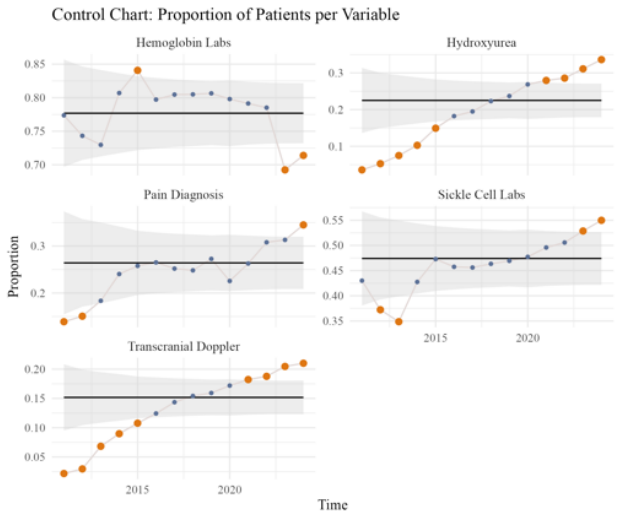
B



C

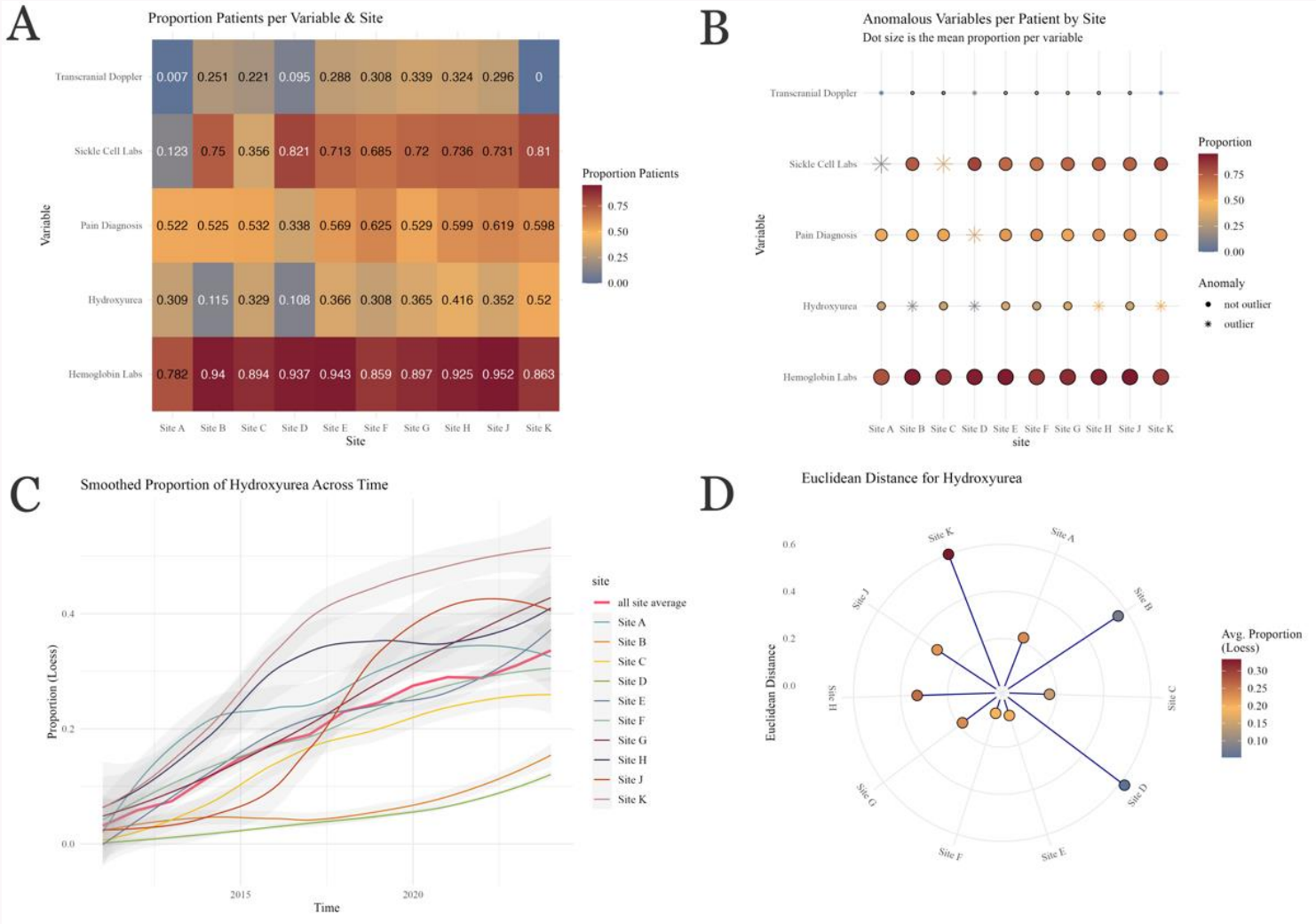


D



PEDSnet
a pediatric learning health system

Stratifying Data



PEDSnet
a pediatric learning health system

How does SQUBA compare to other tools?

