# SimuTrial

## Improving Trial Outcomes

By Gnaneswar, Dhruv, Pranav, Krithish, Srikar, David, Om, Arjun, Arm

# Patient Recruitment!

*"Over **90% of new drugs fail** during clinical trials.*

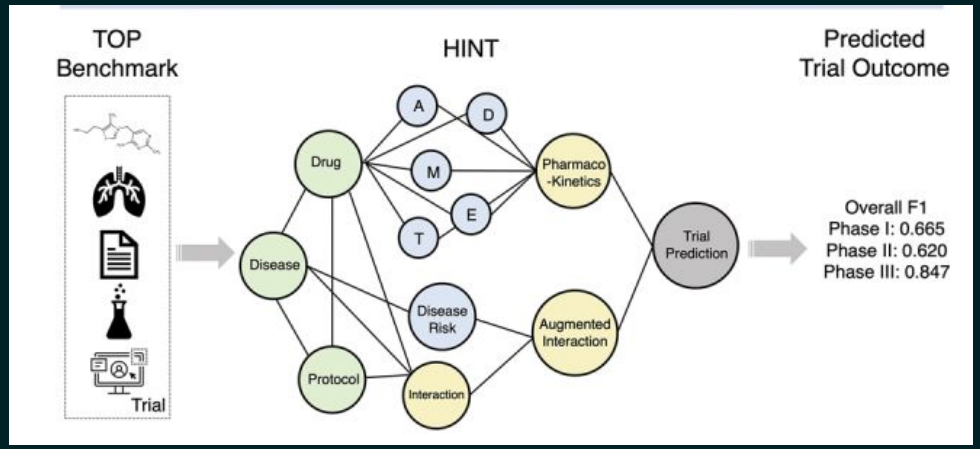
*The average cost of phase 1, 2, and 3 clinical trials across therapeutic areas is **around $4, 13, and 20 million** respectively. Pivotal studies for new drugs approved by the Food and Drug Administration (FDA) of the **United States cost a median of $41,117 per patient.***

*It costs an average of **$6,533 to recruit one patient** and **$19,533 to replace a patient** who drops out"*

# SOLUTION

?

Leverage Generative AI (GenAI) and Agent-Based Modeling (ABM) to streamline, simulate and optimize patient recruitment.

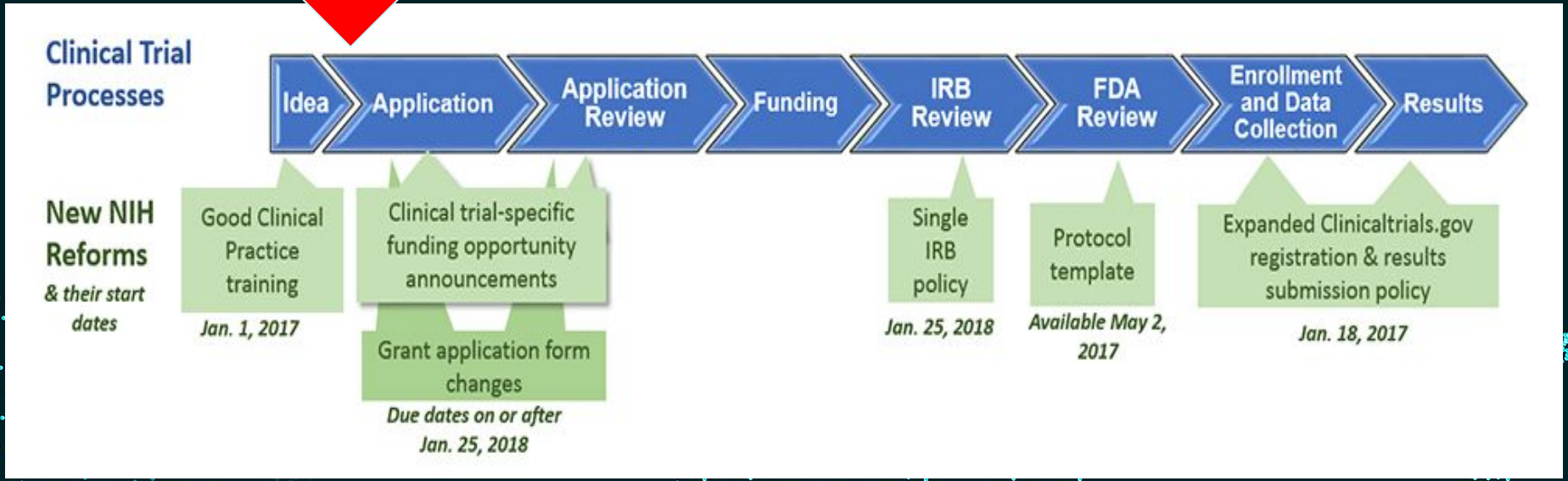



Logistic regression to predict willingness of patient to accept the trial invitation based on the data.

