TrialFORMA: Formal Reasoning and Modelling Agent with an Application to Clinical Trial Matching

Cyrus Z Zhou*, Yufei Jin†, Yilin Xu‡, Chieh-Ju Chao*§, Monica S Lam*

*Stanford University §Mayo Clinic †UCLA ‡JHU

AI+ HEALTH
Stanford University • Online Conference

Motivation

Clinical trials are the driving force behind the development of new treatments.
Recruiting patients for clinical trials **is hard.**

Stats on recruitment difficulty:

💰 **1.9B USD** spent annually [1],
😞 **80% of trials still** miss recruitment targets [2].

Consequences of recruitment failures:

⚠️ Money burns.
⚠️ Scientific progress stalls.
⚠️ Patients miss life-saving opportunities.

Why recruitment is difficult:

🚧 Clinical trial eligibility criteria are complex.
🚧 Search systems not efficient and usable.
🚧 Heavily rely on labor-intensive, manual search.

Prior Works

***Problem:* Finding clinical trials for a patient.**

Prior Works: e.g., *TrialGPT* [3]
Dense & Keyword Retrieval + LLM Checks

Shortcomings:

❌ Retrieval based on text similarity, not logic.
❌ Black box, cannot be audited.
❌ Hard to be adapted for real deployment.

Our Approach

TrialFORMA turns free-text trial criteria and patient records into a **shared symbolic representation,** uses SQL to quickly retrieve candidate trials, and applies formal logic (SMT) to soundly decide eligibility with clear, auditable explanations.

References:

[1] Brøgger-Mikkelsen, Mette, et al. "Online patient recruitment in clinical trials: systematic review and meta-analysis." Journal of medical Internet research 22.11 (2020): e22179.

[2] Desai, Mira. "Recruitment and retention of participants in clinical studies: critical issues and challenges." Perspectives in Clinical Research 11.2 (2020): 51-53.

[3] Jin, Qiao, et al. "Matching patients to clinical trials with large language models." Nature communications 15.1 (2024): 9074.

[4] Koopman, Bevan, and Guido Zuccon. "A test collection for matching patients to clinical trials." Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. 2016.

Goals and Problems

❓ **Embedding-Based Retrieval Approach:**
🎯 **Minimize False Negatives and False Positives.**

✅ **Our approach (Formal Semantics, SMT):**
🎯 **Eliminate False Negatives and False Positives.**

Problems?

🔍 **Errors in symbol canonicalization.**
Same concept, different symbols → no match.

🔍 **Errors in semantic parsing.**
NL criteria mistranslated or dropped in SMT.

🔍 **Incompleteness in ontology.**
Symbol canonicalization and entailment break.

Solution: A Family of Agentic Workflows

FORMA Workflow #1: Precise and comprehensive ***symbol canonicalization*** linked to a unified terminology system (e.g., UMLS).

Entity Span Identification → **Dense Retrieval in UMLS ontology** → **Concept Reranking & Filtering**

FORMA Workflow #2: *Semantic parsing* of clinical trials and patient records to ***SMT and formal representations*.**

Free Text Preprocessing → **Canonicalized Concept Schematization** → **SMT Programming**

FORMA Workflow #3: *Curation of incomplete relationships* in the ontology.

Dense Candidate Concept Search → **Src-Candidate Relationship Identification** → **Relationship Persistence**

System Optimization

Core Idea: SQL as an efficient prefilter for SMT

Satisfiability Modulo Theories (SMT)

↓ **Projection (Quantifier Elimination)**

Propositional Logic (PL)

↓ **Expression Normalization Database Ingestion**

Queryable Database Tables

Preliminary Evaluation & Results

Dataset:

- SIGIR dataset [4], **59** patients, **3621** trials.

Results:

- **Recall**
  - TrialFORMA is able to find **all** trials for 8 sampled patients through *manual checking*. (105 trials fetched and not explicitly contradicted)
  - On all 59 patients with *LLM judge*, TrialFORMA has an adjusted recall of 96.2% with about 140 trials that are not explicitly contradicted fetched per patient.
- **Precision** (4-patient sample; 10% of retrieved trials per patient; potentially relevant & eligible)
  - TrialFORMA: 89%
  - TrialGPT [3]: 56%
- **Speed** (M2 MacBook, SQLite, yet to be optimized)
  - 2.95s per patient against 3621 trials

Next Steps

🚀 **Enable the Trial2Patient Direction.**

🚀 **Integration with lengthy, heterogeneous, multimodal, real patient records.**

🚀 **A highly usable conversational assistant for patient-trial matching built on our backend.**

🚀 **Curation of more relationships among biomedical concepts (e.g., "may lead to").**