

CTInfer: Natural Language Inference for clinical trials

Mathilde Aguiar, Pierre Zweigenbaum, Nona Naderi

{first, last}@liscn.upsaclay.fr

Université-Paris-Saclay, CNRS, Laboratoire Interdisciplinaire des Sciences du Numérique



Introduction

- Example: **Hypothesis:** *The patient does not suffer from diabetes.*
Premise: *The patient took 1 unit of insulin at 7:00 a.m.*
→ **Contradiction**
- **Objective:** Determine techniques, resources and models currently available to apply Natural Language Inference (NLI) in a clinical context, as well as tackling its gaps and challenges.

Research questions

- 1: What are the existing methods, models and datasets available to resolve the NLI task in the clinical domain?
- 2: How can NLI be beneficial to the clinical domain?
- 3: What are the gaps of NLI applied to the clinical domain?

Methods

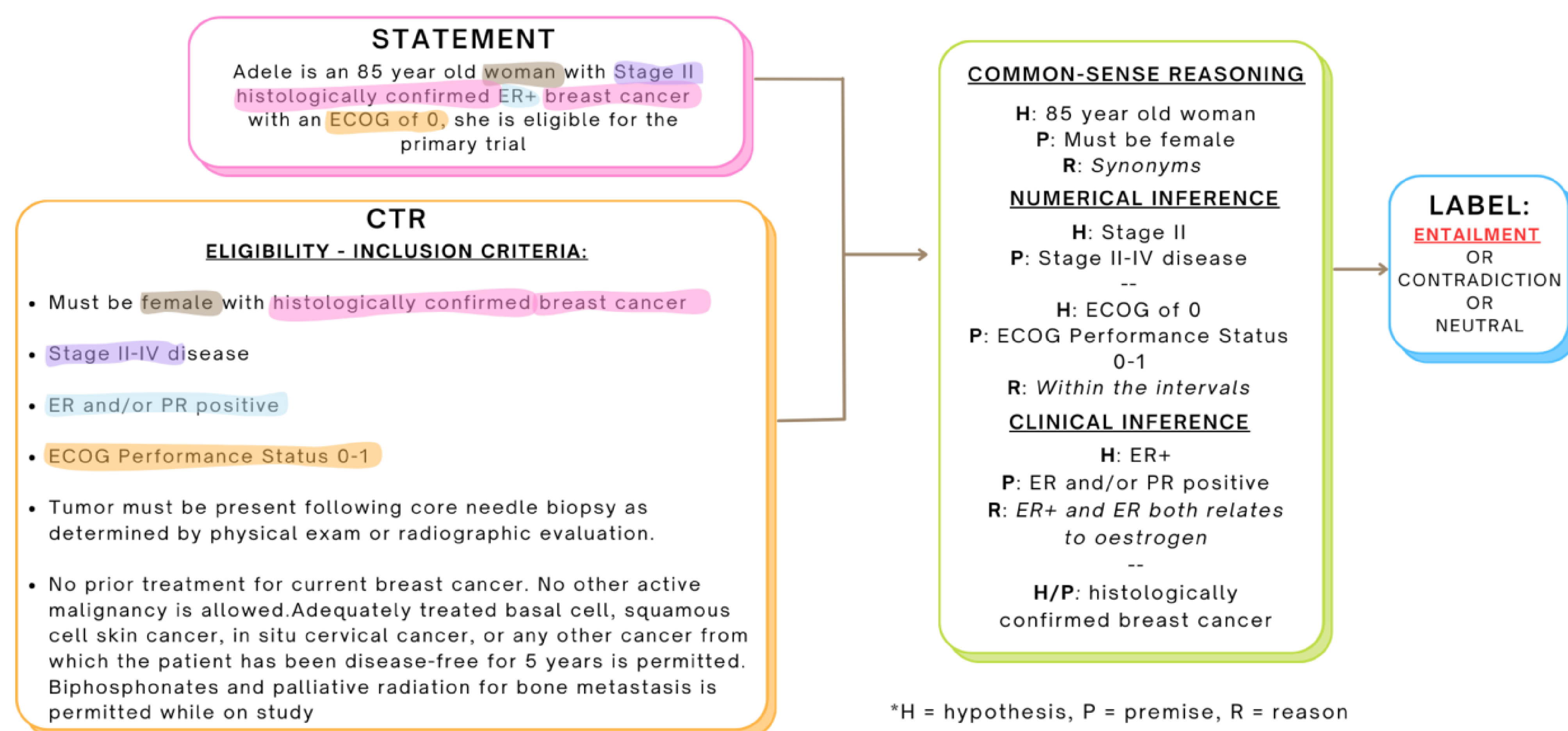
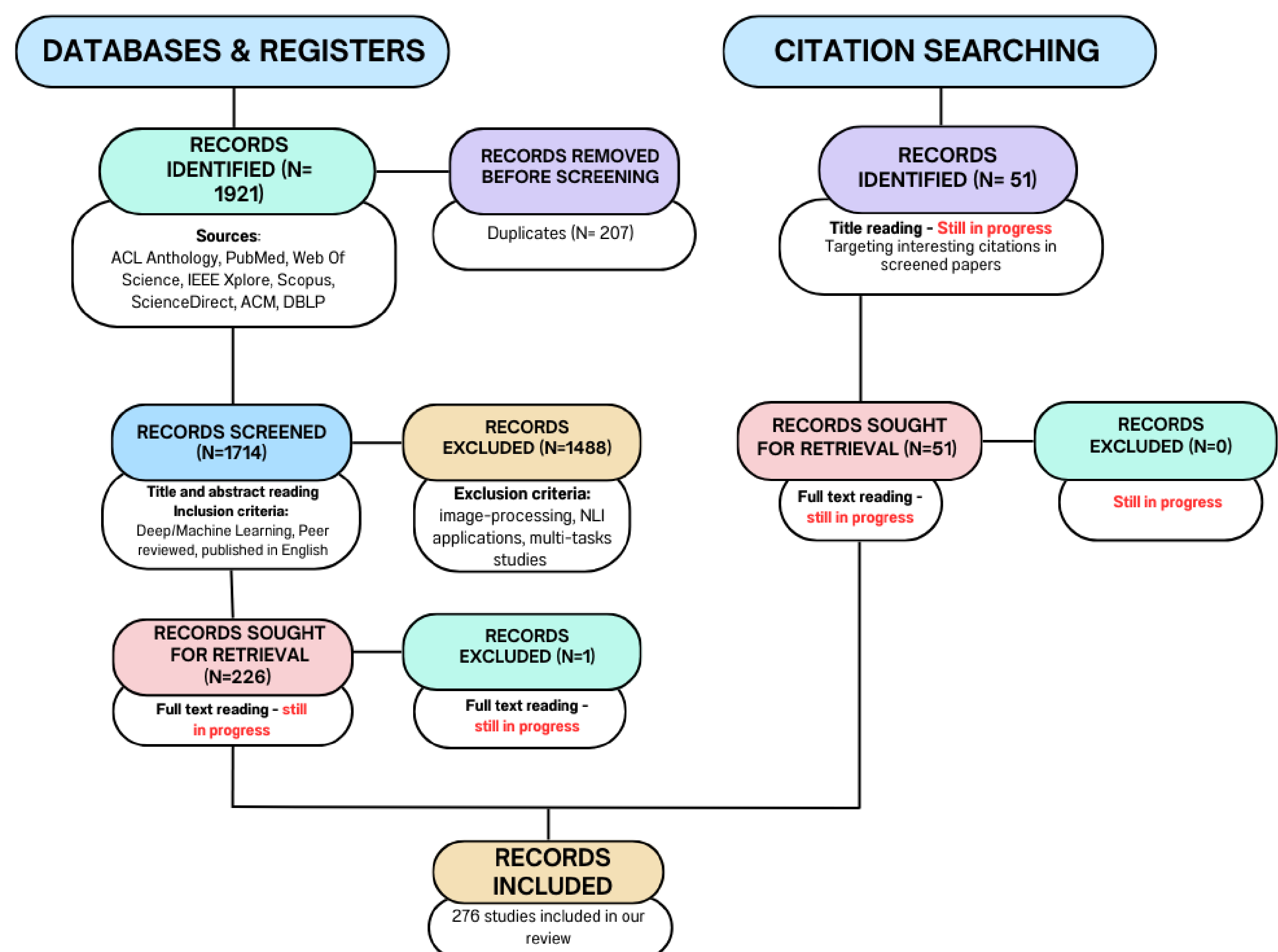
- PRISMA [1] protocol to conduct the systematic review.
- Databases screened:



Queries:

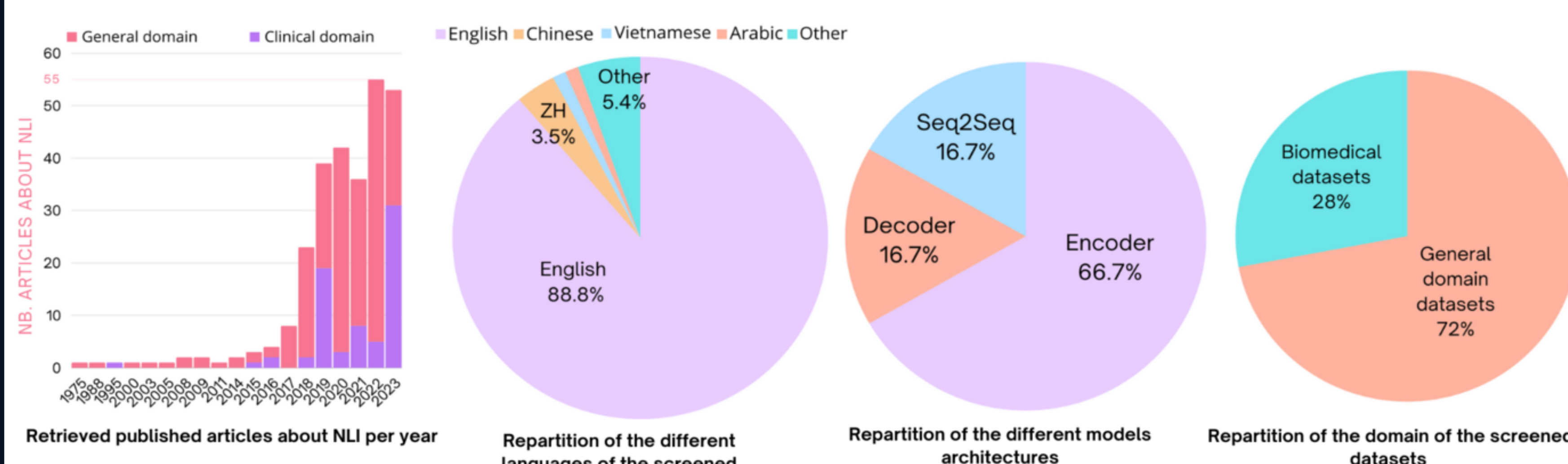
1. *clinical AND "Natural Language Processing" AND "Natural Language Inference" OR NLI*
2. *clinical AND "Textual Entailment" OR TE*

- **Inclusion criteria:** Deep/machine learning, peer reviewed paper, published in English
- **Exclusion criteria:** Multimodal/ image-processing, NLI applications, multi-tasks studies



Results

1700 records screened and 276 studies included in the systematic review.



Conclusion

- Complex task requiring different kinds of reasoning (real-world, numerical, domain-specific, etc.).
- Lack of non-English datasets.
- Lack of domain-specific datasets.

References

- [1] Page et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372:n71, 03 2021.

Project funded by CNRS
80|PRIME grant.

