A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks

Hanieh Razzaghi, PhD

12/02/2025

PEDSnet
a pediatric learning health system

Acknowledgments

This project was funded through a Patient-Centered Outcomes Research Institute® (PCORI) Infrastructure Award (RI-CHOP-01-PS2). The statements presented in this work are solely the responsibility of the author(s) and do not necessarily represent the views of PCORI or its Board of Governors.

PEDSnet Institutions:

- Children's Hospital of Philadelphia
- Cincinnati Children's Hospital
- Children's Hospital of Colorado
- Lurie Children's Hospital
- Nationwide Children's Hospital
- Nemours Children's Hospital
- Children's National
- Seattle Children's Hospital
- Stanford Children's Hospital
- Texas Children's Hospital

Study Team

- Charles Bailey, MD, PhD
- Chris Forrest, MD, PhD
- Kimberley Dickinson, MS
- Kaleigh Wieand
- Samuel Boss
- Hunter Weidlich
- Yungui Huang, PhD
- Keith Morse, MD
- Sujan Kumar Mutyala, MS
- Jyothi Priya Alekpatti Nandagopal, MS
- Karthik Viswanathan, MS

PEDSnet
a pediatric learning health system

Background

Previous PEDSnet Data Quality Program

- Reliant on Github
- Lacked visualizations and researcher access
- Checks were oriented for an informatics audience

Current Research

- Common vocabularies and taxonomies
- Learn from other networks and programs

PEDSnet
a pediatric learning health system

Objectives

Assess research readiness

- Flexible checks that can be customized
- Results accessible to wide audience

Portable across networks

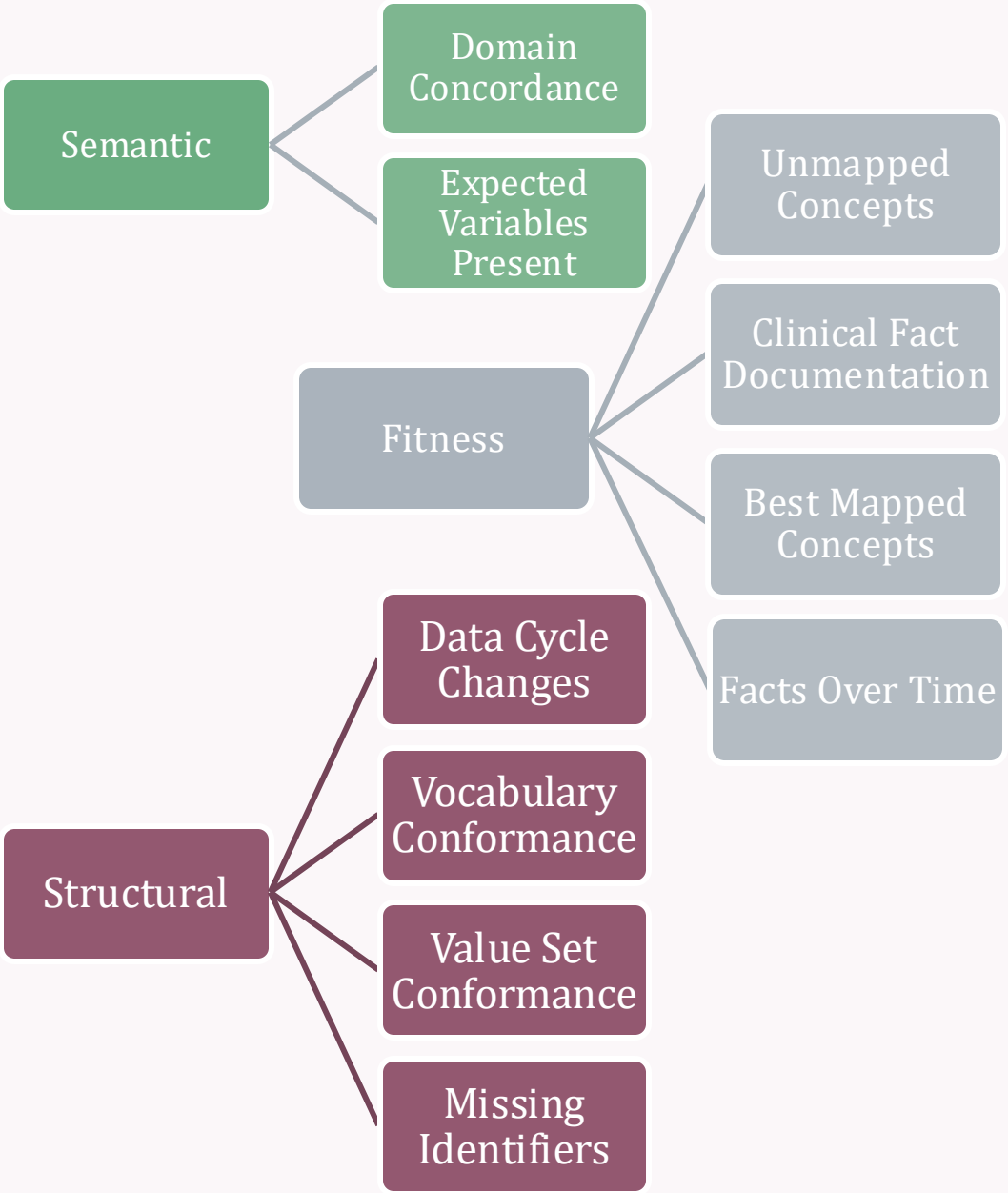
- Compatible with OMOP data model
- Eventual compatibility with PCORnet
- Packaged on Github

