

Extraction of Crohn's Disease Clinical Phenotypes from Clinical Text Using Natural Language Processing

Linea Schmidt^{a,b,c,*}, Susanne Ibing^{a,b,c,1,*}, Florian Borchert^a,
Julian Hugo^{a,b,c}, Allison Marshall^d, Jellyana Peraza^e, Judy H. Cho^e,
Erwin P. Böttinger^{a,b,c}, Ryan C. Ungaro^{f,1}

^a Hasso Plattner Institute, Digital Engineering Faculty, University of Potsdam, Germany

^b Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, USA

^c Windreich Dept. of Artificial Intelligence & Human Health, Icahn School of Medicine at Mount Sinai, New York, USA

^d Department of Medicine, Mount Sinai Health System, New York, USA

^e Department of Pathology, Molecular, and Cell Based Medicine, Icahn School of Medicine at Mount Sinai, New York, USA

^f The Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, USA

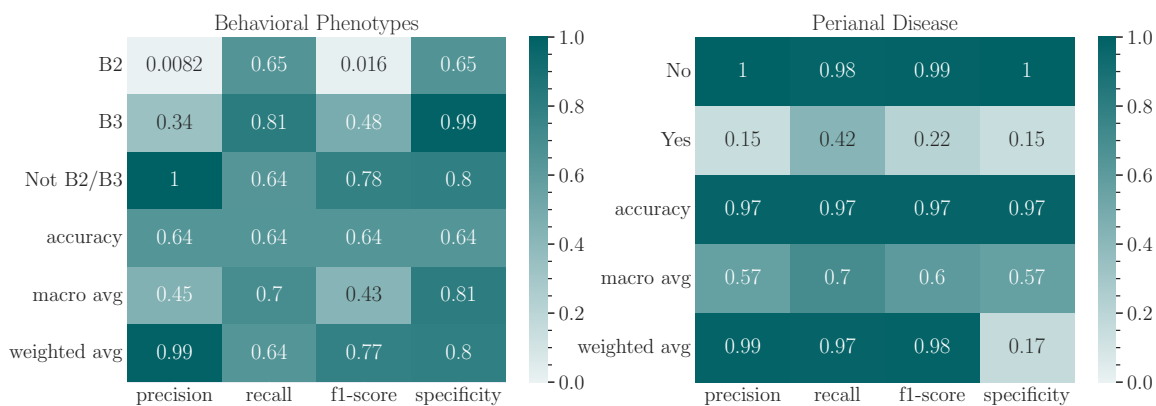
¹ Corresponding authors:

Susanne Ibing, susanne.ibing@hpi.de

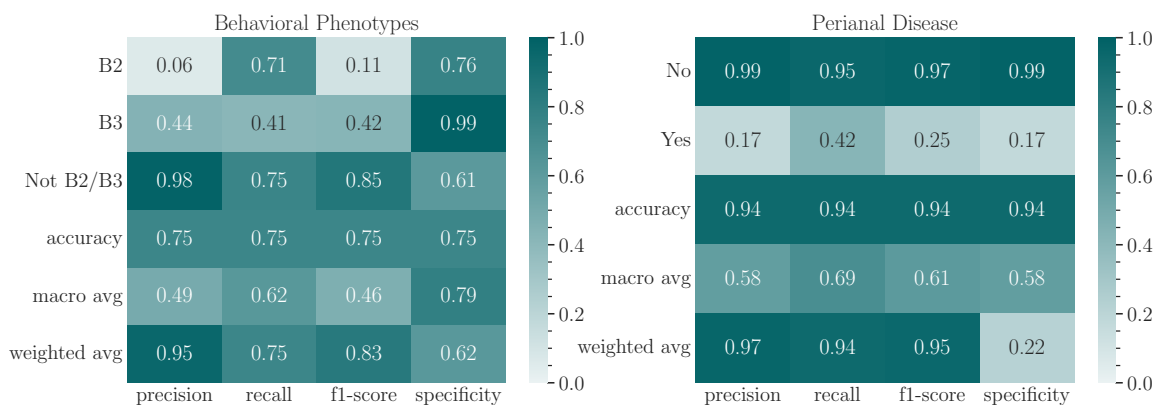
Ryan C. Ungaro, ryan.ungaro@mssm.edu

* These authors contributed equally to this work.