Our simulation/model will help clinicians plan and adjust their clinical trials and logistics based on past clinical trial results.

Data analytics tool for Clinicians based on previous trial data, patient trends and new data.

# This is where we fit in



# Patient Recruitment

- Patients must meet certain criteria to be eligible to participate.
- Successful recruiting leads to reliable and significant results for efficacy of therapies and devices

Current strategy involves defining criteria and employing outreach efforts to recruit.

- Eligibility criteria includes **age, gender, medical history, current condition, etc.**
- Participants are gathered with **advertisements, physician referrals, EMR's, and community postings.**
- Participants are enrolled after **screening, ensuring consent, and understanding purposes and risks involved.**

**Table 1: Race/Ethnicity of Participants in Pfizer-BioNTech and Moderna COVID-19 Vaccine Clinical Trials**

| | Total US Population Age 16+ | Pfizer-BioNTech* | Moderna |
|---|---|---|---|
| **Total** | **258 million** | **40,277** | **27,817** |
| **Race** | | | |
| White | 73.6% | 81.9% | 79.4% |
| Black | 12.3% | 9.8% | 9.7% |
| Asian | 5.9% | 4.4% | 4.7% |
| American Indian/Alaska Native | 0.8% | 0.6% | 0.8% |
| Native Hawaiian or Other Pacific Islander | 0.2% | 0.2% | 0.2% |
| **Ethnicity** | | | |
| Hispanic | 17.6% | 26.2% | 20.0% |
| Non-Hispanic | 82.4% | 73.2% | 79.1% |

Pfizer-Biontech, Moderna COVID-19 Vaccine

- Aimed to recruit over 44,000 patients; resulted in **~68,000 patients**
- Over **40%** U.S. participants were racially and ethnically diverse.
- **40% +** were between ages of 56 - 85 (Group with established COVID-19 mortality rates)

Figure: Artiga el al, Racial/ethnic distribution of total population age 16 or older based on KFF analysis, Briefing Document: Pfizer-BioNTech COVID-19 Vaccine

# Unmet Needs

- **Limited Awareness:** Most patients don’t know about clinical trials and their purpose.
- **Strict Criteria:** Exclusion of participants that may be willing.
- **Geography:** Patients may not live nearby academic centers or research institutions; may not be in an urban setting at all.
- **Retention:** Enrolled patients drop out often due to protocol issues, side effects, or inconvenience.
- **Bias:** Historically, trials underrepresented minorities = lack of tailored therapies and treatments.
- **Cultural Beliefs:** Certain communities are skeptical about medical research and ad’s fail due to cultural competence often.
- **Resources:** Cost and Time  Efficient recruitment leads to significant cost reduction and time reduction. (Clinical Trials are about 6-7 in tradition drug dev.)

# Example Use Case

Imagine that Clinician A wants to run a clinical trial for a new asthma treatment.

Step 1: Goes to SimuTrial and inputs patient data

Step 2: SimuTrial uses patient data to predict how successful a trial recruitment process will be by using ABM to create simulations based on patient demographics and behaviors to predict recruitment success across various populations.

Step 3: Clinician A can revise recruitment strategy (number of sites, site locations, marketing strategy, etc.)

# Input/Output

Simulation/Model
Agent-based modeling (ABM)
Bayesian logistic regression
Previous clinical trial Data

## Patient Data

| CENSREG | InvitedClinTrial | ParticipatedClinTrial | Age | BirthGender | Occupation_Employed | Occupation_1YUnemployed | C |
|---|---|---|---|---|---|---|---|
| 1 | 2 | -1 | -9 | 2 | 2 | 2 | |
| 3 | 2 | -1 | 74 | 2 | 2 | 2 | |
| 1 | 2 | -2 | 51 | 2 | 1 | 2 | |
| 1 | 2 | -1 | 31 | 2 | 1 | 2 | |
| 1 | 2 | -1 | 70 | 1 | 1 | 2 | |
| 2 | 2 | -1 | 64 | 1 | 2 | 2 | |
| 1 | 2 | -1 | 71 | 1 | 1 | 2 | |
| 2 | 1 | 2 | 67 | 1 | 1 | 2 | |
| 3 | 2 | -1 | 65 | 1 | 2 | 2 | |
| 3 | 2 | -1 | 62 | 1 | 2 | 2 | |
| 1 | 2 | -1 | 39 | 1 | 2 | 2 | |
| 3 | 2 | -1 | 83 | 1 | 2 | 2 | |
| 1 | 2 | -1 | 70 | 1 | 2 | 2 | |