Comprehensive system

- Focus on automation
- Comprehensive tracking and metadata

PEDSnet
a pediatric learning health system

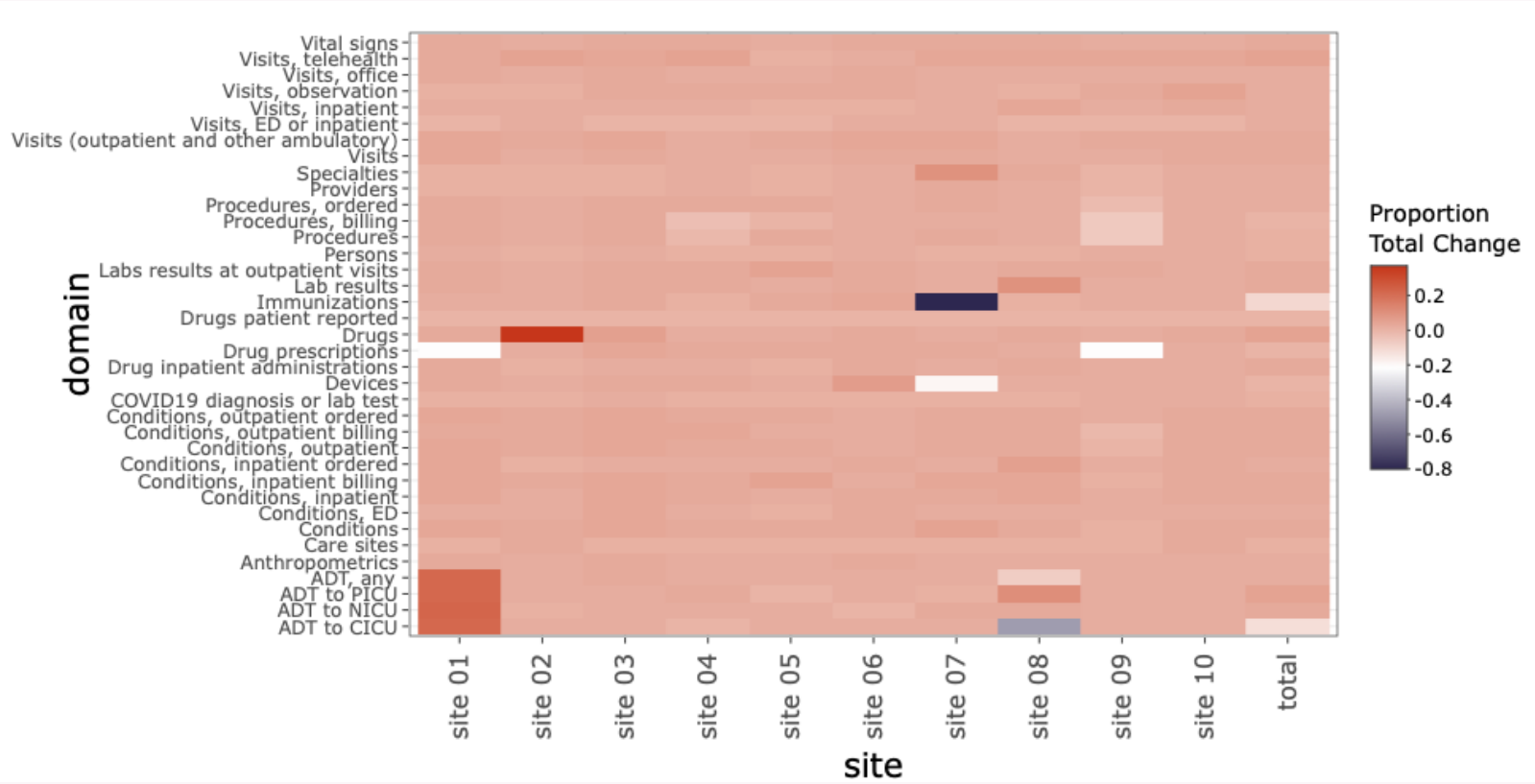
Check Development: Number / Types



- ~ 350 checks across the network
- 10 Check Types
- Applied with every data refresh
- New checks added every cycle