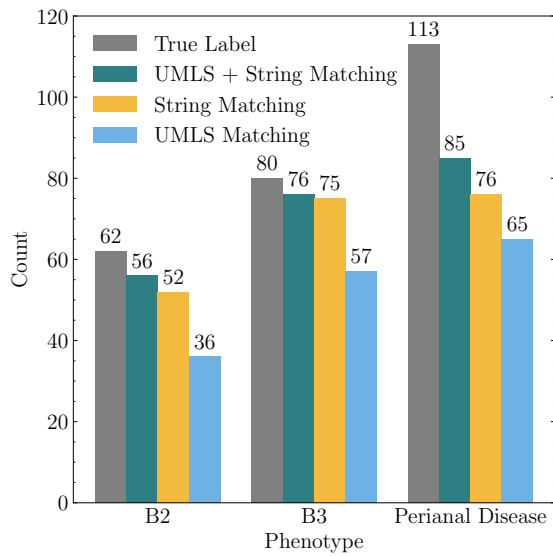
1 Supplementary Figures



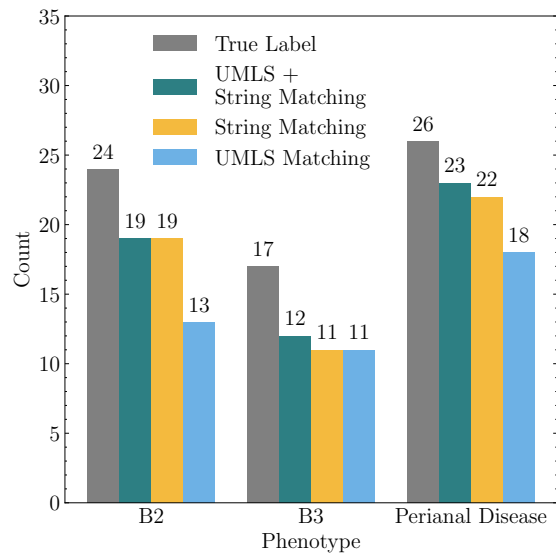
Suppl. Figure 1: Performance of zero-shot transformer on sentence-level based on clinical notes



Suppl. Figure 2: Performance of zero-shot transformer on sentence-level based on radiology reports

