Design and Assessment of Representative Hybrid Clinical Trials using Health Recommender System

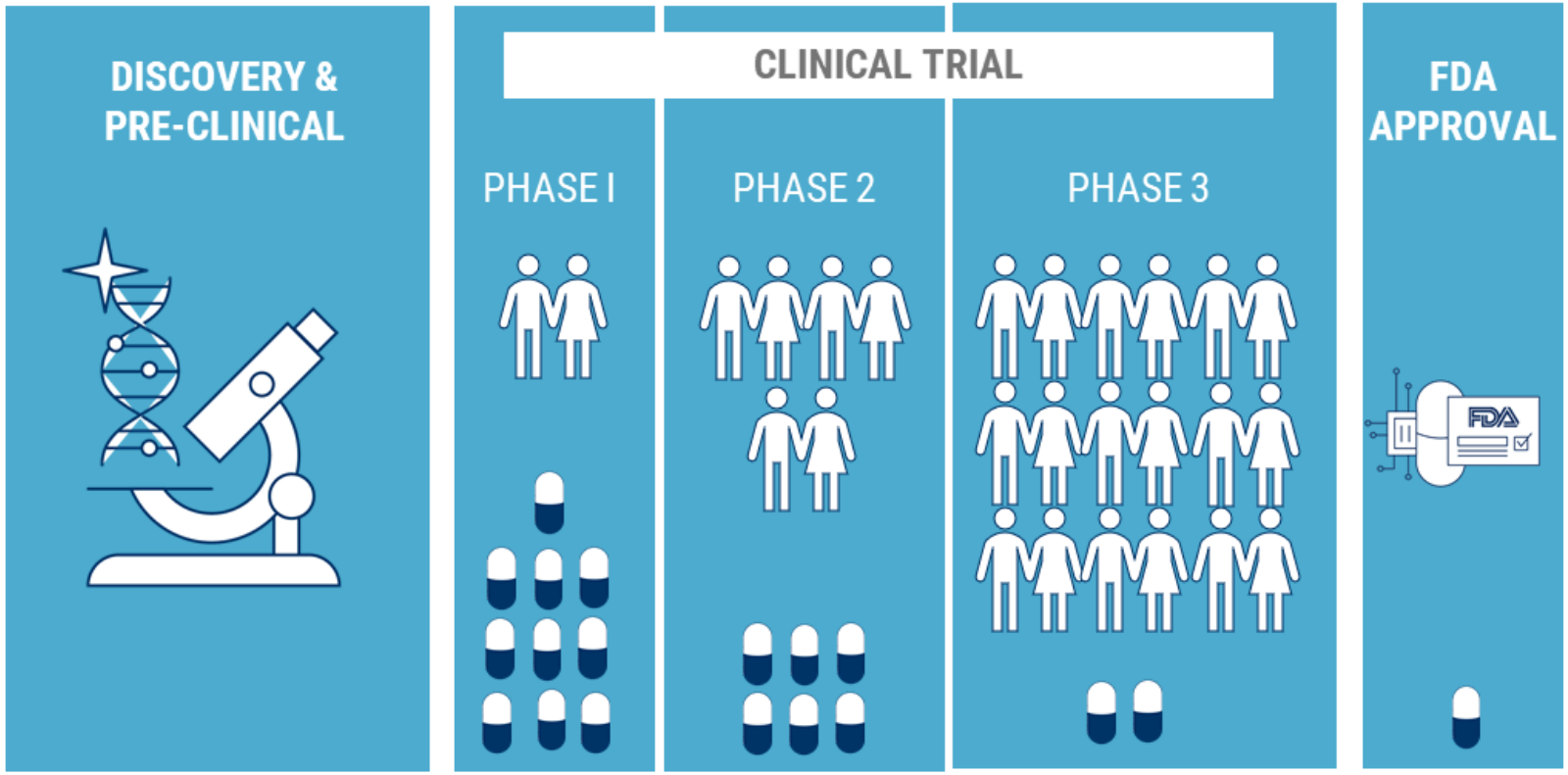
Nafis Neehal[1], Vibha Anand[2], Kristin P. Bennett[1]