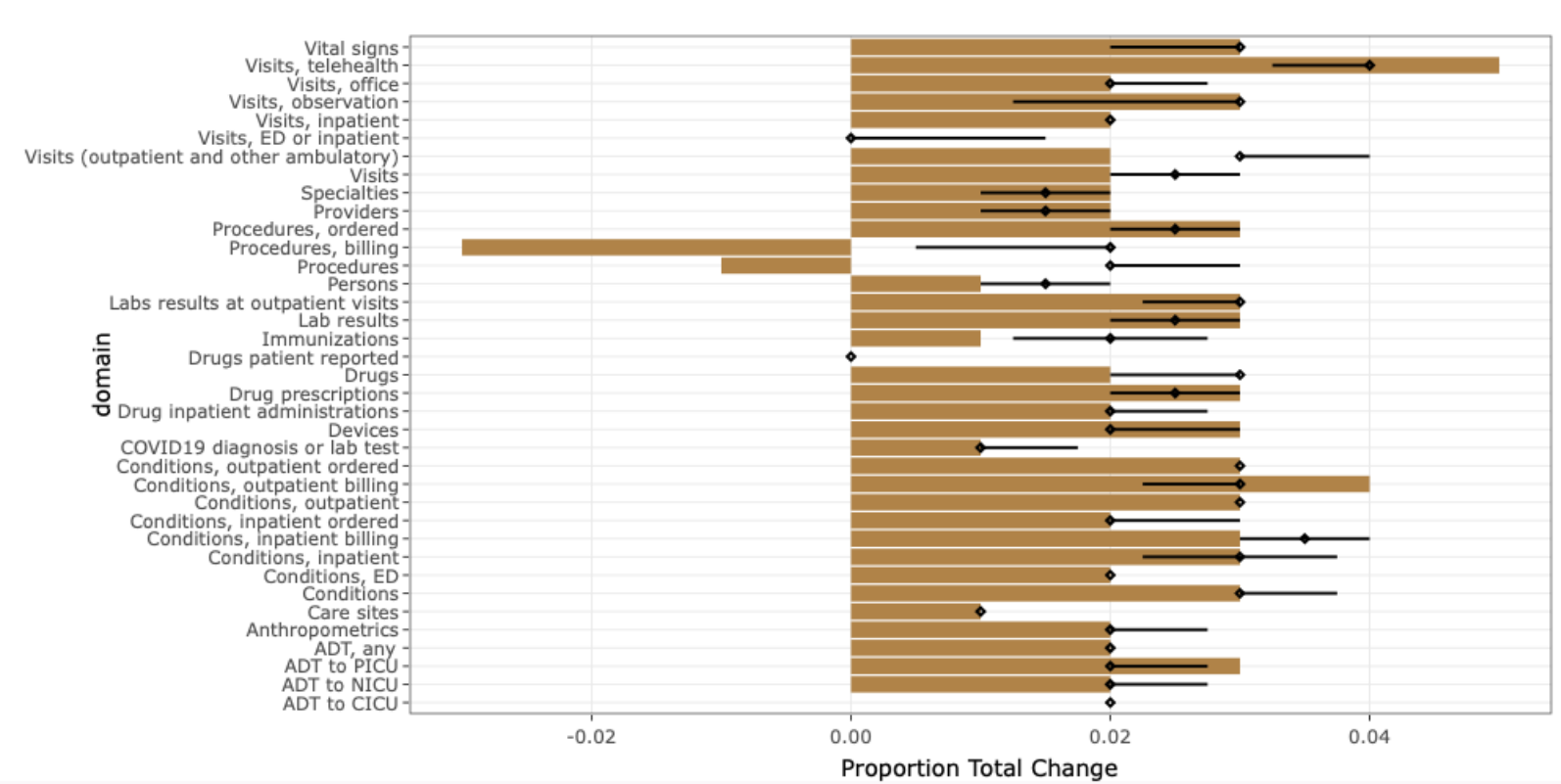
PEDSnet
a pediatric learning health system