Mean Consent Rate: 28.968142968142967%

Confidence Interval: +/- 0.09261065970225675

Average Staff Needed: 185.5

Average Sites Needed: 92.5

### Calculated Willingness Scores

| | Age | CENSREG | BirthGender | RaceEthn | WillingnessScore |
|---|---|---|---|---|---|
| 0 | 32 | 3 | 0 | -9 | 0.2332 |
| 1 | 29 | 3 | 0 | 4 | 0.3158 |
| 2 | 18 | 3 | 1 | -9 | 0.2332 |
| 3 | 41 | 3 | 1 | 2 | 0.3352 |
| 4 | 52 | 3 | 0 | 3 | 0.2932 |
| 5 | 66 | 3 | 1 | 2 | 0.3362 |
| 6 | 49 | 3 | 0 | 2 | 0.3998 |
| 7 | 15 | 3 | 0 | -9 | 0.2332 |
| 8 | 51 | 3 | 1 | 3 | 0.2701 |
| 9 | 42 | 3 | 1 | 2 | 0.3353 |

Mean Willingness Score Across Agents: 0.2832805674246766

# Tech Stack

**Backend**

Python (Models (MESA, SKLearn, Numpy/PyArrow), Pandas and Polars)

**Frontend**

Python (Streamlit)

**Cloud**

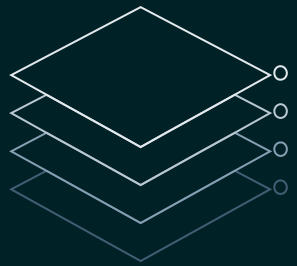
AWS (EC2, S3 (S3 can be used for data/matrix storage of prob scores and supporting files))

**Data**

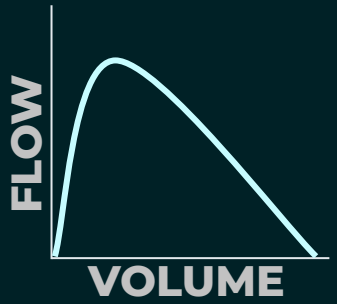
(Traditionally would connect to hospital/institute EMR - uploaded data sources from Kaggle)

# ANALYSIS PROCESS

1. *Data Ingestion*



2. *GRAPH GENERATION*

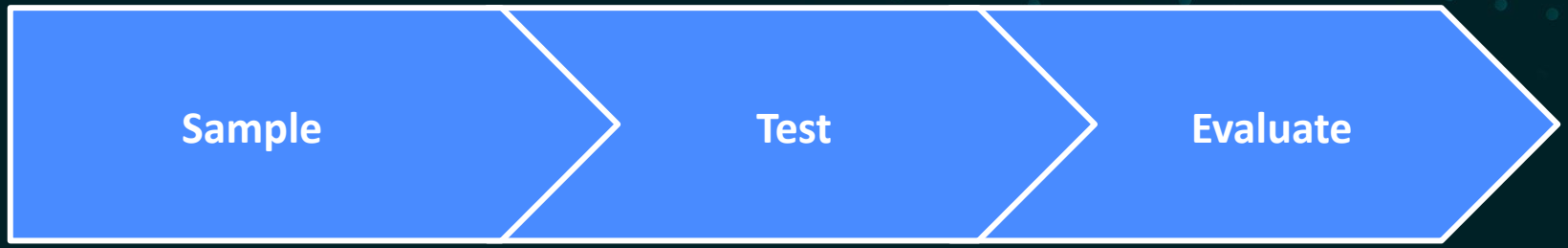



3. *REFERENCE VALUES*

4. *MACHINE LEARNING*

## Solution:

**Determine patient recruitment success for clinical studies is with the use of ABM's (Agent-Based Modeling)**




- Input: Live census data, social media scraping, cultural change
- Determine the sample
- Design patient survey accordingly

- Multiple populations, races, ethnicities.
- Use feature selection to determine optimal covariates.
- Simulate recruitment multiple times

- Evaluate which set of regions, peoples, backgrounds yield the best recruitment, are best to support inclusion/health equity efforts, and optimize resources

