Framework for Research in Equitable Synthetic Control Arms (FRESCA)

Session Title: Clinical Trials & Patient Eligibility - "Put It to the Test"

Session Number: S83

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

[1]Rensselaer Polytechnic Institute (RPI), [2]IBM

Email: neehan@rpi.edu

#AMIA2023

Disclosure

1. This work was primarily funded by IBM Research.
2. This manuscript was prepared using Systolic Blood Pressure Intervention Trial (SPRINT) study research materials obtained from the NHLBI Biologic Specimen and Data Repository.
3. Information Coordinating Center and does not necessarily reflect the opinions or views of the SPRINT or the NHLBI.

Learning Objectives

After participating in this session, the learner should be better able to:

- Understand the equity-related issues that arise when conducting a hybrid clinical trial.
- Explain how propensity and equity adjustments of synthetic control arms can improve both the accuracy of treatment effect estimation and equity.
- Create, assess, and compare the performance of methods for constructing equitable RCT with Hybrid Control Arms using the proposed FRESCA framework.
- Learn the potential benefits of augmenting Synthetic Controls into an equitable RCT (reduced trial length, reduced cost, increased statistical power, etc.)

Introduction

- RCT – the gold standard for measuring an intervention's efficacy
- We focus on two RCT design-related issues -
  - Issue 1: Challenges exist in RCT Control patient recruitment (e.g. rare aggressive diseases)
  - Issue 2: RCT conclusions need to be equitable (i.e., generalizable on target population)
- We aim to tackle both issues with our proposed framework - FRESCA
- For Issue 1: FRESCA allows borrowing Synthetic Controls from RWD and form a Hybrid RCT
  - Potential to achieve reduced trial length, reduced cost, and increased statistical power
- For Issue 2: FRESCA allows calculating adjusted treatment effects using Hybrid RCT Population for an intended target population

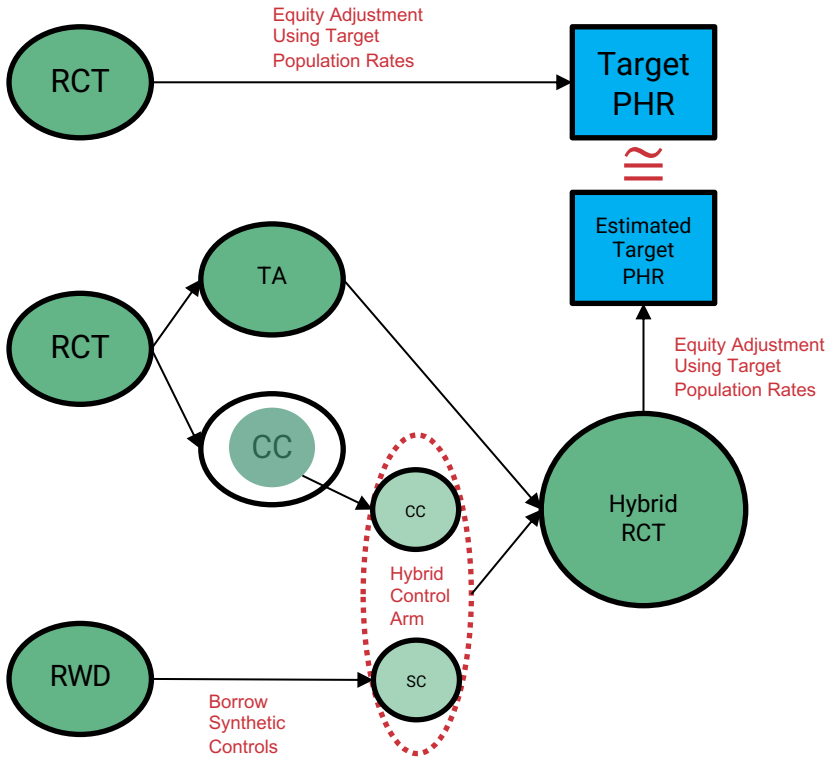
Major Contributions

- Identify and define the issue of generalizability/equity in Hybrid RCTs
- Propose and investigate a method for Hybrid RCT formation and calculating adjusted treatment effect with respect to a target population
- Utility to evaluate and compare the performance of different methods and also different RCTs, RWDs and Target Populations
- Empirically demonstrate that Matching and Equity Adjustment lead to accurate Target PHR* estimations
- Empirically demonstrate that accurate Target PHR estimations are possible with lesser recruitment in RCT Controls (i.e., augmenting synthetic controls)
- An open-source R Package (ongoing work)

* Target PHR = estimated treatment effect on the target population using RCT / Hybrid RCT population. PHR means Population Hazard Ratio.