Structural Checks: Data Cycle Changes



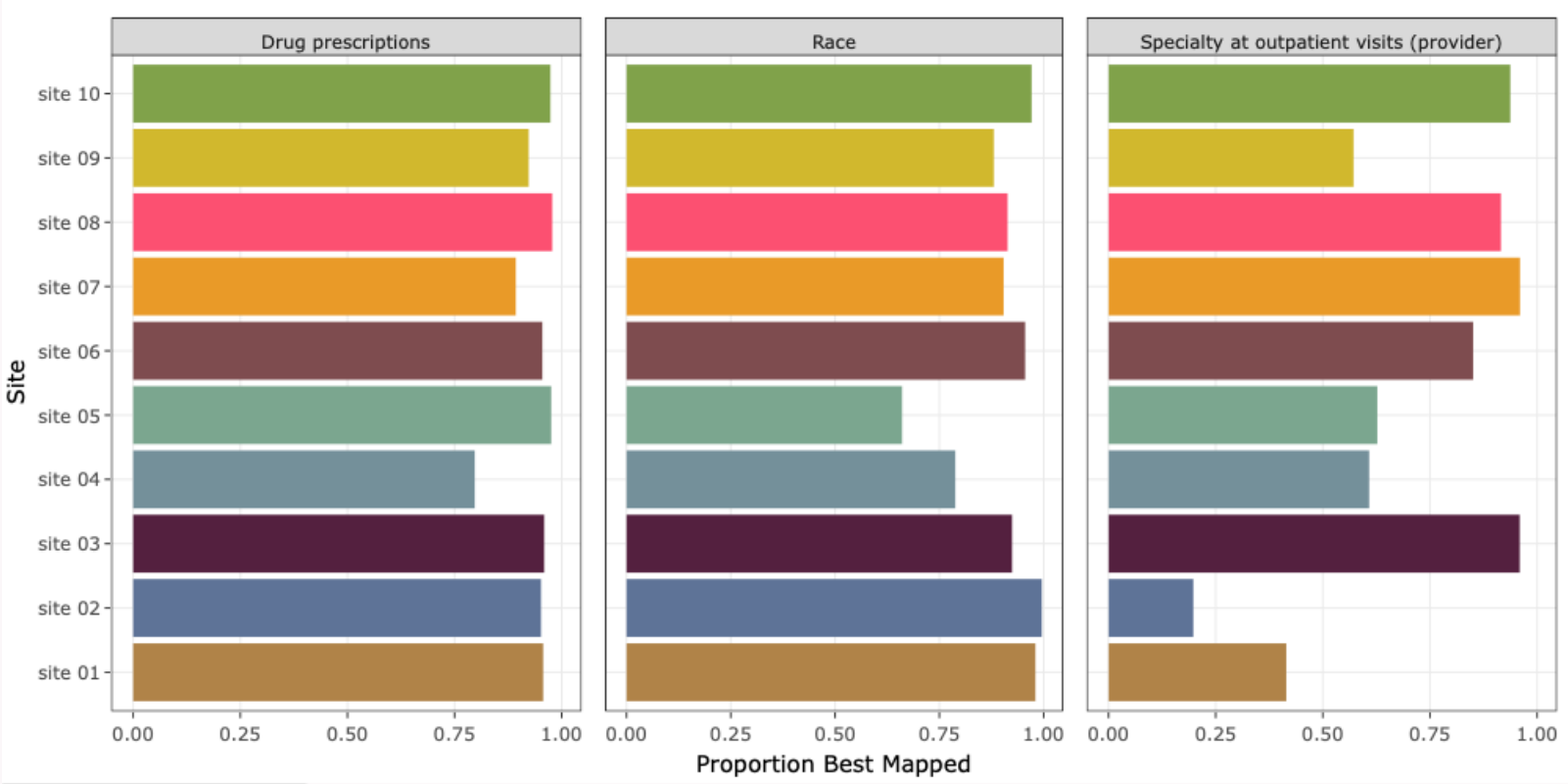
PEDSnet
a pediatric learning health system

