JEBS
A Fine-grained Biomedical Lexical Simplification Task

William Xia, Ishita Unde, Brian Ondov, Dina Demner-Fushman

Tufts University, Johns Hopkins University, Yale University, National Library of Medicine

INTRODUCTION

Motivation

- Medical literacy is critical for making informed healthcare decisions, but is challenging due to jargon.
- Past solutions prioritize end-to-end simplification, which can easily distort or drop important information.

Jargon Explanations for Biomedical Simplification (JEBS)

- We divide term simplification into 3 sub-tasks:

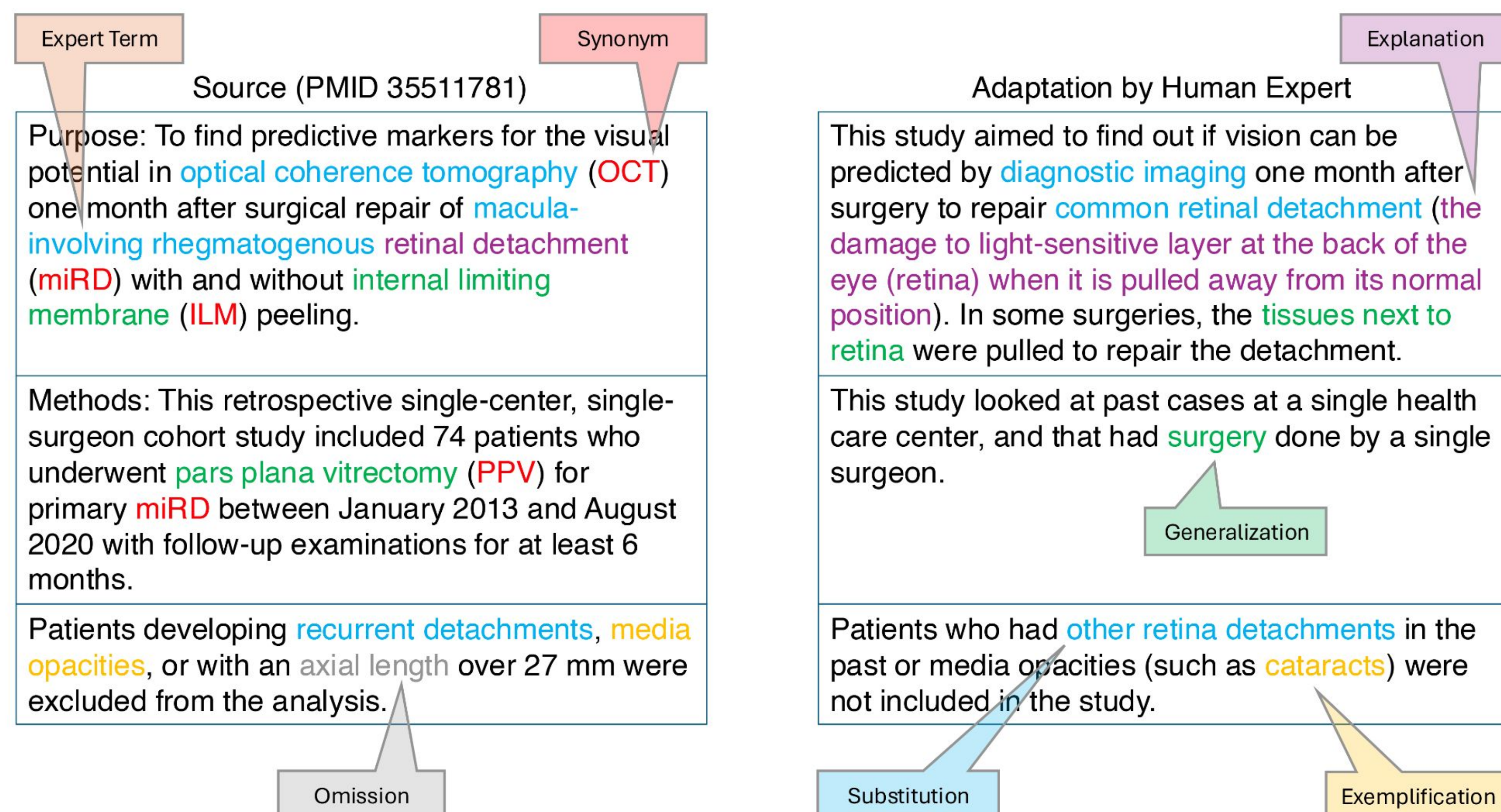| 1 Identify | ⇨ | 2 Classify | ⇨ | 3 Generate |

TASK DEFINITION

Figure 1. Demonstration of how expert terms and their in-text synonyms were annotated according to 5 types of simplification.