Goal

- Define a ``ground-truth'' target PHR using equity adjustment on whole RCT data
- Take a smaller set of In-Trial / Concurrent Controls (CC) from RCT
- Borrow Synthetic Controls (SC) to augment in-trial control group
- Form a Hybrid Control Arm (CC + SC)
- Form a Hybrid RCT (TA + Hybrid Controls)
- Estimate equity adjusted PHR using Hybrid RCT
- Goal: estimated PHR ≅ target PHR



Data

- Target Population: National Health and Nutrition Examination Survey (NHANES) 2015-2016
- RCT Population: Systolic Blood Pressure Intervention Trial (SPRINT)
  - After pre-processing, 4234 Treated and 4200 Controls
- RWD is simulated using RCT data
- Outcome Analyzed: Primary Outcome of SPRINT
- Equity was evaluated based on protected attributes[1] (Age, Race, Gender)
  - e.g., P_{SPRINT}(Asian, Female, 50+) = P_{NHANES}(Asian, Female, 50+)

[1] Qi M, Cahan O, Foreman MA, Gruen DM, Das AK, Bennett KP. *Quantifying representativeness in randomized clinical trials using machine learning fairness metrics.* JAMIA Open. 2021;4(3):ooab077.

Metrics

- Measure of Treatment Effect: Hazard Ratio (Cox's PH Model)
  - Target PHR in our case = equity adjusted treatment effect using full RCT (SPRINT) data
  - Goal is to achieve estimated PHR similar to target PHR using Hybrid RCT data
- Measure of Equity: Log Disparity[1] measured in cohort level
  - Equivalent to the ratio of enrollment odds of subjects of the protected group $g(x)$ in the observed cohort y' to the odds of protected subjects in the ideal cohort y :

$$\frac{odds\ (g(x) = 1 |\ y' = 1)}{odds\ (g(x) = 1\ | y = 1)}$$

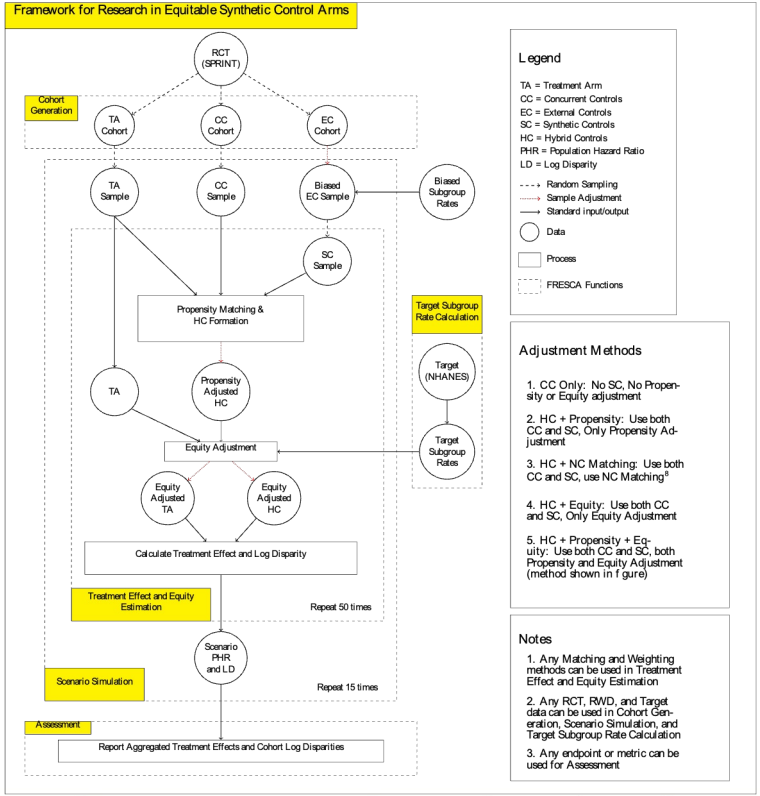
  - Protected group can be any lowest-level subgroup combining multiple features
    - E.g. consider two features – Gender (Male, Female), Race/Ethnicity (NH White, NH Black)
    - Protected group (i.e. lowest-level subgroup) can be: NH White Female, NH Black Male etc.