Alternate View: Data Cycle Changes



PEDSnet
a pediatric learning health system

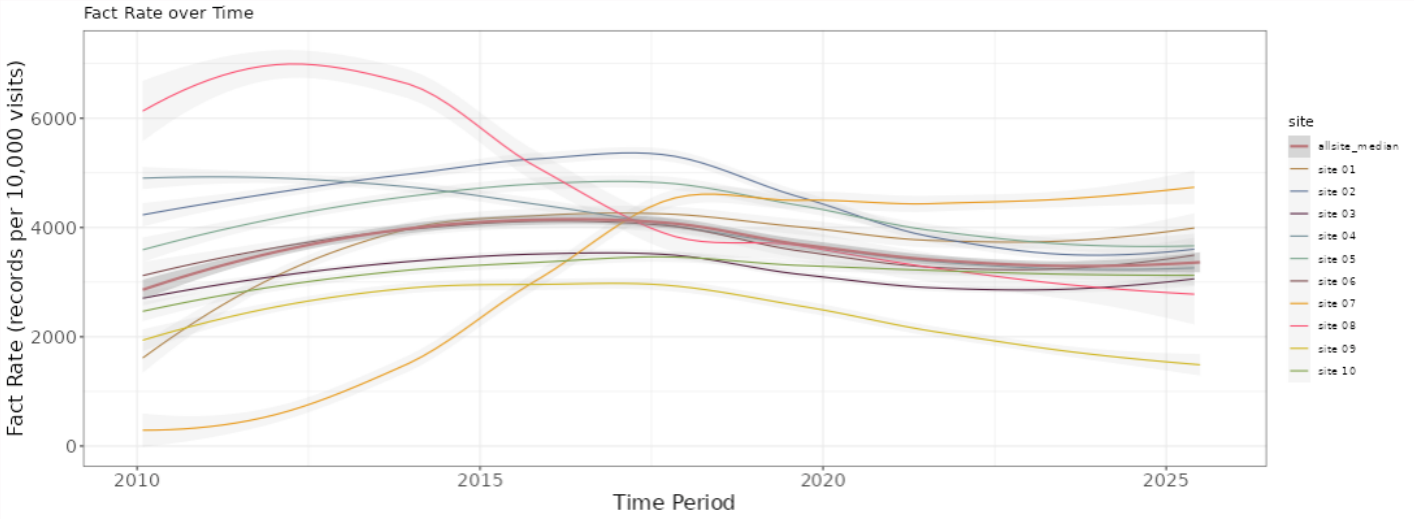
Data Fitness Checks: Best Mapped Concepts (Rx, Race, Provider Specialties)



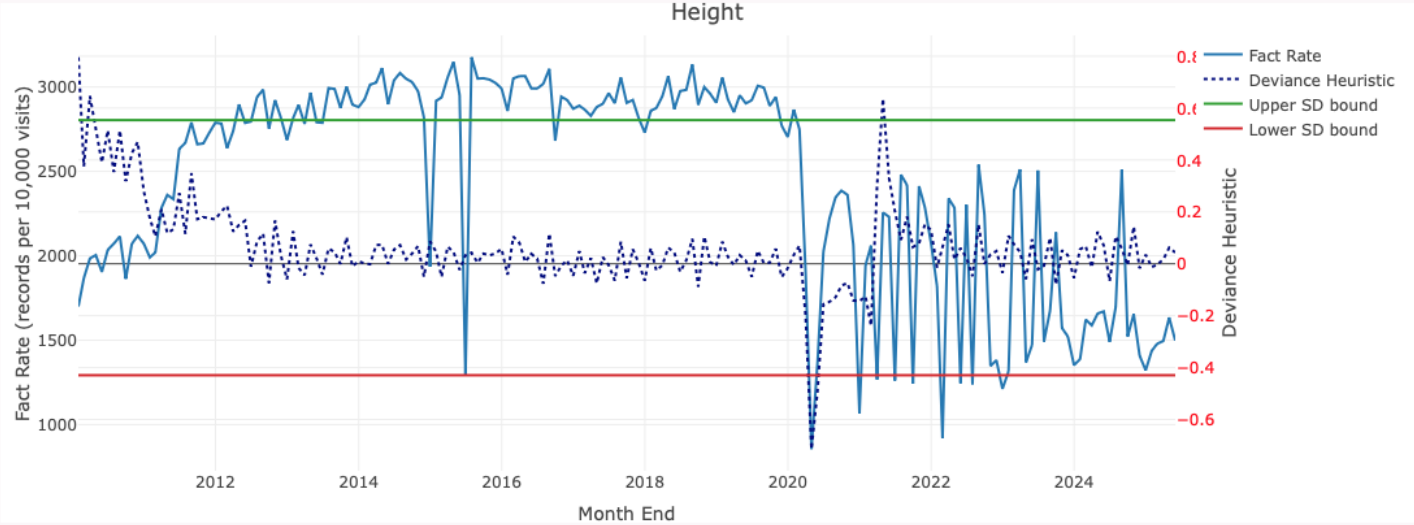
PEDSnet
a pediatric learning health system

Data Fitness Checks: Facts Over Time (Measured Heights)

