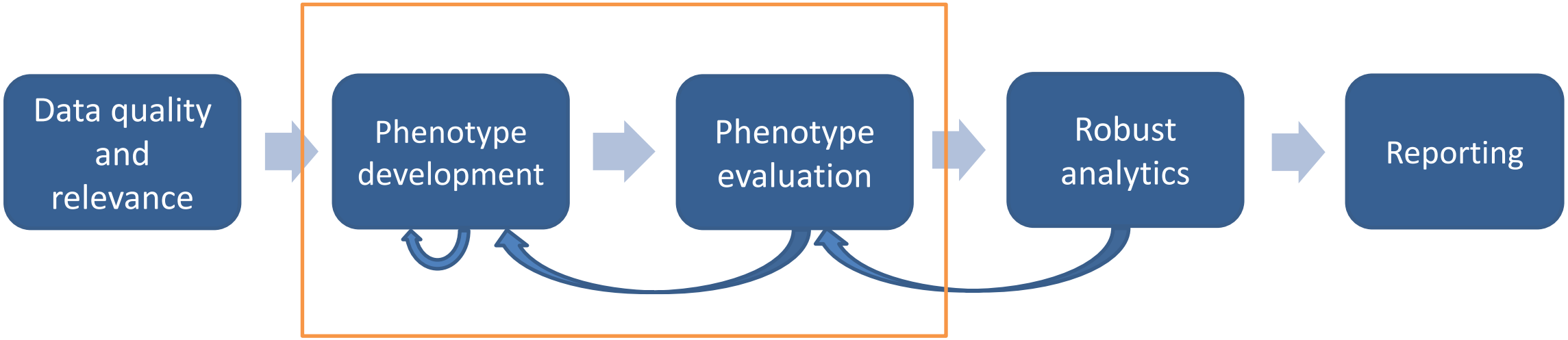
Scalable and interpretable alternative to chart review for phenotype evaluation using standardized structured data from electronic health records