DATASET & BASELINE MODELS

Dataset Overview

- JEBS is a dataset for targeted, term-level biomedical simplification.
- Derived from 400 abstracts and ~1,200 adaptations from PLABA (Attal et al. 2023).
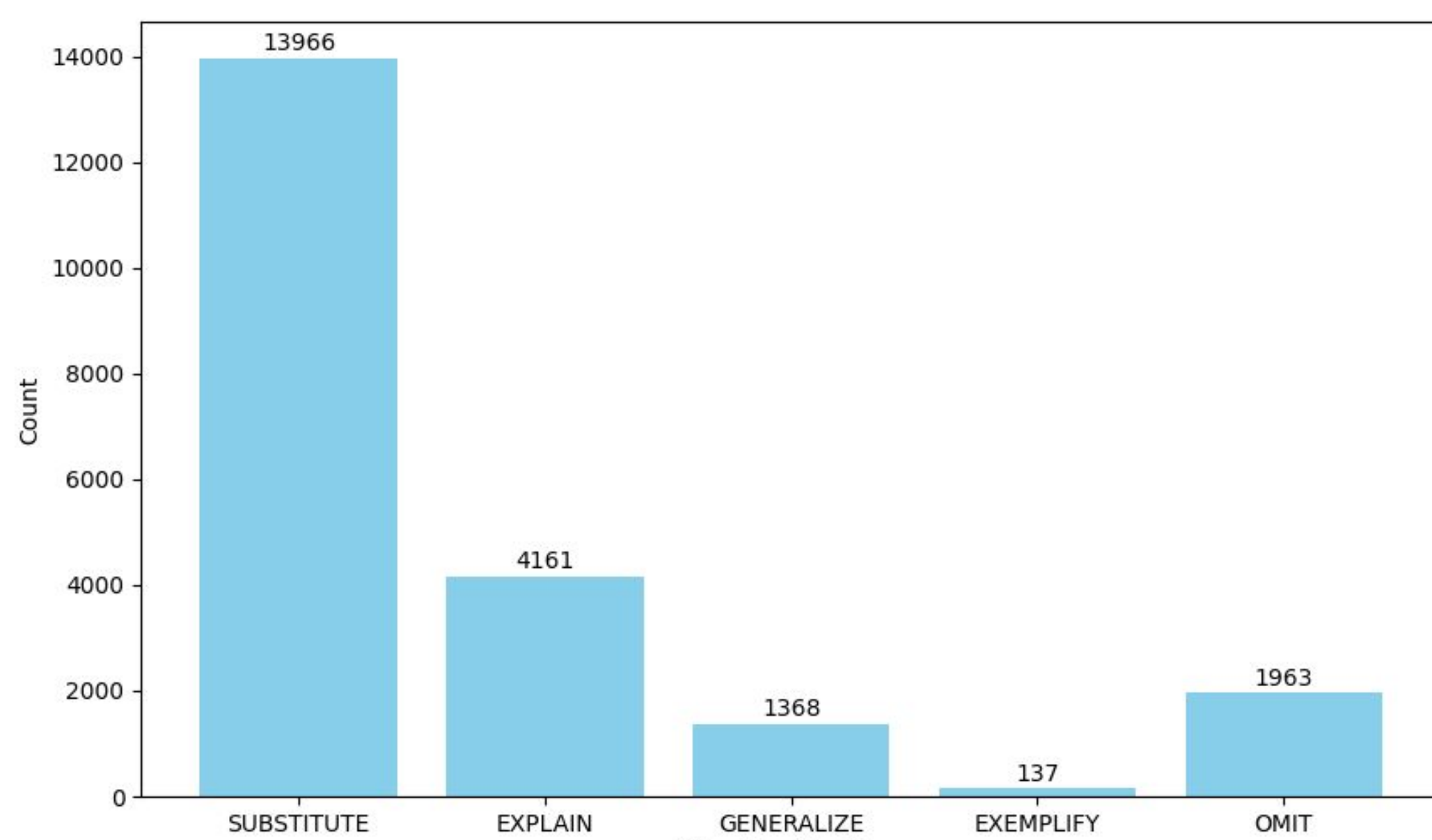- Manual annators (IAA: 0.52 F1) identified 21,595 simplifications for 10,314 expert terms.

Figure 2. Distribution of simplification types.

Sub-task 1: Identification

- Framed as a named entity recognition (NER) task.

```
"Patients received a ring suture"
   O       O     O   B    I
```

Model	Avg F1	Union F1	Pyramid
MetaMapLite	0.2097	0.2487	0.2916
DeBERTa	**0.4317**	**0.5255**	0.6014
Llama3	0.3678	0.4085	0.4692
End-to-end	0.3923	0.3918	**0.6207**

Sub-task 2: Classification

Model	Avg F1	Union F1
BERT Frozen	0.0337	0.0334
DeBERTa Frozen	0.1823	0.1856
BERT NER	**0.3588**	**0.3413**
DeBERTa NER	0.3300	0.3363

Sub-task 3: Generation

- Manual evaluations across 4 axes:
  - **Simplicity:** Simplification is easy to understand.
  - **Accuracy:** Simplification contains correct information.
  - **Completeness:** Simplification minimizes information lost from the original text.
  - **Brevity:** Simplification is concise.

Model	SUB	EXP	GEN	EXE	OMI
Llama3 $_{1\text{-}shot}$	0.788	0.465	0.840	0.740	0.701
Llama3 $_{3\text{-}shot}$	0.853	0.492	0.805	0.767	0.747
Llama4 $_{1\text{-}shot}$	0.764	0.879	**0.927**	0.548	0.697
Llama4 $_{3\text{-}shot}$	0.801	**0.918**	0.906	0.514	0.695
GPT-4o $_{1\text{-}shot}$	0.833	0.514	0.904	0.801	0.745
GPT-4o $_{3\text{-}shot}$	**0.922**	0.544	0.855	**0.822**	**0.754**

Table. Aggregate results of each baseline model on each simplification type.

CONCLUSION & FUTURE WORK

JEBS enables precise and interpretable biomedical simplification by focusing on individual terms, providing a unique resource for training and benchmarking future biomedical simplification tools.

Future work may consider grammar-aware omission models, addressing the issue of class imbalance in our dataset, or performing more rigorous evaluation of an end-to-end text simplification tool based on the JEBS task.

JEBS Dataset Release

NIH National Library of Medicine