[1] Qi M, Cahan O, Foreman MA, Gruen DM, Das AK, Bennett KP. *Quantifying representativeness in randomized clinical trials using machine learning fairness metrics.* JAMIA Open. 2021;4(3):ooab077.

FRESCA Framework

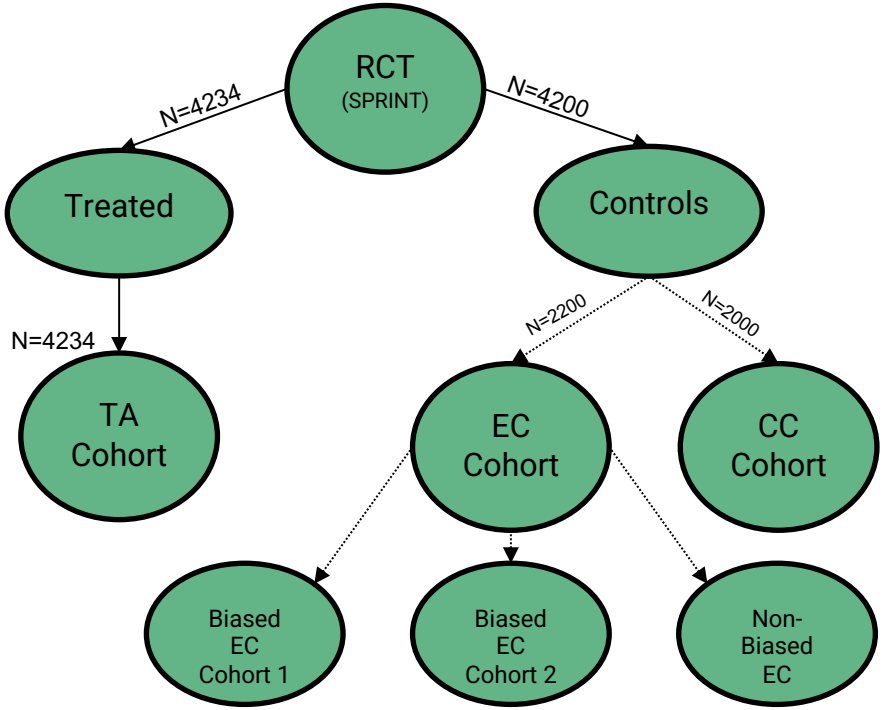
FRESCA has 5 main functions –

1. Cohort Generation
2. Scenario Simulation
3. Target Subgroup Rates Calculation
4. Treatment Effect and Equity Estimation
5. Assessment