Anna Ostropolets

Evidence generation

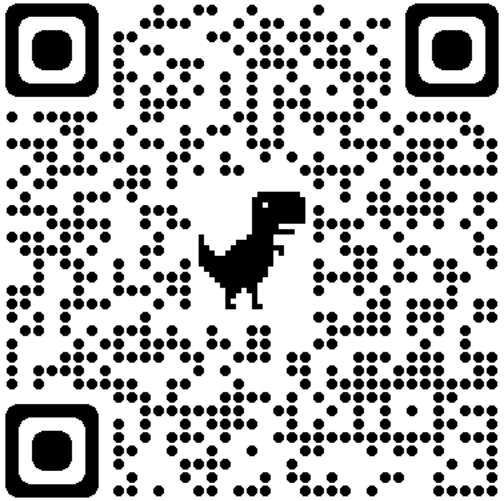


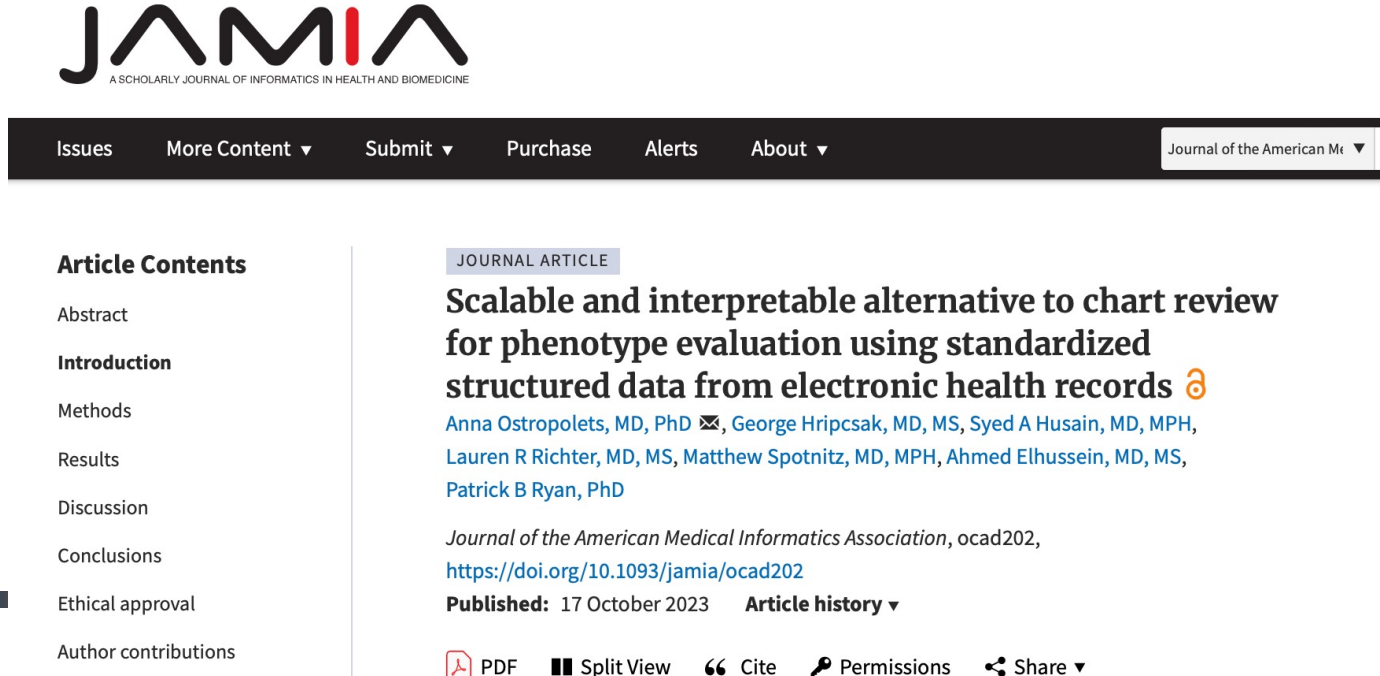
- Manual chart review - gold standard but resource- and time-consuming
- Do nothing (borrow from the literature)
+ other tools in the community (CohortDiagnostics, PheValuator)

Main challenge of chart review for phenotype evaluation

Challenge: high volume of data, which is hard to navigate and interpret

Solution: KEEPER - Knowledge-Enhanced Electronic Profile
Review system on structured data from EHR or claims data sources





JAMIA
A SCHOLARLY JOURNAL OF INFORMATICS IN HEALTH AND BIOMEDICINE

Issues | More Content | Submit | Purchase | Alerts | About | Journal of the American M

Article Contents

Abstract
Introduction
Methods
Results
Discussion
Conclusions
Ethical approval
Author contributions

JOURNAL ARTICLE

Scalable and interpretable alternative to chart review for phenotype evaluation using standardized structured data from electronic health records

Anna Ostropolets, MD, PhD, George Hripcsak, MD, MS, Syed A Husain, MD, MPH, Lauren R Richter, MD, MS, Matthew Spotnitz, MD, MPH, Ahmed Elhussein, MD, MS, Patrick B Ryan, PhD

Journal of the American Medical Informatics Association, ocad202, https://doi.org/10.1093/jamia/ocad202

Published: 17 October 2023 **Article history**

PDF | Split View | Cite | Permissions | Share

KEEPER principles

Principle 1: Adherence to clinical reasoning

KEEPER applies general principles and steps of diagnostic clinical reasoning

(clinical elements to extract: presentation, plausibility = prior history of disease, demographics, differential diagnosis, diagnostic procedures and labs, treatment and complications)