Summary Plot

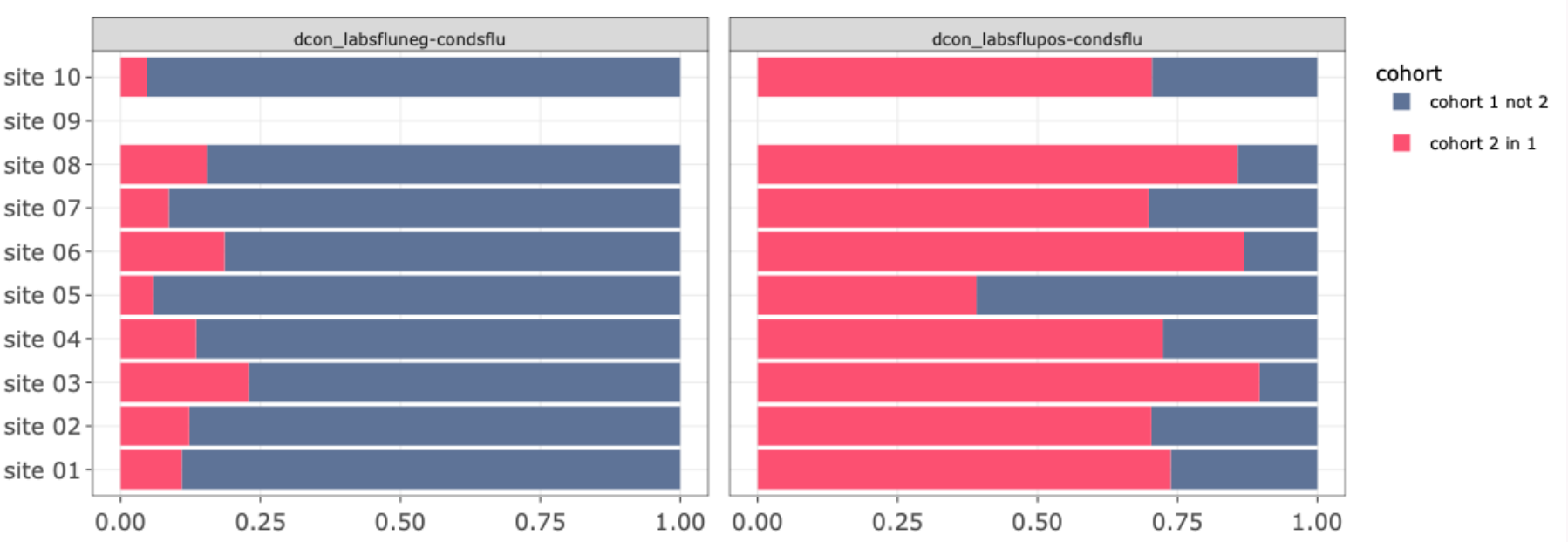


Site Specific Facts Over Time

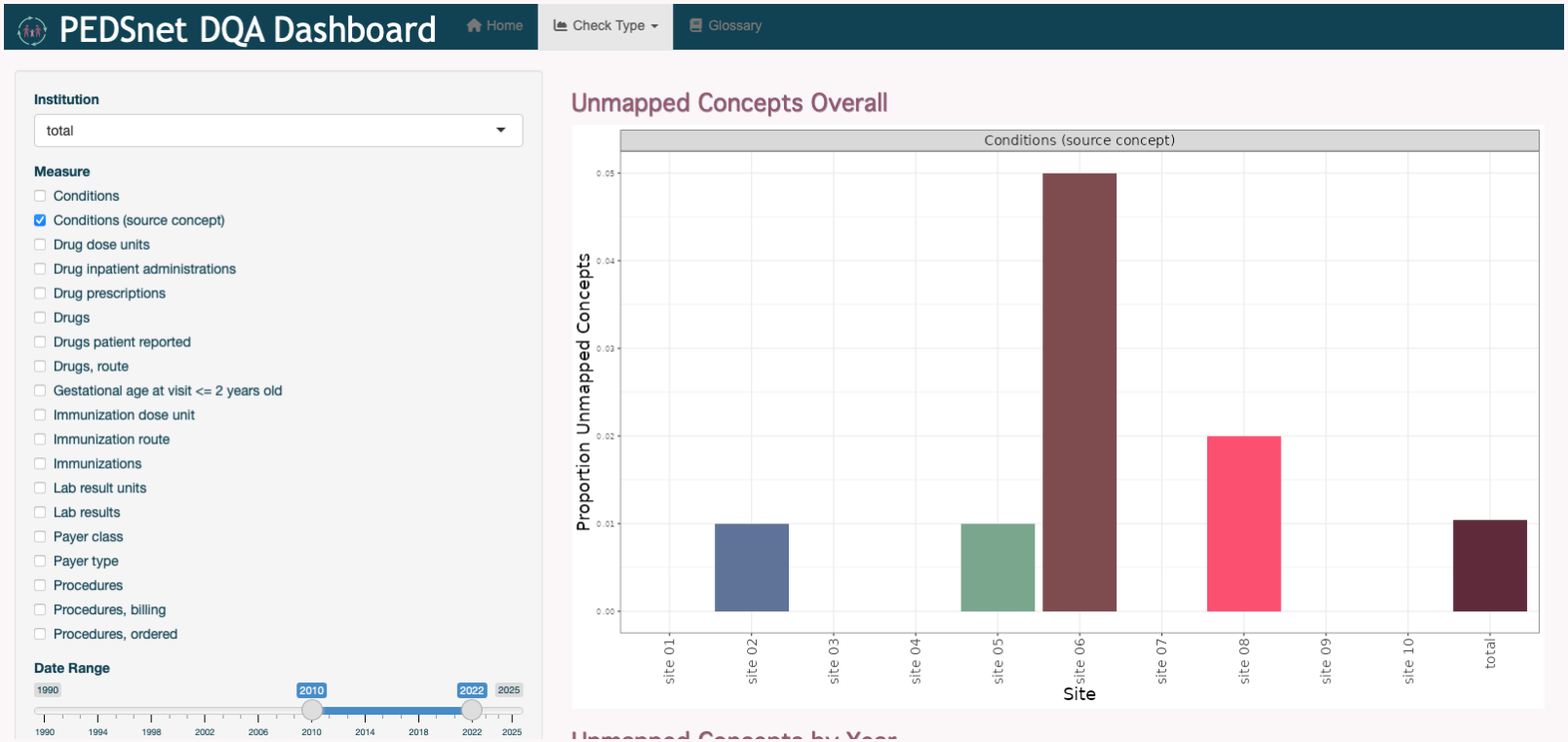


PEDSnet
a pediatric learning health system

Semantic Checks: Domain Concordance (Flu test / Flu dx)



PEDSnet DQ Dashboard
https://pedsnet.org/ndq/



Tracking and Reporting: Data Cycle Level

PEDSnet Current Data Quality Summary

	Plausibility	Consistency	Concordance	Completeness	Conformance	Total
	% Passed (N = 15)	% Passed (N = 72)	% Passed (N = 9)	% Passed (N = 63)	% Passed (N = 15)	% Passed (N = 174)