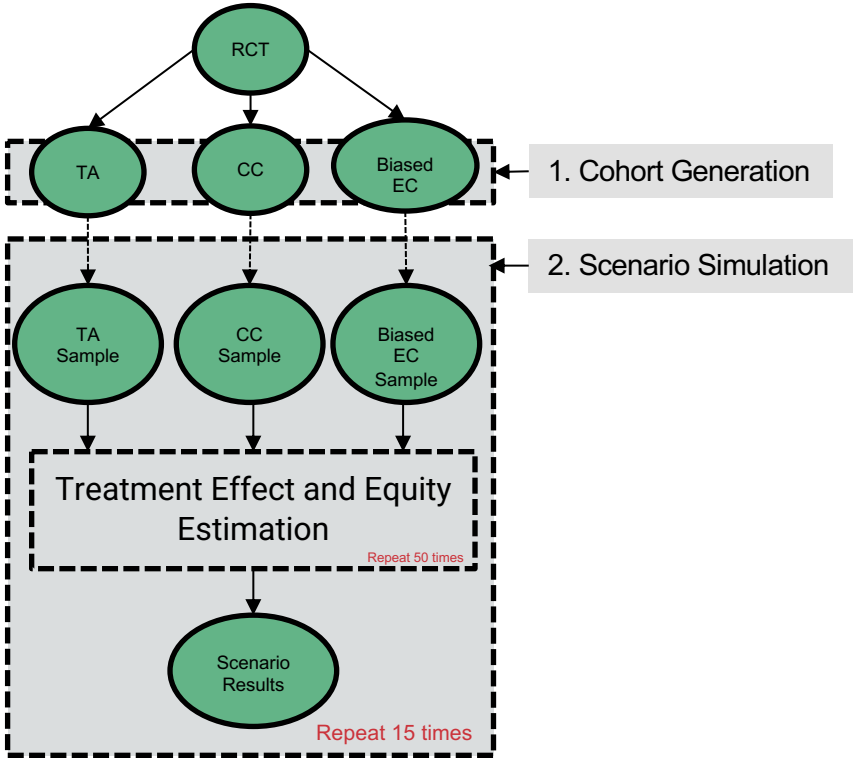
Function 1: Cohort Generation

- TA Cohort (Treatment Arm) contains all treated patients in SPRINT
- All SPRINT control patients are randomly divided into EC Cohort (External Controls) and CC Cohort (Concurrent / In-Trial Controls)
- Distribution of initial EC cohort is further manipulated to create 3 different Biased EC Cohorts



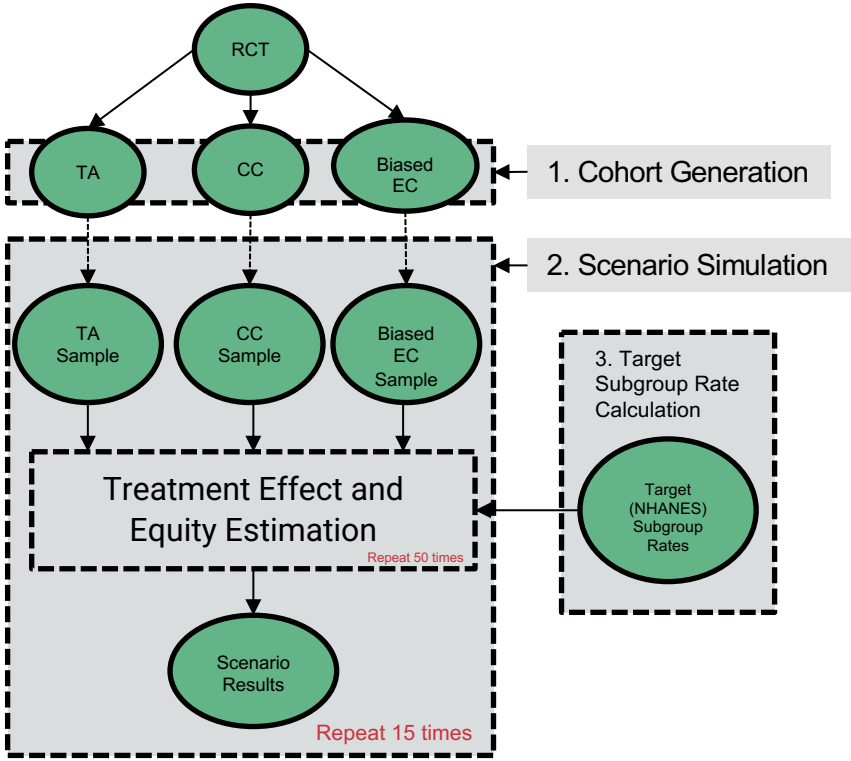
Function 2: Scenario Simulation

- Generates hypothetical clinical trials
- TA and CC samples represent trial patients with random assignment
- Biased EC Sample represents RWD
- We experiment with varying CC sizes across different trial scenarios
- Throughout all experiments –
  - TA Sample size = 2000
  - CC Sample size = {0, 500, 1000, 1500, 2000}