Principle 2: Dimensionality reduction

Only extract relevant information

Principle 3: Standardization

Both input and output are standardized across data sources and condition

KEEPER as an OHDSI package

Per disease:
Concept sets per
KEEPER category

Ex: ESRD Symptoms:
vomiting, edema, dyspnea

CSV table:
record per person,
column per element

Cohort

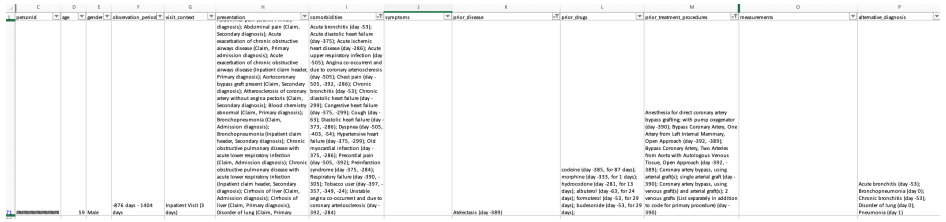
Ex: ESRD

KEEPER
data extraction

Time
windows
per category

Ex: Symptoms: -30d to
0d before index date

Person_id	Symptoms
1	Vomiting and nausea (day -29); Dyspnea (day -11);…



Experiment: Data preparation

GOLD STANDARD (AO, GH)

Random sample of 20 patients per eMERGE algorithm
Iterative review on full chart + all structured data

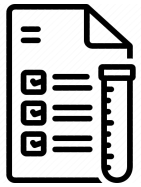
	T1DM	Acute append	COPD	ESRD
Case	12	15	11	13
Control	8	5	9	7

KEEPER PROFILES

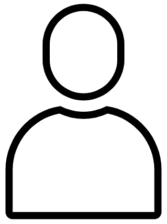
Created KEEPER profiles for 80 patients

Columbia University EHR

Experiment



EXPERIMENT (AE, LR, MS, SAH)



2 reviewers
T1DM, Appendicitis

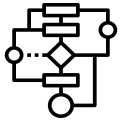
2 reviewers
COPD, ESRD

7-day washout

ROUND 1 ROUND 2

10 patients

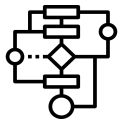
10 patients



KEEPER

Chart

Chart



KEEPER

Dataset:

- 160 patients adjudicated with KEEPER
- 160 patients adjudicated with chart review

