

Brandão AAGS¹, Leão D¹, Konopelko M², Amtmaski W², Basak G^{2,3}, Verboski BM¹, Assunção CL¹, Lis K², Dabrowski MJ^{2,4}, Dudziski M^{2,5}
¹Hospital BP - A Beneficência Portuguesa de São Paulo, São Paulo, SP, Brasil ²Saventic Health, Torun, Poland ³Department of Hematology, Transplantation and Internal Medicine, Medical University of Warsaw, Warsaw, Poland ⁴Institute of Computer Science of the Polish Academy of Sciences, Computational Biology Group, Warsaw, Poland ⁵University Teaching Hospital in Rzeszow, Rzeszow, Poland

INTRODUCTION

- Idiopathic multicentric Castleman's disease (iMCD): ultra rare lymphoproliferative disorder.
- iMCD: difficult to diagnose due to its rarity (prevalence of 6.9-9.7 cases per million in the United States) and the lack of awareness among physicians.
- We propose that an artificial intelligence (AI)-based algorithm working on electronic health records may help to identify patients suspected of iMCD that were missed during medical evaluation.

METHODS

- Symptoms listed in the CDCN iMCD criteria were defined by medical terms. Medical definition contained ICD - 10 codes, lab test results and phrases that may be found in the medical texts (descriptions of imaging tests, interviews, epicrisis).

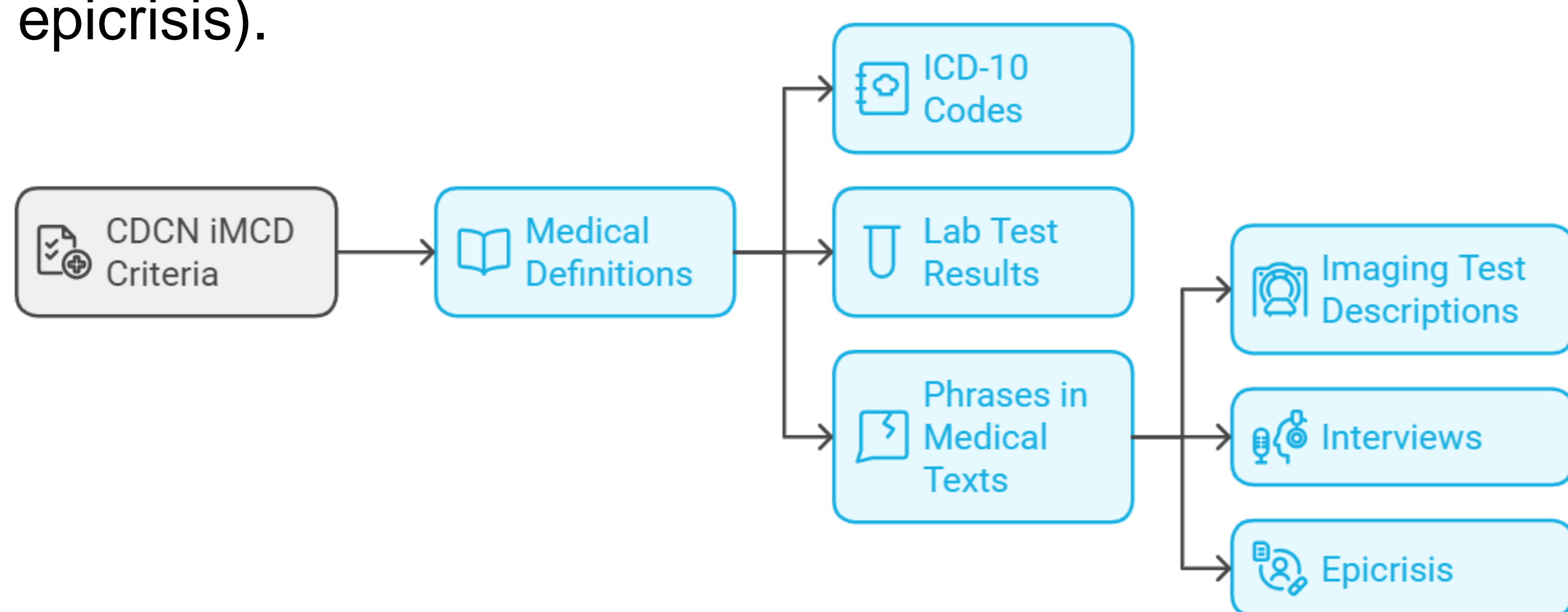


Figure 1: Medical definitions process

- SPACY models with in house developed regular expressions (regex) were used to identify phrases (entities) assigned to symptoms in the unstructured medical texts.
- Entity context defined as negation, family relation, historical or hypothetical event was determined by the BERTH model that went through additional training on the Saventic database.

