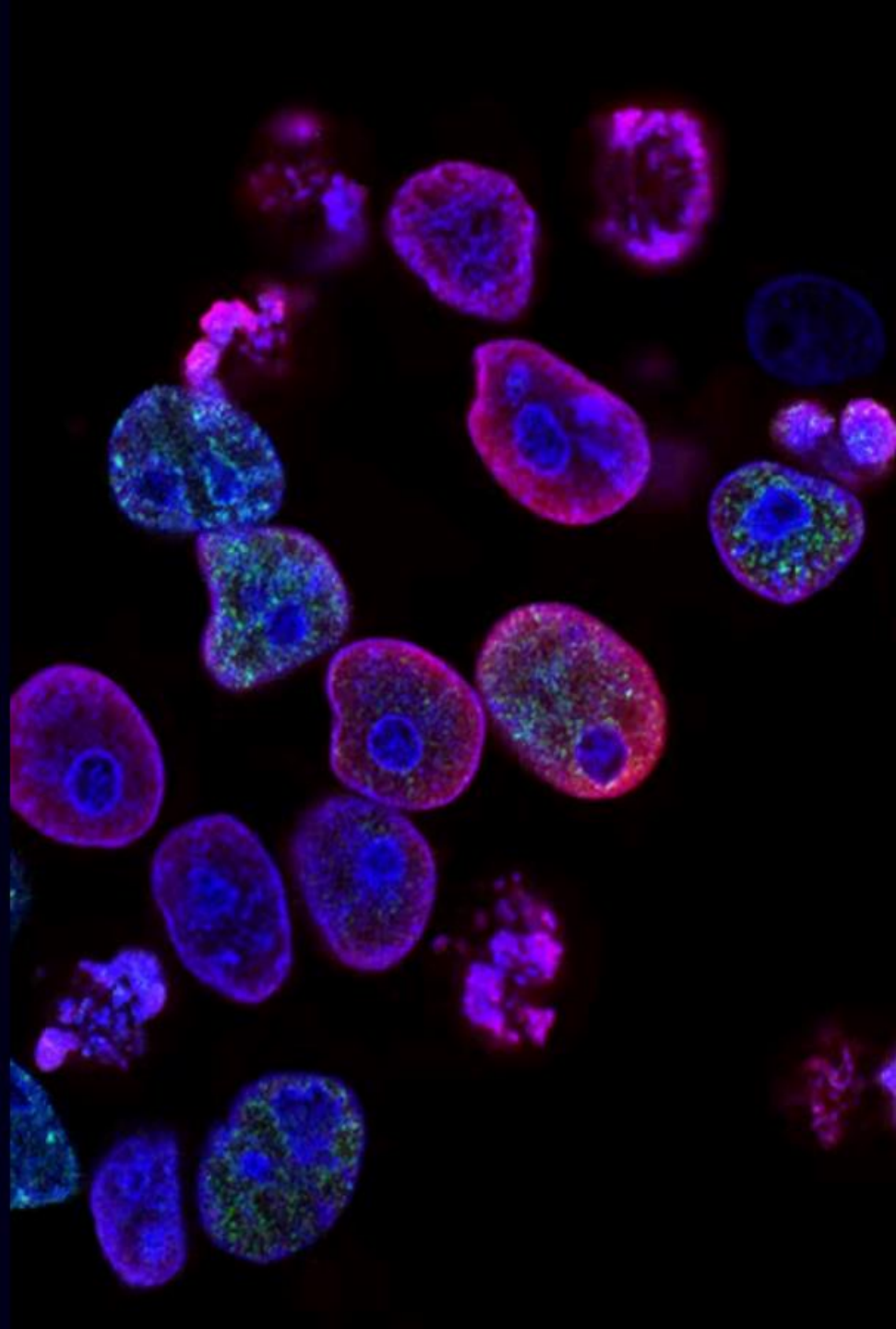
HelixRare: A Novel Approach for Clinical Trial Matching

Problem: Rare Disease Trial Recruitment

- Team DeepDx



Relevant Literature:

TrialGPT

AI model matching patients to clinical trials

Huntington's Disease

HTT gene contains a CAG trinucleotide repeat

GenomeEHR Integration

Genomic data's presence in EHRs is increasing

Data: Genomic Datasets

gnomAD

Publicly available genomic dataset.