[1]Rensselaer Polytechnic Institute

[2]IBM Research

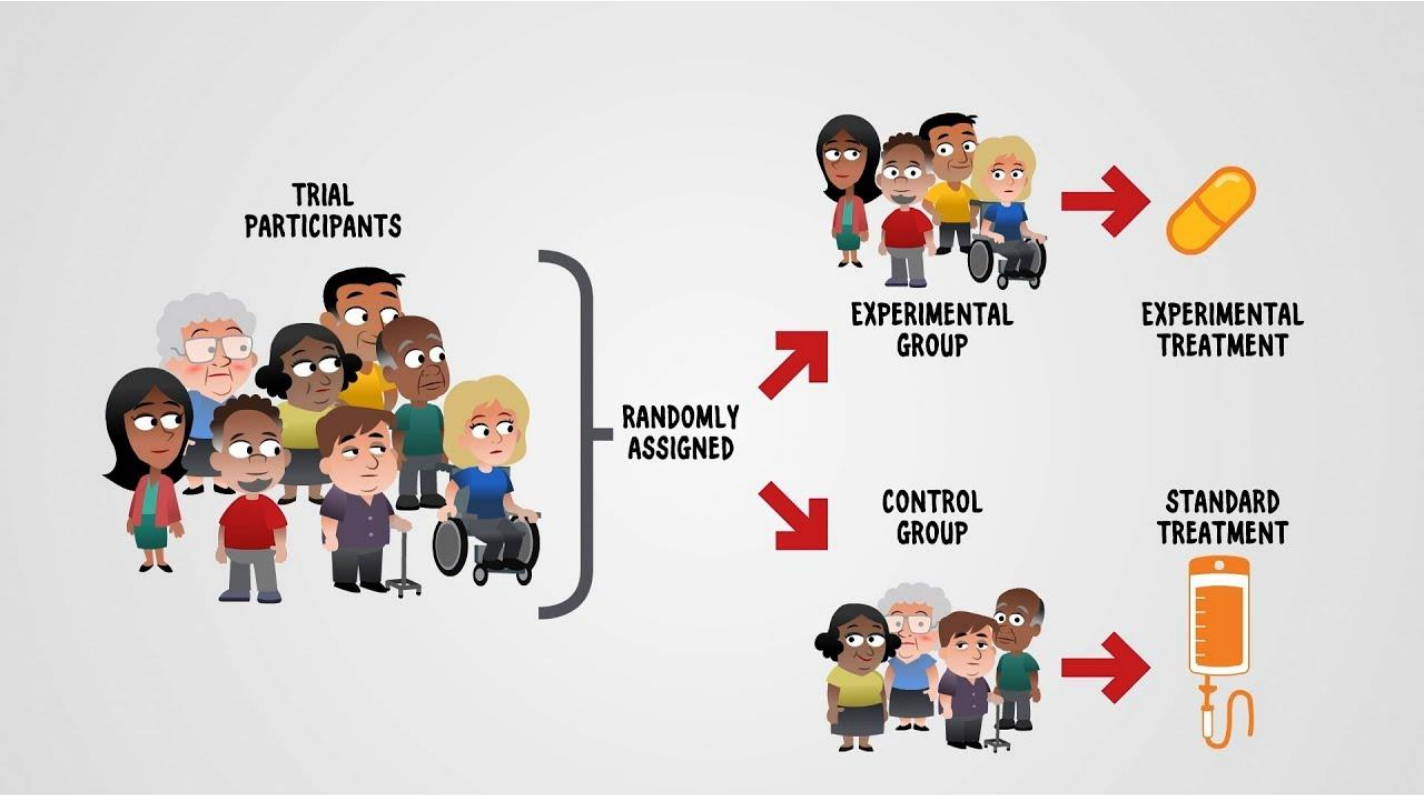
What are Clinical Trials?

- *Clinical trials* are a type of research that studies new tests and treatments and evaluates their effects on human health outcomes.
- Happens in multiple stages
- Involves complex design choices in different stages
- Average time span 10-15 years, costing millions of USD



Source: cbinsights.com

How do we know if a Drug Works?



Source: https://www.youtube.com/watch?v=QL3gvDEr9C0

Compare the features of patients (that are indicative of a patient's health outcome) in both groups after the clinical trial runs for a period of time – then calculate treatment effect