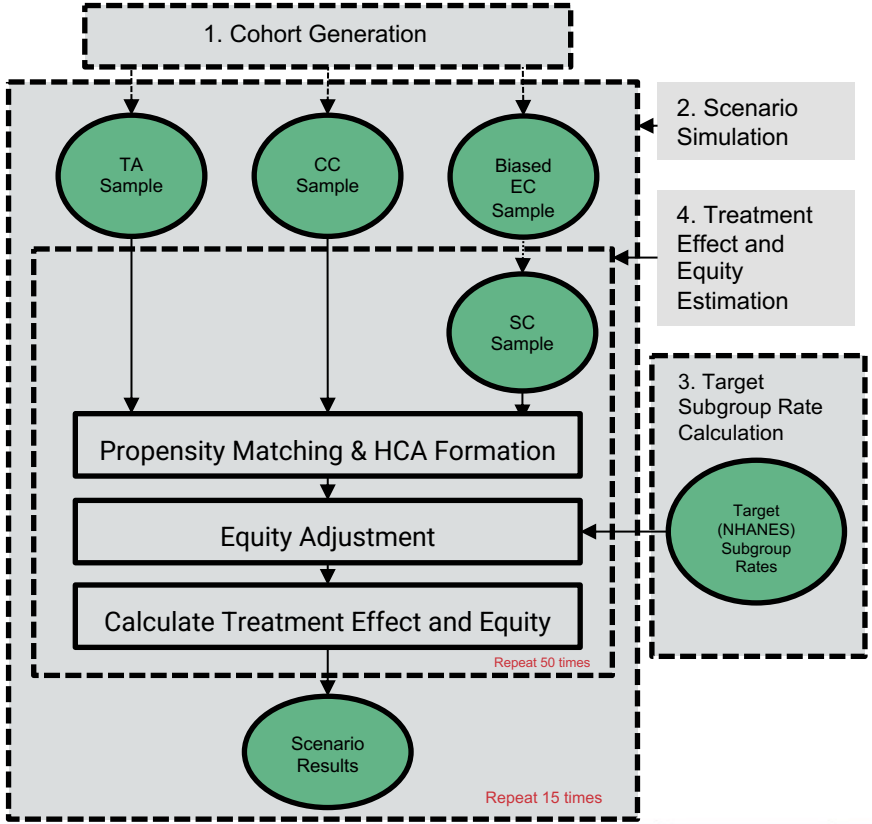
Function 3: Target Subgroup Rate Calculation

- Desired Target Subgroup rates are calculated from NHANES data using Survey-Weighted Analysis [1]
- Only target marginal distributions from the target population is required
- Any Target population can be used



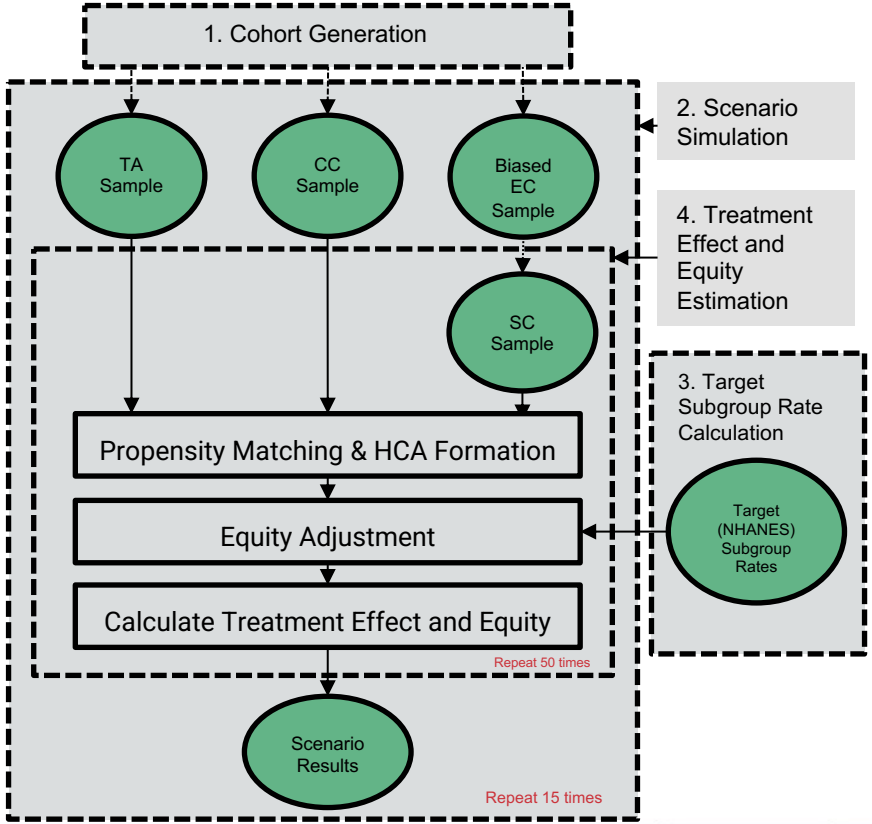
Function 4: Treatment Effect and Equity Estimation

- SC Sample is randomly drawn with a size of $N_{SC} = N_{TA} - N_{CC}$
- SC patients most likely to be in the trial are then selected using propensity matching
- Formation of the HCA = Hybrid Control Arm (CC + SC)



Function 4: Treatment Effect and Equity Estimation (Cont.)