ClinVar

Database of genetic variations and phenotypes.

VariationID	AlleleID(s)	Germline classification	Chromosome	Position	Reference	Alternate	Allele Frequency
1547158	1571701	Likely benign	4	3239993	C	T	7.494640e-06
1453852	1451378	Likely benign	4	3240008	T	C	8.467060e-04
1366273	1411615	Uncertain significance	4	3240012	C	T	1.436550e-05
1662541	1642764	Likely benign	4	3240020	G	A	6.248969e-07
1429507	1477240	Uncertain significance	4	3240030	C	T	5.008230e-06

Methods: AI-Driven Matching

1

Clinical Match

Rule-based system for eligibility criteria.

2

Variant Novelty

Assess novelty of patient genetic variants.

3

LLM Evaluation

LLM evaluates eligibility criteria.

Results

patientID	CAG_Repeats	Age	Travel_Preference
HD-P001	35	63	Yes
HD-P002	27	56	No
HD-P003	40	44	Yes
HD-P004	51	46	Yes
HD-P005	26	15	No
HD-P006	51	47	Yes
HD-P007	44	33	Yes
HD-P008	39	27	No
HD-P009	28	85	No
HD-P010	41	61	No

```
Patient: HD-P006, Score: 0.83
Patient: HD-P004, Score: 0.82
Patient: HD-P003, Score: 0.82
Patient: HD-P007, Score: 0.81
Patient: HD-P001, Score: 0.65
Patient: HD-P009, Score: 0.37
Patient: HD-P008, Score: 0.16
Patient: HD-P010, Score: 0.15
Patient: HD-P002, Score: 0.12
Patient: HD-P005, Score: 0.08
```

Clinical Criteria

Identifies potential participants based on criteria.

Novel Variants

Prioritizes patients with novel genetic variants.

patientID	Variants
HD-P001	[{'Chromosome': '4', 'Position': 3049495, 'Reference': 'A', 'Alternate': 'G'}]
HD-P002	[{'Chromosome': '4', 'Position': 3089481, 'Reference': 'T', 'Alternate': 'G'}]
HD-P003	[{'Chromosome': '4', 'Position': 3212885, 'Reference': 'A', 'Alternate': 'G'}]
HD-P004	[{'Chromosome': '4', 'Position': 3069192, 'Reference': 'A', 'Alternate': 'T'}]
HD-P005	[{'Chromosome': '4', 'Position': 3193646, 'Reference': 'G', 'Alternate': 'T'}]
HD-P006	[{'Chromosome': '4', 'Position': 3072925, 'Reference': 'T', 'Alternate': 'G'}]
HD-P007	[{'Chromosome': '4', 'Position': 3212948, 'Reference': 'C', 'Alternate': 'T'}]
HD-P008	[{'Chromosome': '4', 'Position': 3209548, 'Reference': 'T', 'Alternate': 'G'}]
HD-P009	[{'Chromosome': '4', 'Position': 3123508, 'Reference': 'A', 'Alternate': 'G'}]
HD-P010	[{'Chromosome': '4', 'Position': 3217129, 'Reference': 'C', 'Alternate': 'A'}]

Challenges & Future Directions

1. **EHR Integration**

   Integrate EHR data for patient profiles.

2. **Expand Database**

   Expand variant database.

3. **User Interface**

   Develop a user interface.

4. **Fine Tune**

   Fine tune the LLM prompt.

5. **Expert's Opinion**

   Seek expert opinions to refine & adjust the scoring weights

Conclusion

1. A novel and effective solution
2. Has the potential to discover rare variants
3. Disease and Trail specific model