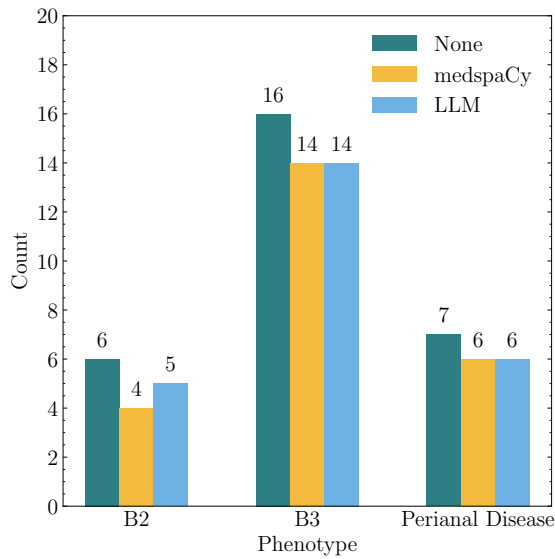


(a) Clinical Notes

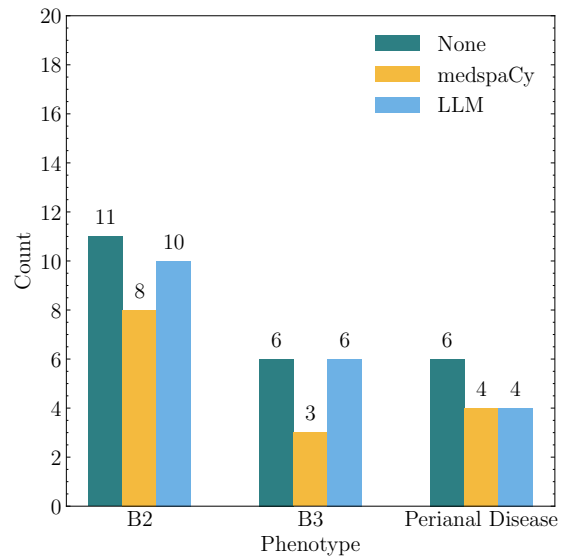


(b) Radiology Reports

Suppl. Figure 3: Correctly identified phenotypes by patterns. The count of correctly identified phenotypes by different usage of patterns versus the true count per phenotype on sentence level in the annotated clinical notes and radiology reports. Unified Medical Language System (UMLS) matching refers to patterns using the matched UMLS codes, while string matching refers to patterns manually created to match specific phenotypes.

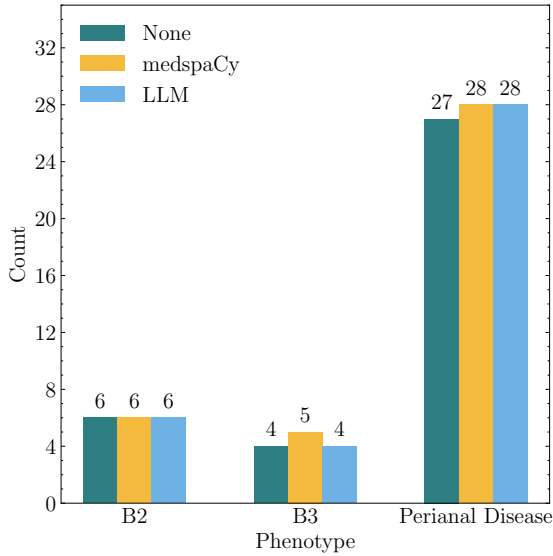


(a) Clinical Notes

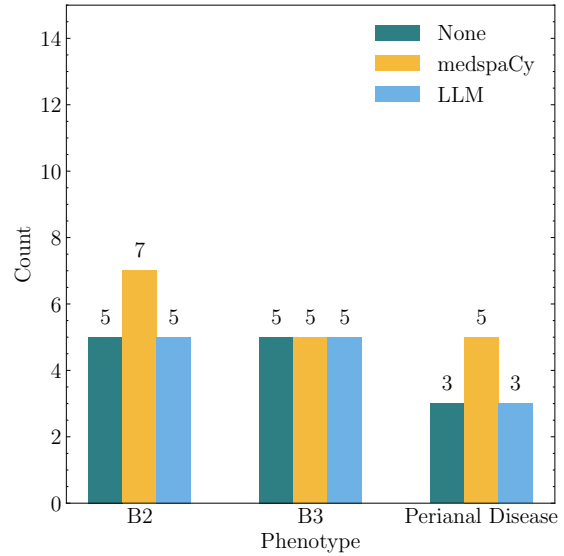


(b) Radiology Reports

Suppl. Figure 4: Incorrect identified phenotypes by negation detection methods. Differences in negation detection methods when analyzing the count of incorrectly identified phenotypes in the annotated clinical notes and radiology reports. "None" means, that no additional negation detection besides the manually defined rules for uncertainty and exclusion is used. "medspaCy" refers to negation detection based on the medspaCy ConText component, and "LLM" refers to negation detection via the clinical-assertion-negation-bert classifier.



(a) Clinical Notes



(b) Radiology Reports

Suppl. Figure 5: False negative phenotypes by negation detection methods. Differences in negation detection methods when analyzing the count of false negative phenotypes in the annotated clinical notes and radiology reports. "None" means, that no additional negation detection besides the manually defined rules for uncertainty and exclusion is used. "medspaCy" refers to negation detection based on the medspaCy ConText component, and "LLM" refers to negation detection via the clinical-assertion-negation-bert classifier.