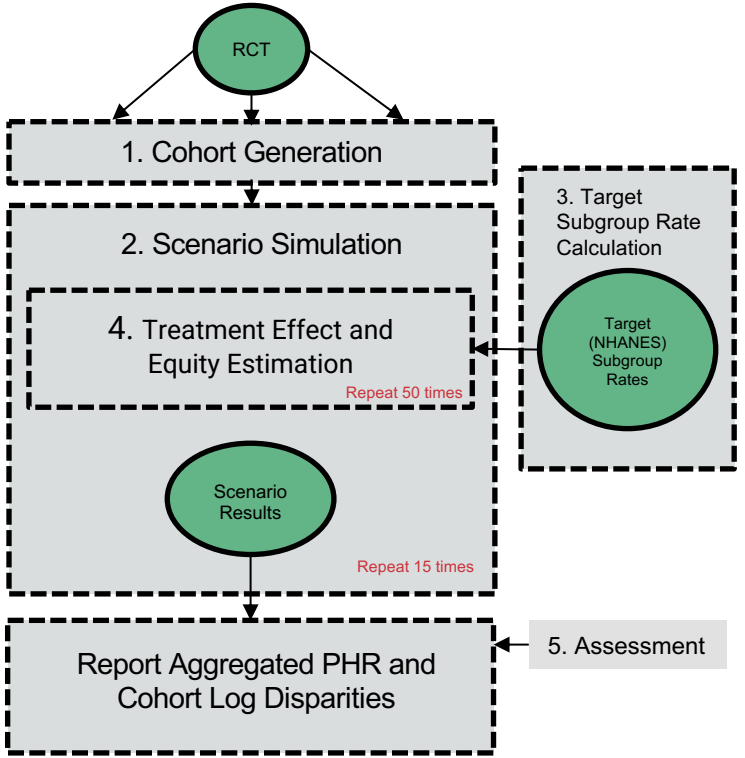
- Iterative Proportional Fitting (IPF)* is used to make equity adjustments on both TA and HCA
- IPF returns new adjusted weights for each patient in TA and HCA cohorts
  - Tries to match Hybrid RCT to Target
- Bootstrap samples using these weights – new adjusted TA and HCA
- Measure treatment effect and equity in new adjusted Hybrid RCT



* IPF adjusts an initial set of weights defining a seed distribution to match various target marginal distributions

Function 5: Assessment

- Aggregated PHR estimates and Cohort Log Disparities (equity measure) are reported
- Aggregated Cohort Log Disparity values need to be within [0, 0.22][1]
- Aggregated PHR estimates are expected to be within 95% CI of the target PHR



Results: Comparison of Different HCA Construction Methods

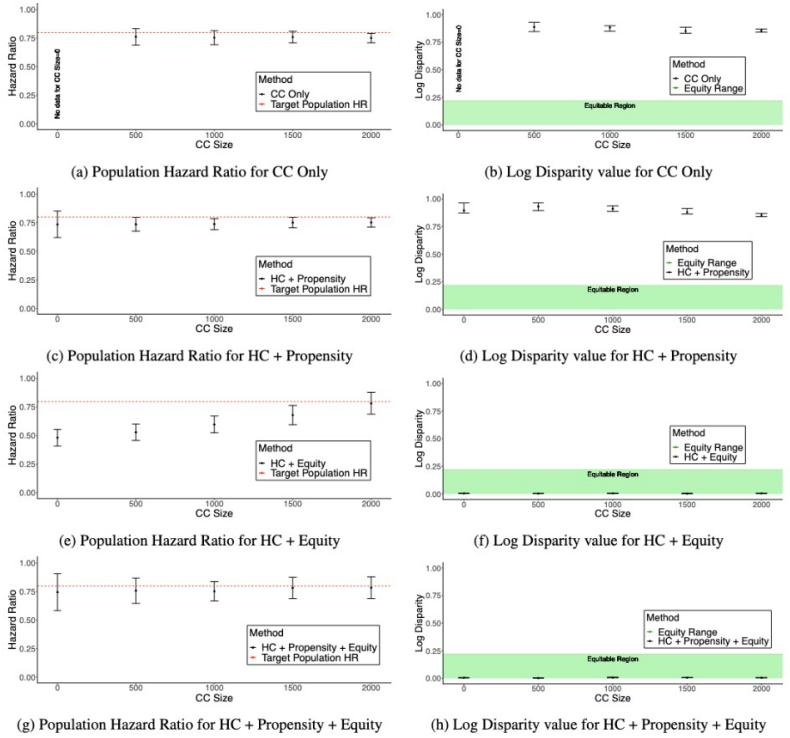
- Examine all 3 different Biased EC Cohorts (3 different cases)
- Ablation study – 4 different cases + One state-of-the-art matching algorithm
- Failing to perform equity adjustment leads to inequitable trials in all cases
- Only equity adjustment produced an inaccurate PHR estimate (“High Risk” cohort)
  - only performing equity adjustment may not be sufficient to estimate PHR accurately
- Hybrid RCTs with equity adjustment
  - Accurate equity-adjusted PHR estimates in all cases