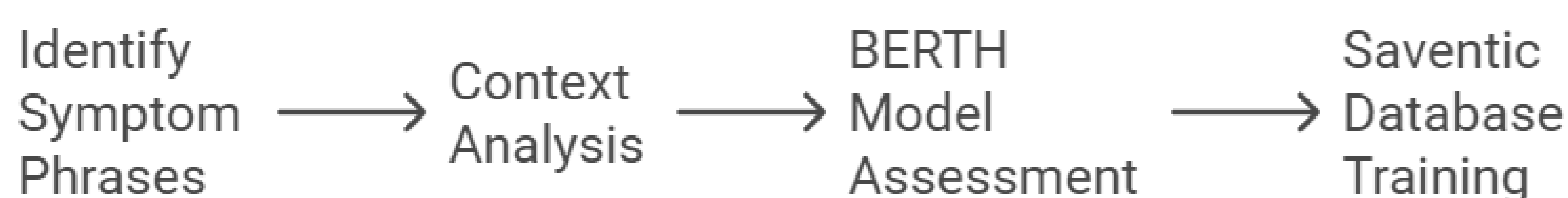


Figure 2: NLP construction

- Patients were ranked based on the symptoms they possessed. Patients were ranked only when lymphadenopathy was identified, and patients already diagnosed with iMCD (four cases) were excluded. Physicians with expertise in iMCD led the review of patients' charts with the highest score.

RESULTS

- 594,953 patients' charts from a tertiary Brazilian hospital were screened by the algorithm.
- iMCD clinical and/or laboratory features were identified in 102 patients (higher probability of having iMCD). 12 had mimicker conditions with overlapping iMCD features and three others required further assessment.

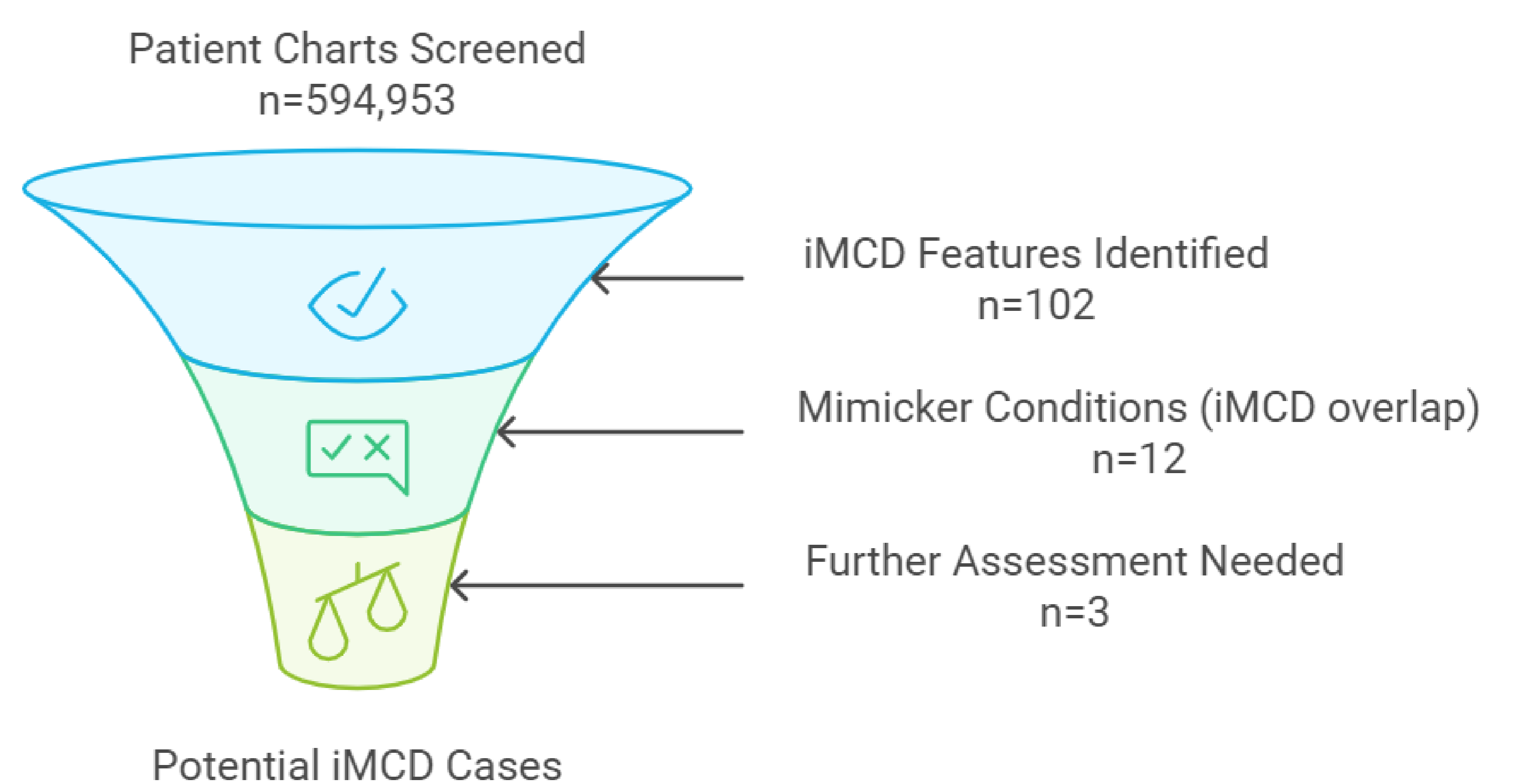


Figure 3: Screening process

CONCLUSIONS

- This AI driven tool was able to identify patients who have iMCD features.
- Screening with this tool found at least three candidates for additional evaluation.
- In the future, algorithms like the one we present may facilitate screening for iMCD.

REFERENCES: Fajgenbaum DC, Uldrick TS, Bagg A, et al. International, evidence-based consensus diagnostic criteria for HHV-8-negative/idiopathic multicentric Castleman disease. *Blood*. 2017;129(12):1646-57.

CONTACT INFORMATION: antonio.brandao@hc.fm.usp.br