Site A	93.33%	88.89%	66.67%	93.65%	100.00%	90.80%
Site B	100.00%	98.61%	88.89%	96.83%	100.00%	97.70%
Site C	93.33%	100.00%	88.89%	84.13%	100.00%	93.10%
Site D	93.33%	81.94%	77.78%	93.65%	100.00%	88.51%
Site E	93.33%	93.06%	88.89%	96.83%	93.33%	94.25%
Site F	46.67%	91.67%	88.89%	73.02%	93.33%	81.03%
Site G	86.67%	95.83%	88.89%	95.24%	93.33%	94.25%
Site H	93.33%	100.00%	77.78%	84.13%	100.00%	92.53%
Site J	100.00%	90.28%	88.89%	92.06%	86.67%	91.38%
Site K	66.67%	97.22%	77.78%	84.13%	93.33%	88.51%

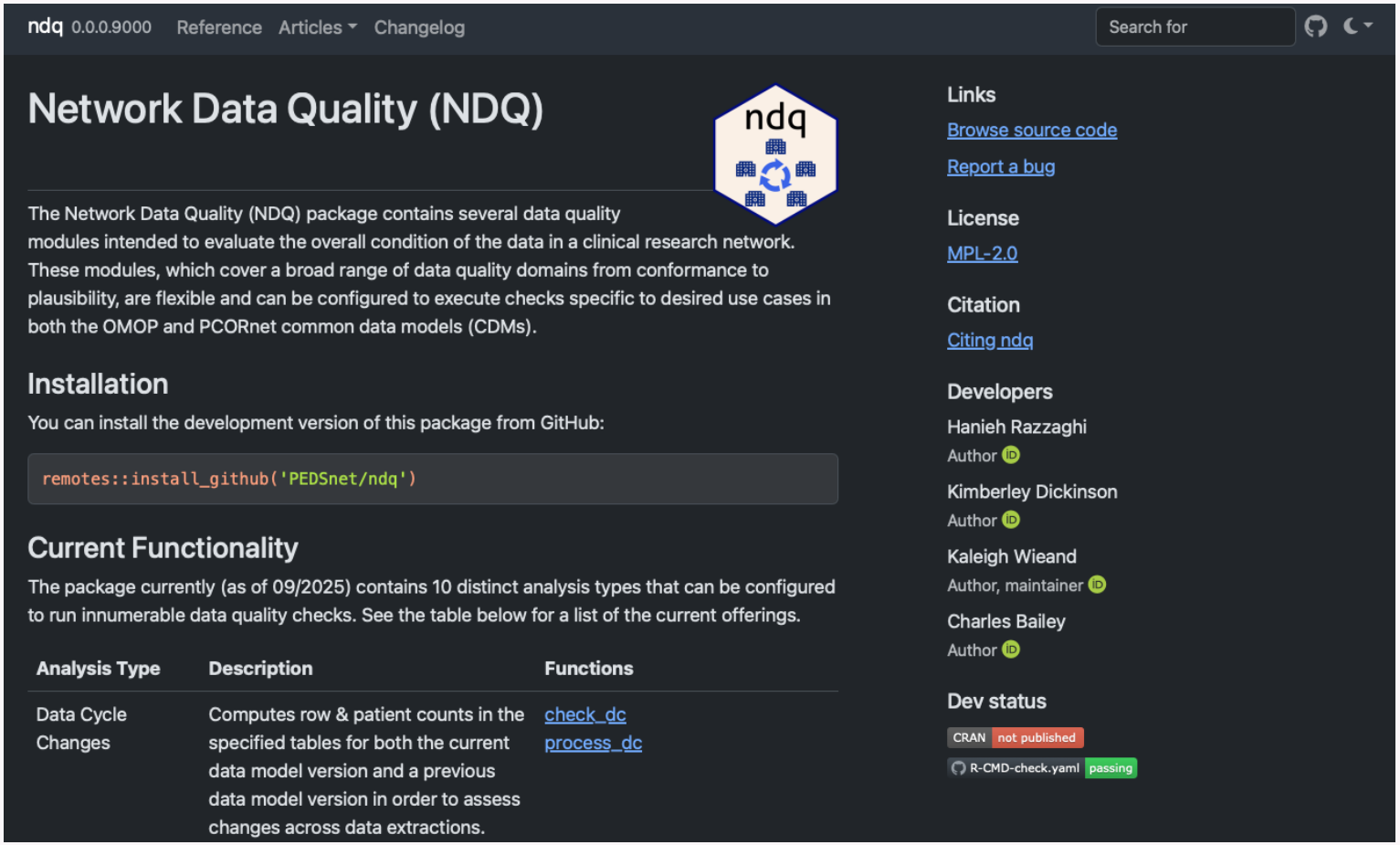
PEDSnet
a pediatric learning health system

