Impact of participant–, clinician–, and EHR–derived phenotyping for rare disease diagnosis

Rachel A. Ungar*, **Jordan L. Cahoon***, Chloe Reuter, Emily Alsentzer, Meghan Halley, Stephen B. Montgomery

*These authors contributed equally to this work

Context

Motivation

- Rare diseases collectively affects 300 million patients worldwide, predominantly affecting children [1].
- Diagnosis is critical for effective symptom management, risk prognosis, and treatment development [2].
- Best practice in diagnosis relies on clinician review to collect Human Phenotype Ontology (HPO) terms, creating a bottleneck that limits scalability and consistency.

Goals

- We explore two alternatives to clinician to phenotype collection:
  - Direct reporting from patients
  - Automatic extraction from electronic health records (EHRs)
- We hypothesize that collaborative phenotyping can improve performance on diagnostically relevant tasks.

Figure 1: Mixed methods study overview

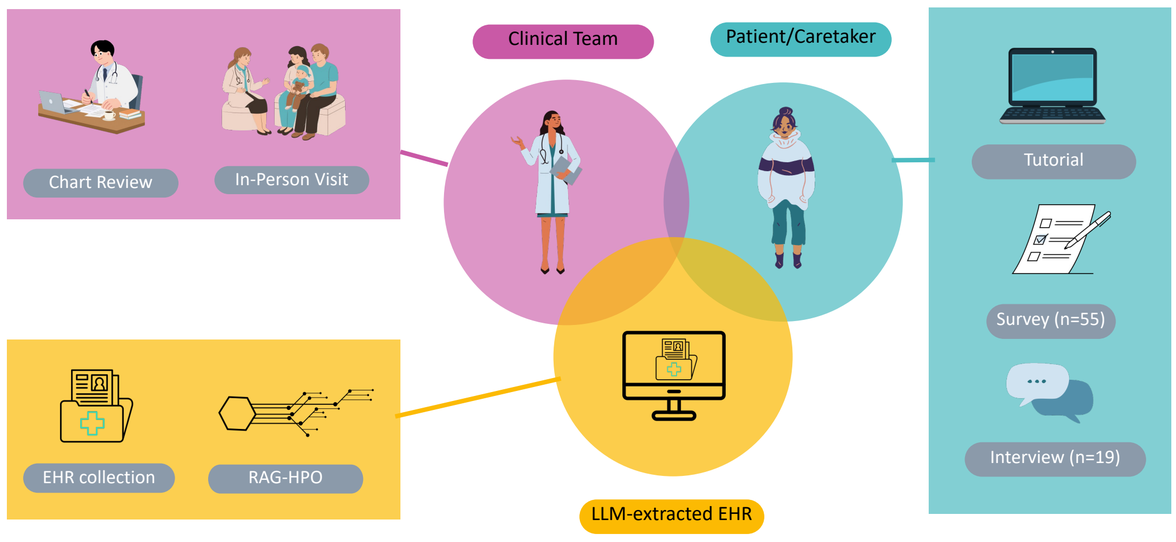