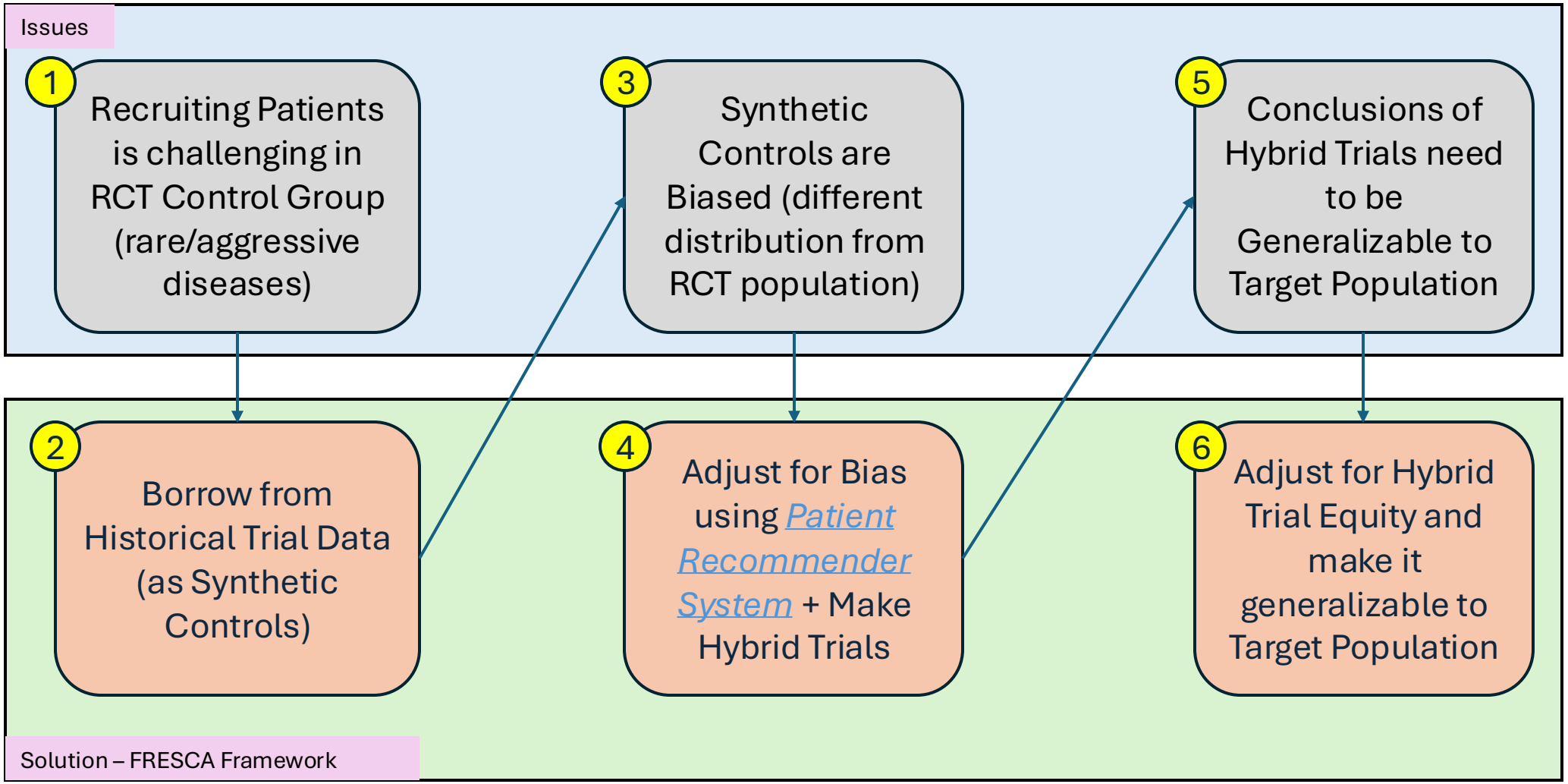
Motivation

- Randomized Controlled Trials (RCT)
  - gold standard for measuring an intervention's efficacy
- Synthetic Controls incorporate Real World Data (RWD) into RCTs
  - creates a *hybrid RCT population* (Trial Population + RWD from historically compatible trials)
  - has the potential to produce more effective studies