# Datasets Used


# MARKET Details

**$6.5k**

To Recruit one patient

*Source: GVR*

**$58B**

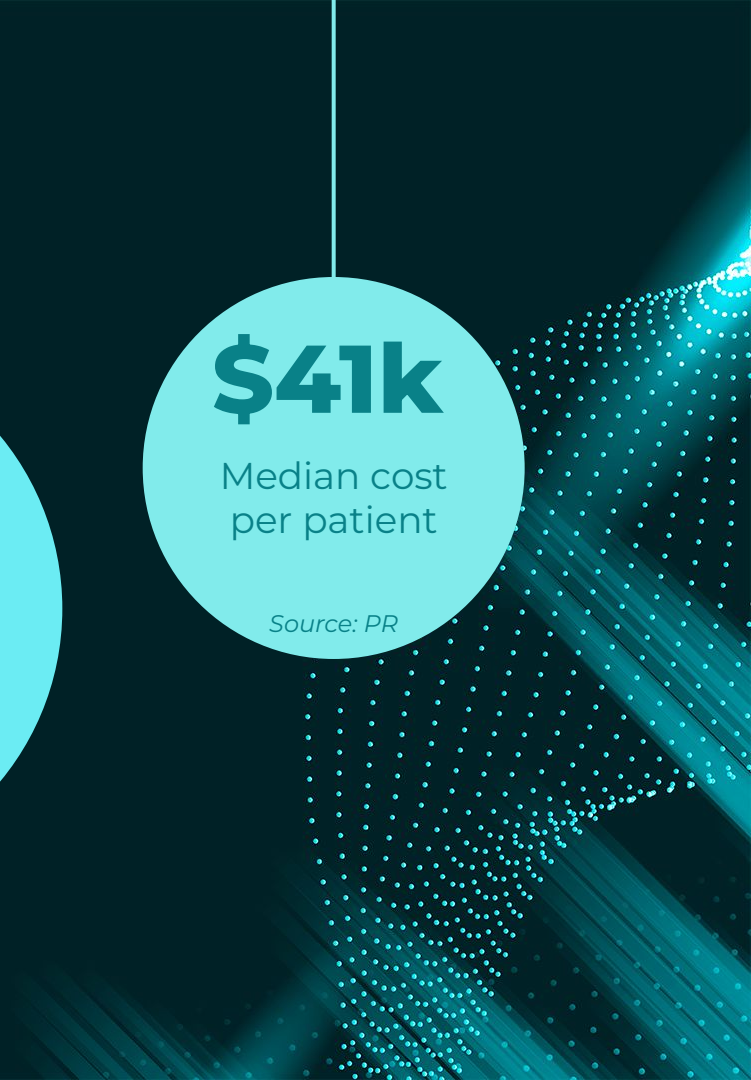
Clinical Trials
**CAGR: 7.1%**

*Source: BusinessForum*

**$41k**

Median cost per patient

*Source: PR*



# Impact/Feasibility & Scalability:

Give clinicians a tool based of Data to assist in making decisions that will cost a lot of time and money.

Reduced Dropout Rates: Consistent engagement (reminders, updates) lowers dropout rates, preserving data quality and shortening timelines.

Enhanced Data Accuracy: Real-time patient data collection improves accuracy and reduces manual data cleaning.

Cost Reduction: Streamlined recruitment, monitoring, and compliance reduce operational costs, benefiting both large and niche trials.

# THANK YOU!

SimuTrial

# Resources


Fu, T., Huang, K., Xiao, C., Glass, L. M., & Sun, J. (2022). HINT: Hierarchical interaction network for clinical-trial-outcome predictions. Patterns (New York, N.Y.), 3(4), 100445. https://doi.org/10.1016/j.patter.2022.100445

Jacques, R. M., Ahmed, R., Harper, J., Ranjan, A., Saeed, I., Simpson, R. M., & Walters, S. J. (2022). Recruitment, consent and retention of participants in randomised controlled trials: a review of trials published in the National Institute for Health Research (NIHR) Journals Library (1997-2020). BMJ open, 12(2), e059230. https://doi.org/10.1136/bmjopen-2021-059230



Fu, T., Huang, K., Xiao, C., Glass, L. M., & Sun, J. (2022). HINT: Hierarchical interaction network for clinical-trial-outcome predictions. Patterns (New York, N.Y.), 3(4), 100445. https://doi.org/10.1016/j.patter.2022.100445



Williams, C. (2021, September 29). Factors associated with clinical trial invitation and participation in the US. JAMA Network Open. https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2784556