Figure 2: Number and type of HPO terms reported by different sources

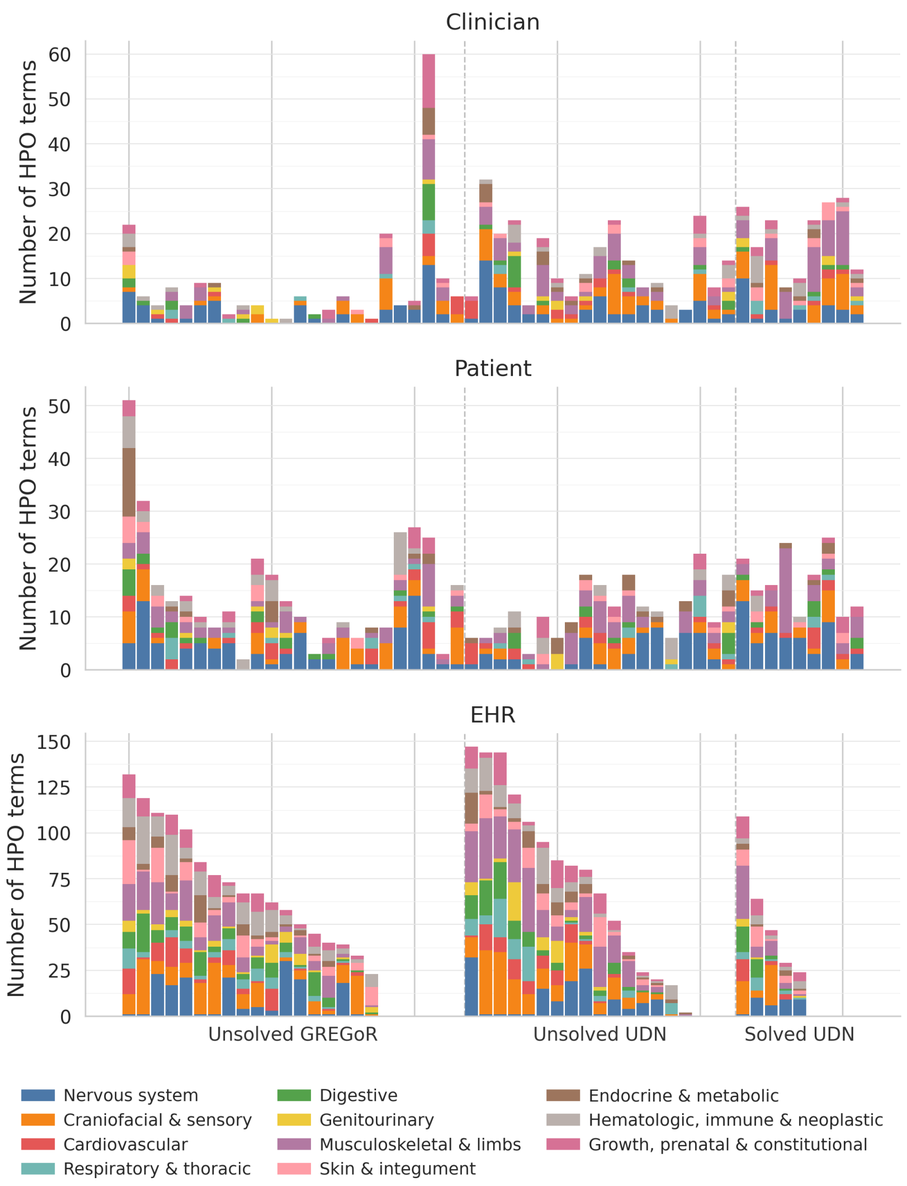
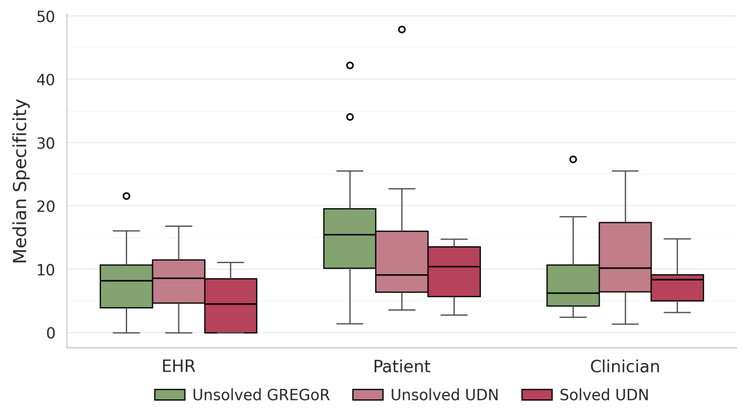


Figure 3: Disease specificity of HPO terms across patient groups



HPO Term Collection

Patients

- Created survey with self–phenotyping video instructions for patients participating in the Undiagnosed Disease Network (UDN) [3] and Genomics Research to Elucidate the Genetics of Rare Diseases (GREGoR) Consortium [4].
- 55 patients completed the survey, of which 9 were UDN participants who knew their diagnoses, 20 were UDN participants who did not know their diagnoses, and 28 were GREGoR participants who did not know their diagnoses.
- 15 patients were interviewed about their self phenotyping experiences

EHR

- Extracted 16,089 notes across 52 patients (mean notes per patient = 309, SD = 438) from an Epic Clarity Database
- A large–language–model–based approach [5] was used to extract HPO terms directly from clinician notes