Metric	SQ	CD	DG	DQR	DQD	EC	EN	MIR	N3C	PC	SPF
Study cohorts	3	3	3	3	1	1	1	1	1	1	3
Theoretical framework	3	2	3	3	3	1	1	3	2	1	1
Check diversity	3	2	3	3	3	1	1	3	3	2	2
Check names/catalog	3	2	3	3	3	2	1	3	1	1	2
Extensibility	3	1	2	2	3	2	1	2	1	1	3
Customization	3	1	3	3	2	1	1	3	1	1	3
Data Models	3	2	0	1	2	3	2	3	3	2	1
Packaging	3	3	0	3	3	2	3	3	2	3	0
Availability	3	3	0	3	3	2	3	2	1	1	0
Dependencies	3	2	0	3	3	3	3	3	1	3	0
Coding required	3	3	0	3	3	3	3	3	3	3	0
Integration	2	3	0	2	3	3	3	3	2	3	0
Pass/Fail Checks	3	1	3	3	3	1	1	3	3	3	3
Tabular output	3	3	0	3	3	3	3	3	3	3	0
Visualizations	3	1	0	3	1	2	3	3	2	1	0
Temporal dimension	3	1	1	1	1	1	1	1	1	1	1

Labels: SQ – SQUBA; CD – OHDSI CohortDiagnostics; DG – DataGauge; DQR – dataquieR; DQD – OHDSI Data Quality Dashboard; EC – DQe-c v2; EN – ENACT Data Quality Explorer; MIR – MIRACUM data quality workbench (DQAstats/DQAgui) ; N3C- National COVID Cohort Collaborative data quality program; PC – PCORnet® Data Curation; SPF – SPFID2

PEDSnet
a pediatric learning health system

Making Metadata Available

PEDSnet
a pediatric learning health system

Our Network | Research | Youth & Families | Learning | Database | Work with Us | Log In

Collections & Domains | All of PEDSpace | Statistics

Home • Data Quality Modules • Cohort Identification • Sensitivity to Selection Cr… • Sensitivity to Selection Cr…

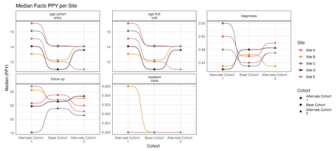
Data Quality Check

Sensitivity to Selection Criteria: Multi-Site, Exploratory, Cross-Sectional Analysis

DOI

Created

2024-12-17



Click on the thumbnail above to preview images.

Files

2 Downloads

Domain

Cohort Identification

Category

Plausibility

Parameters

Multi-Site Analysis | Exploratory Analysis | Cross-Sectional Analysis

Abstract

The Sensitivity to Selection Criteria module measures how different cohort inclusion definitions can impact the makeup of a cohort, including demographics, utilization, and outcomes. Each check will compare a user-provided base cohort definition to each provided alternate cohort definition to evaluate how the new definitions impact each measured characteristic.

Data Requirements

`base_cohort` , `alt_cohorts` , `multi_or_single_site` , `anomaly_or_exploratory` , `person_tbl` , `visit_tbl` , `provider_tbl` , `care_site_tbl` , `demographic_mappings` , `specialty_mappings` , `specialty_concepts` , `outcome_concepts` , `domain_tbl` , `domain_select`

Probe

Misclassification Detection | Eligibility Criteria Assessment | External Benchmarking | Selection Error or Bias Detection

Clinical Assessment

Targeted Patient Population | Clinical Data Distributions

Access Package

```
# install.packages("devtools")
devtools::install_github('ssdqa/sensitivityselectioncriteria')
```

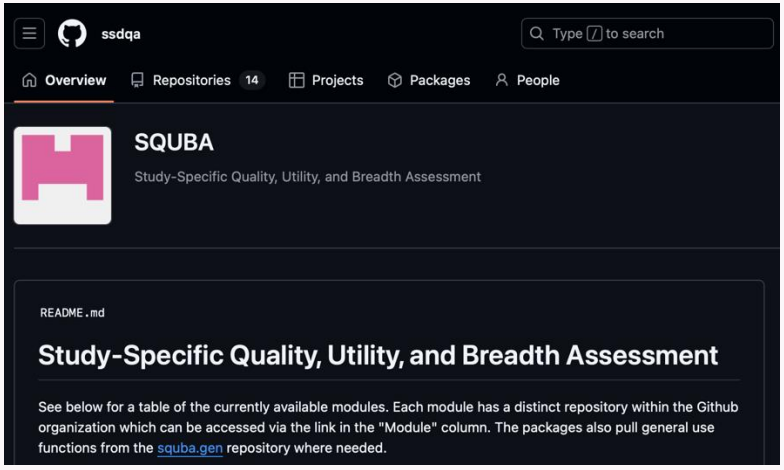
Visualization Output

PEDSnet
a pediatric learning health system

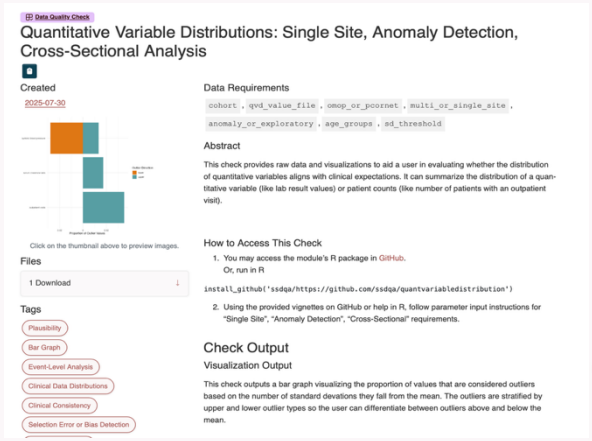
Thank you!

[SQUBA on PEDSpace]

Available in Github (SQUBA)



More descriptions with example output on PEDSnet metadata repository



PEDSnet
a pediatric learning health system