Data Quality Cataloging

Check Type	Check Name	Check Domain	Check Application	Full Description	Check Type Short	Check Category
Data Cycle changeds	dc_demog	Person	Demographics/Person	counts in the full person table in the previous and current data submissions	dc	Consistency
Data Cycle changeds	dc_drugsall	Drug Exposure	All Drugs	counts in the full drug_exposure table in the previous and current data submissions	dc	Consistency
Data Cycle changeds	dc_condsall	Condition Occurrence	All Conditions	counts in the full condition_occurrence table in the previous and current data submissions	dc	Consistency
Unmapped Concepts	uc_drugsall	Drug Exposure	All Drugs	for the full drug_exposure table, counts of drug_concept_ids mapped to 44814650, 0, 44814653, or 44814649	uc	Completeness
Unmapped Concepts	uc_drugsip	Drug Exposure	Inpatient Administrations	for inpatient administrations (drug_type_concept_id = 38000180), counts of drug_concept_ids mapped to 44814650, 0, 44814653, or 44814649	uc	Completeness
Unmapped Concepts	uc_drugspresc	Drug Exposure	Prescription Drugs	for prescription drugs (drug_type_concept_id = 38000177), counts of drug_concept_ids mapped to 44814650, 0, 44814653, or 44814649	uc	Completeness
Unmapped Concepts	uc_drugs-doseunits	Drug Exposure	Drug Dose Unit	for the full drug_exposure table, counts of dose_unit_concept_ids mapped to 44814650, 0, 44814653, or 44814649	uc	Completeness

PEDSnet
a pediatric learning health system

NDQ (Network Data Quality) Package



PEDSnet
a pediatric learning health system

Flexibility

- Packaged on GitHub
- Applied to multiple networks
- Can be distributed (returns aggregate counts)
- Modular and extensible
- Can be applied to any OMOP database

Executable Function

```
unmapped_concepts <- check_uc(concept_list = unmapped_concepts_args,
                              produce_mapped_list = TRUE,
                              check_string = 'uc')
```

User-Defined Input List

```
unmapped_concepts_args <- list(

  "all drugs" = list(cdm_tbl('drug_exposure'),
                     'drug_concept_id',
                     'uc_dr',
                     'drug_source_value'),

  "prescription drugs" = list(cdm_tbl('drug_exposure') %>%
                                filter(drug_type_concept_id == 38000180),
                              'drug_concept_id',
                              'uc_dp',
                              'drug_source_value'),

  "all conditions" = list(cdm_tbl('condition_occurrence'),
                          'condition_concept_id',
                          'uc_co',
                          'condition_source_value')

)
```

Components of List Elements

```
1 "all drugs" = list(cdm_tbl('drug_exposure'), 2
                     'drug_concept_id',        3
                     'uc_dr',                  4
                     'drug_source_value')      5
```

1. List Element Identifier
2. Interrogated Table or Object
3. Interrogated Element
4. Application Name
5. Source Value Column (if exists)

PEDSnet
a pediatric learning health system

PEDSnet as a Learning Network

Documentation and learnings (eg, fixable problems vs site characteristic)