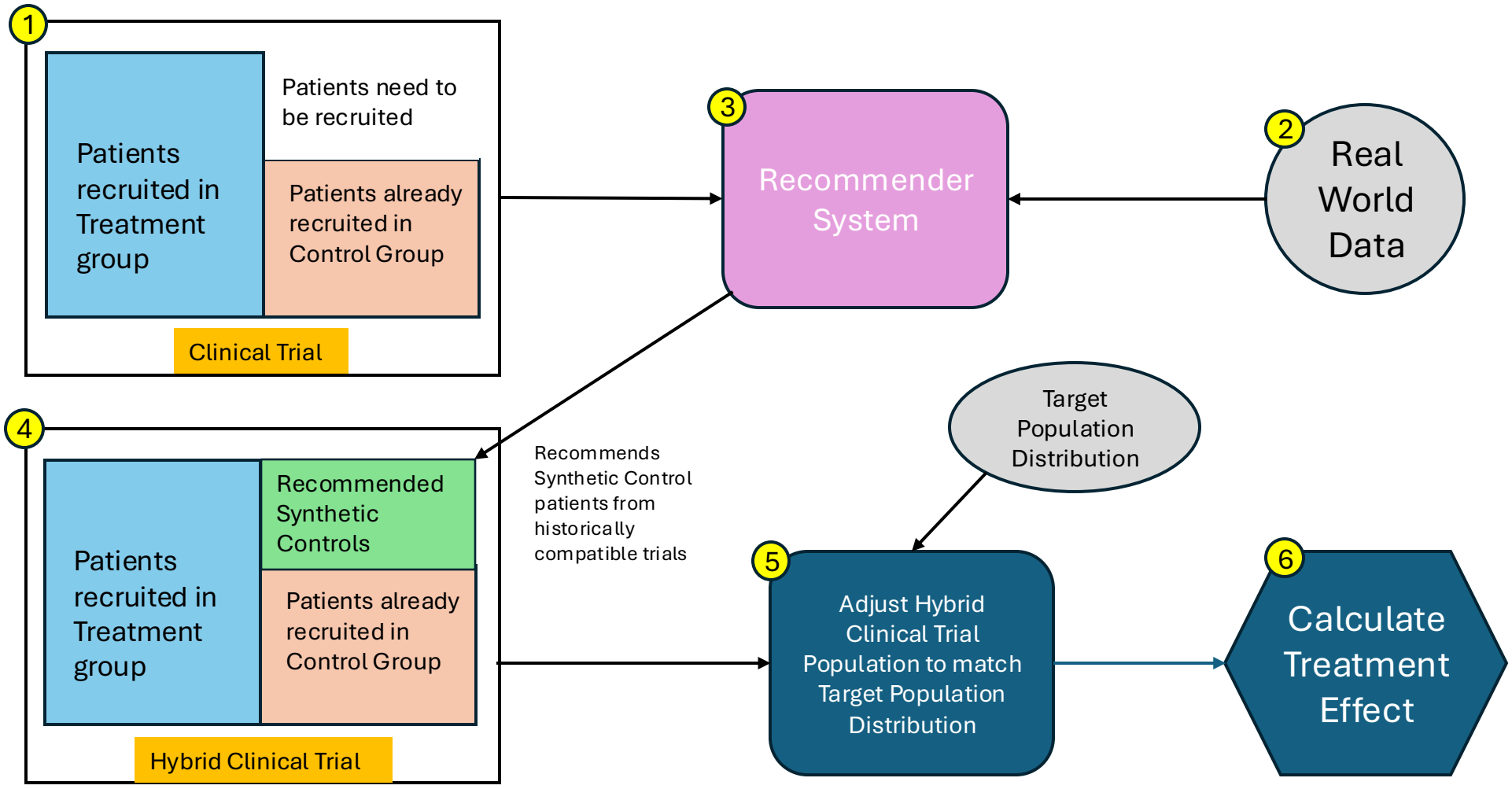
We focus on 3 RCT design-related issues –

- Issue 1: Challenges exist in RCT control patient recruitment
  - (e.g. rare diseases, ethical reasons in case of aggressive diseases, etc.)
- Issue 2: RWD to be used as the source of synthetic controls is biased
  - (i.e. distributions of RCT and RWD data don't match for significant features)
- Issue 3: RCT conclusions (e.g. the drug works / doesn't work) need to be equitable
  - (i.e., generalizable on the target population)