Figure 4: Combined phenotyping improves ranking in diagnosed patients

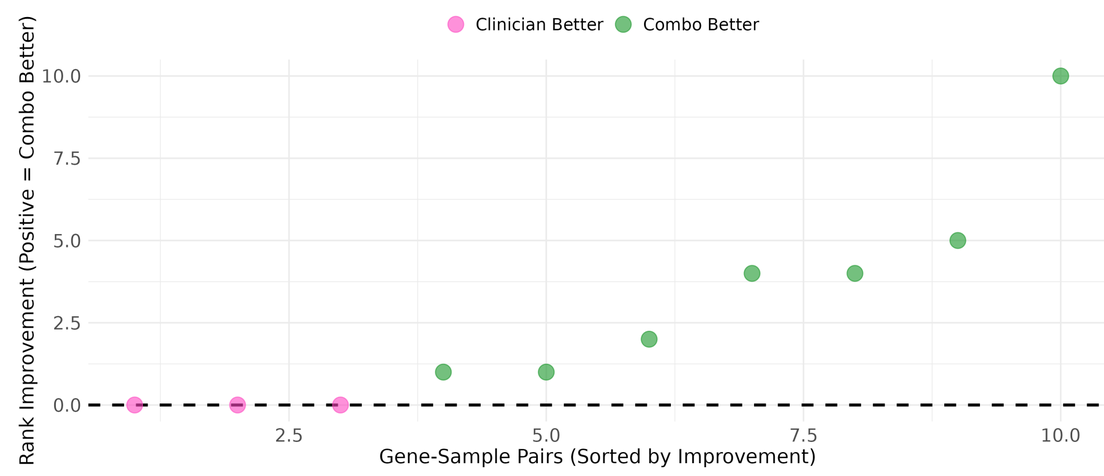
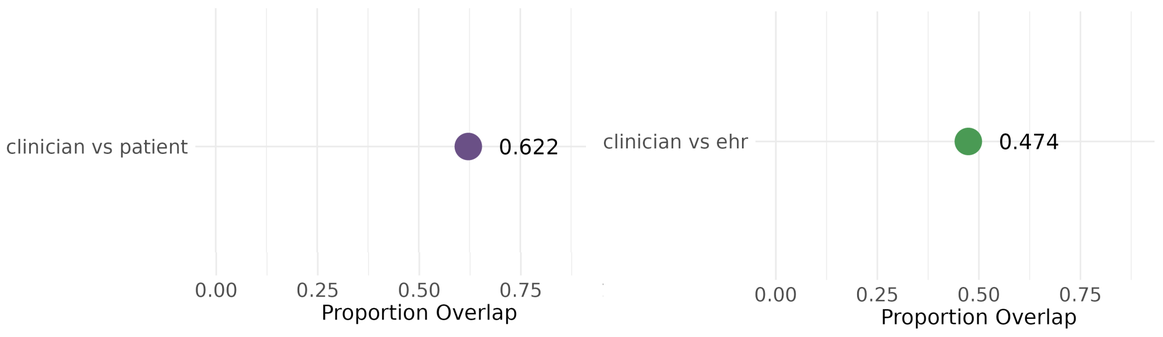


Figure 5: Proportion of overlapping genes in top 30 ranked genes



Evaluation

- HPO term lists were evaluated based on the number, type, and disease specificity (as measured by the number of gene–phenotype relationships in PrimeKG[6]).
- Exomiser was used to conduct causal gene ranking for diagnosed UDN patients using variant information and HPO terms.

Future Work

- Complete interview transcript qualitative analysis to characterize self phenotyping experiences
- Prospectively identify cases where added phenotypes identified plausible diagnoses

Sources [1] The Lancet Global Health. The landscape for rare diseases in 2024. Lancet Glob Health. 2024 Mar;12(3):e341. doi: 10.1016/S2214–109X(24)00056–1. [2] Glaubitz R et al. The cost of the diagnostic odyssey of patients with suspected rare diseases. Orphanet J Rare Dis. 2025 May 10;20(1):222. doi: 10.1186/s13023–025–03751–y. [3] Ramoni RB et al. The Undiagnosed Diseases Network: accelerating discovery about health and disease. Am J Hum Genet. 2017;100(2):185–192. doi:10.1016/j.ajhg.2017.01.006. [4] Dawood M et al. GREGoR: accelerating genomics for rare diseases. Nature. 2025;647(8089):331–342. doi:10.1038/s41586–025–09613–8. [5] Garcia BT et al. Improving Automated Deep Phenotyping Through Large Language Models Using Retrieval Augmented Generation. : Genome Med. 2025 Aug 18;17(1):91. doi: 10.1186/s13073–025–01521–w. [6] Chandak, P., Huang, K. & Zitnik, M. Building a knowledge graph to enable precision medicine. *Sci Data* 10, 67 (2023)