Table 3: Comparison of Population Hazard Ratio and Log Disparity across different methods. We show this for with 2000 and 1000 patients in TA and CC sample cohort respectively. Bold (*) symbol in Hazard Ratio column indicates estimated HR being significantly different ($p < 0.05$) from “Ground Truth” Target PHR. Bold (†) symbol in Log Disparity column indicates measured Log Disparity not being within equitable range. ($LD > 0.22$)

Population Reference	Control Population	Adjustment Method	Hazard Ratio [95% Confidence Interval]	Cohort Log Disparity [95% Confidence Interval]
High Risk	CC	None	0.753 [0.692, 0.815]	**0.881 [0.807, 0.983]**†
	HC	Propensity	0.737 [0.689, 0.785]	**0.843 [0.792, 0.900]**†
	HC	NC Matching	0.747 [0.681, 0.819]	**0.893 [0.821, 0.950]**†
	HC	Equity	**0.599 [0.525, 0.671]***	0.007 [0.004, 0.013]
	HC	Propensity + Equity	0.752 [0.668, 0.837]	0.007 [0.003, 0.011]
Veterans	CC	None	0.753 [0.692, 0.815]	**0.881 [0.807, 0.983]**†
	HC	Propensity	0.751 [0.718, 0.784]	**0.769 [0.719, 0.820]**†
	HC	NC Matching	0.742 [0.704, 0.793]	**0.797 [0.758, 0.911]**†
	HC	Equity	0.728 [0.662, 0.793]	0.007 [0.004, 0.011]
	HC	Propensity + Equity	0.751 [0.687, 0.816]	0.007 [0.004, 0.011]
Unbiased	CC	None	0.753 [0.692, 0.815]	**0.881 [0.807, 0.983]**†
	HC	Propensity	0.741 [0.697, 0.784]	**0.889 [0.841, 0.937]**†
	HC	NC Matching	0.753 [0.684, 0.821]	**0.874 [0.815, 0.943]**†
	HC	Equity	0.772 [0.701, 0.844]	0.006 [0.004, 0.012]
	HC	Propensity + Equity	0.771 [0.698, 0.843]	0.006 [0.003, 0.012]
Target Population	All Controls	Equity	0.798 [0.781, 0.817]	0.031 [0.022, 0.040]

Results: Examination of Variation in CC Size

- Examining “High Risk” biased cohort
- Fixed TA Size of 2000, Vary CC and SC Size
- Missing/Zero patients in some lowest-level subgroups with small CC Size (Fig 2b)
- Lack of equity adjustment never produces acceptable equity values (Figs 2a, 2b, 2c, 2d)
- PHR Estimation variance decreases with growing CC Size across all 4 cases (Figs 2a, 2c, 2e, 2g)
- In “HC + Propensity + Equity”, accurate PHR estimates with acceptable equity
  - Potential to reduce recruited CC Size by augmenting SC
  - Tradeoff with Variance



(a) Population Hazard Ratio for CC Only
(b) Log Disparity value for CC Only
(c) Population Hazard Ratio for HC + Propensity
(d) Log Disparity value for HC + Propensity
(e) Population Hazard Ratio for HC + Equity
(f) Log Disparity value for HC + Equity
(g) Population Hazard Ratio for HC + Propensity + Equity
(h) Log Disparity value for HC + Propensity + Equity

Conclusion

- Presented novel framework FRESCA
- Presented novel method for equitable Hybrid RCT construction and PHR Estimation
- Compared 5 different methods/strategies for equitable Hybrid RCT construction
- Empirically demonstrated that equity-adjusted Hybrid RCT can achieve similar PHR estimates to target PHR with smaller CC Size
- Ongoing work - several room for improvements

Thank you!

Email: neehan@rpi.edu