Experiment: Performance Metrics

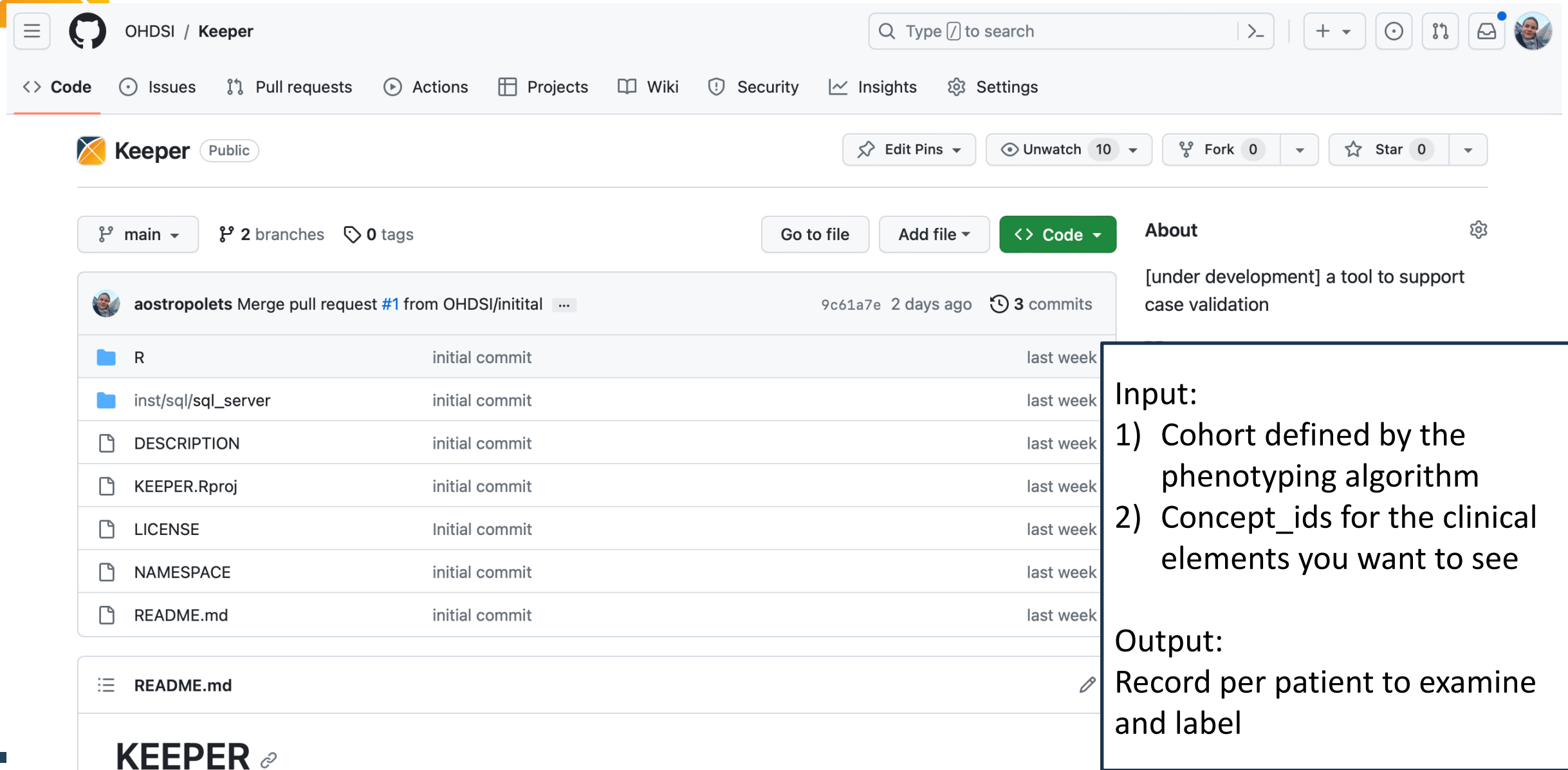
PERFORMANCE METRICS

1. Time to review
2. Agreement:
   - Agreement with the gold standard
   - Agreement of manual chart review and KEEPER
   - Agreement among reviewers

Results: time-to-review and agreement

KEEPER was:	Chart review	KEEPER
1. Twice as fast	67 minutes per 20 patients	30 minutes
2. More consistent	76.3% (61 pts) were classified similarly by 2 clinicians	91.2% (73 pts)
3. As accurate	86.9% (139 pts) were classified similarly the gold standard	88.1% (141 pts)

KEEPER as an OHDSI package



Input:

1) Cohort defined by the phenotyping algorithm
2) Concept_ids for the clinical elements you want to see

Output:

Record per patient to examine and label

More in the OHDSI Symposium 2023 Plenary

Plenary: Improving the reliability and scale of case validation

Case validation is regarded as a necessary element of regulatory-grade evidence, but conducting case validation through human adjudication of source records is time- and resource-intensive, has unknown performance, and is frequently conducted in such a way that does not enable either full caseset review or proper quantitative bias analysis. In this plenary, OHDSI collaborators presented innovative methodological research and open-source development to improve the reliability and scalability of the case validation process, demonstrating that it may be possible to replace source records through an informatics-enhanced patient profile of structured data from the OMOP CDM (KEEPER), and to supplement human review through the use of large language models to estimate measurement error and identify differential misclassification. KEEPER + LLM was empirically evaluated in 10 diseases across 3 experiments in 2 different data sources, and revealed that there can be substantial heterogeneity in agreement between human reviewers but that LLMs agree with humans as much as humans agree with each other. **Speakers included: Patrick Ryan**, Johnson & Johnson, Columbia University; **Anna Ostropolets**, Odysseus Data Service; and **Martijn Schuemie**, Johnson & Johnson, University of California, Los Angeles

Plenary Slides

https://www.ohdsi.org/ohdsi2023/