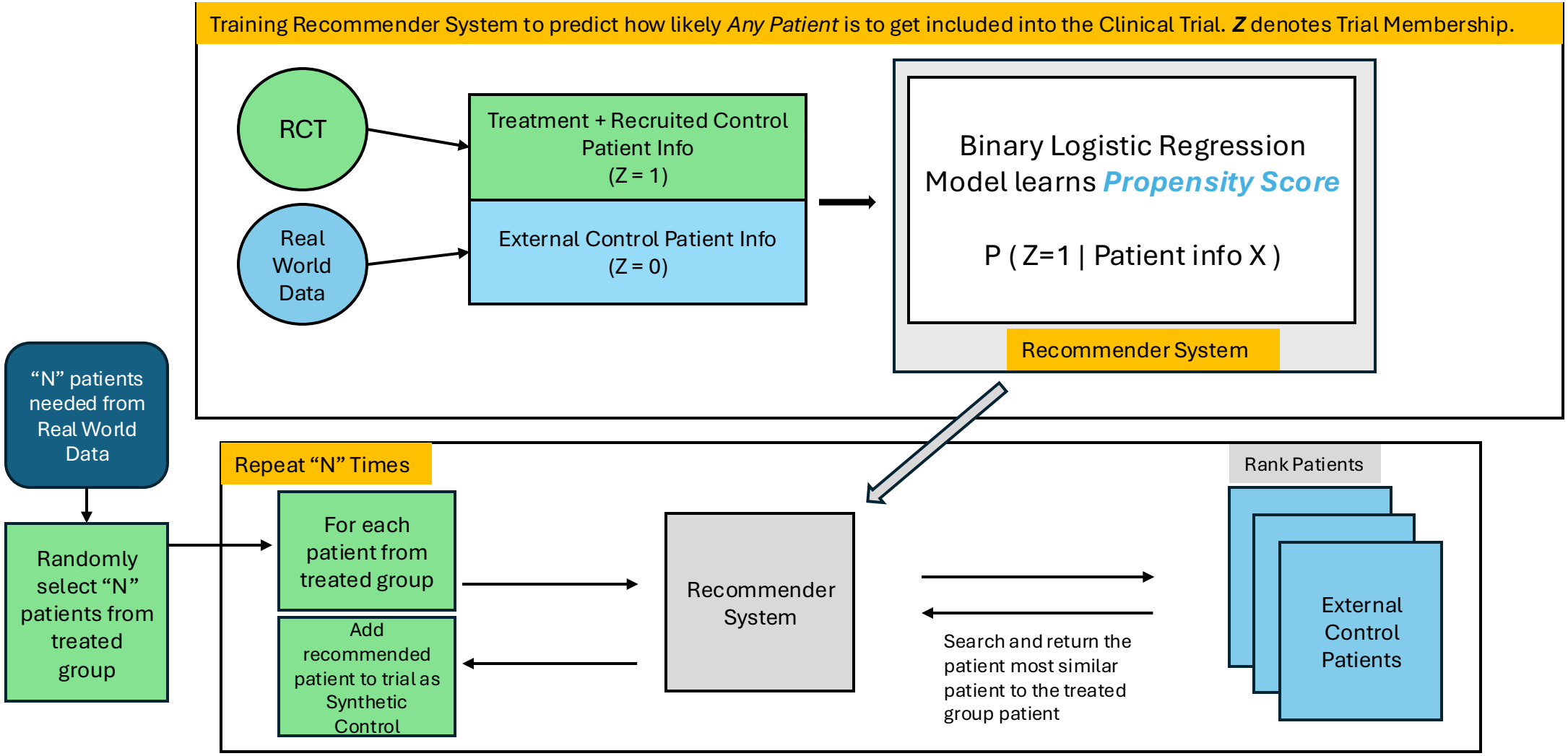
Proposed Solutions



FRESCA Framework



How do the Recommender System Work?



Main Finding

Table 2: Comparison of PHR and CTD across different trials, outcomes and methods. We show this for ALLHAT ($N_{TA} = 4000$, $N_{CC} = 2000$) and SPRINT ($N_{TA} = 2000$, $N_{CC} = 1000$) respectively. Symbol (†) in Cohort-Target Disparity column indicates measured CTD not being within equitable range ($CTD > 0.22$). Bold font indicates the best performing method.

Trial (Study)	Outcome Examined	Control Population	Adjustment Method	Target PHR [95% CI]	Estimated PHR [95% CI]	Cohort-Target Disparity [95% CI]
ALLHAT (Hypertension)	Secondary (Heart Failure)	CC	None	1.38 [1.36, 1.41]	1.39 [1.36, 1.43]	0.89 [0.84, 0.94]†
		Hybrid	NC Matching		1.42 [1.37, 1.48]	0.87 [0.81, 0.94]†
		Hybrid	Propensity Matching + IPF Sampling		1.43 [1.32, 1.49]	0.03 [0.02, 0.04]
		Hybrid	Propensity Matching + IPF Weighting		**1.39 [1.33, 1.46]**	**0.04 [0.03, 0.05]**
SPRINT (Hypertension)	Primary	CC	None	0.79 [0.77, 0.82]	0.75 [0.73, 0.78]	0.91 [0.86, 0.97]†
		Hybrid	NC Matching		0.74 [0.72, 0.77]	0.89 [0.84, 0.96]†
		Hybrid	Propensity Matching + IPF Sampling		0.75 [0.67, 0.84]	0.01 [0.00, 0.01]
		Hybrid	Propensity Matching + IPF Weighting		**0.78 [0.74, 0.81]**	**0.04 [0.03, 0.05]**

We Want –

1. Estimated PHR to be as close to Target PHR as possible
2. Cohort-Target Disparity <= 0.22

We Observe –

1. FRESCA Based Method (Propensity Matching + IPF Weighting) outperforms all other baseline methods in both Clinical Trials

Thank You

neehan@rpi.edu