Recruitment Strategy for Prospective Observational Studies to Mimic Randomization
Up-Front Matching

Ibrahim Turkoz, PhD

15 Feb 2023

Disclosures

- Current employee Janssen R&D LLC
- Stock Johnson & Johnson

Agenda

Prospective Observational Studies

Prospective Observational Studies can be one of the best options when a RCTs are not possible to execute.

Strength of Prospective Observational Studies include:

1. Can provide better quality of data on the primary exposure and on confounding variables.
   a. Clear specification of target patient population(s), treatments, and outcomes of interest for making inferences regarding causal effects.
   b. Generate apriori study protocol, build data collection algorithms, state the purpose or main hypotheses, identifies confounders (whether measured or not), specifies the primary analyses and required sample size.
   c. Implement time and even schedules, clinical scales, patient reported outcomes.
   d. Control quality of data.
2. Since exposures are assessed before outcomes occur, they are less prone to bias.
   a. Collect baseline data on all subjects, before any of them have developed the outcomes of interest.

Disadvantages to Prospective Observational Studies include:

1. They could be more expensive and time consuming.
2. They are not efficient for diseases with long latency.
3. Must account for measured and unmeasured confounders. Losses to follow up can bias the measure of association.

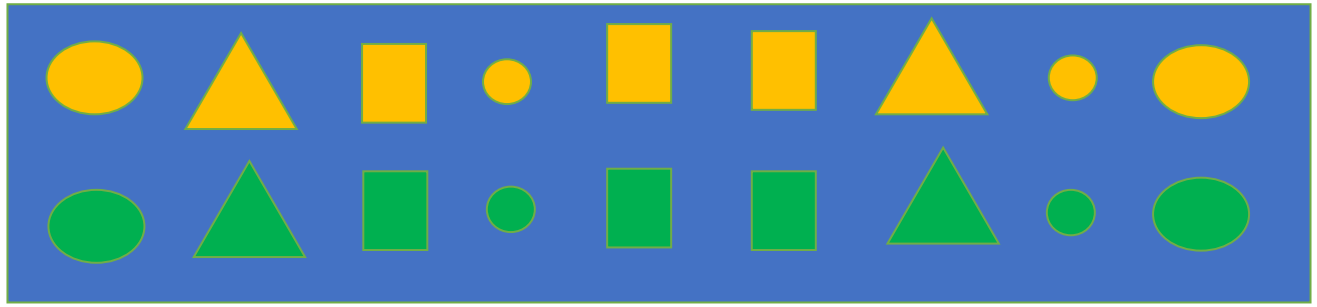
Challenges Associated with the Use of RWE/Observational Studies

- Selection Bias
- Information Bias
  -memory bias
  -interviewer bias
- Confounding

Randomized Clinical Trial vs. Observational Study

Green: patients who receive treatment A
Yellow: patients who receive treatment B

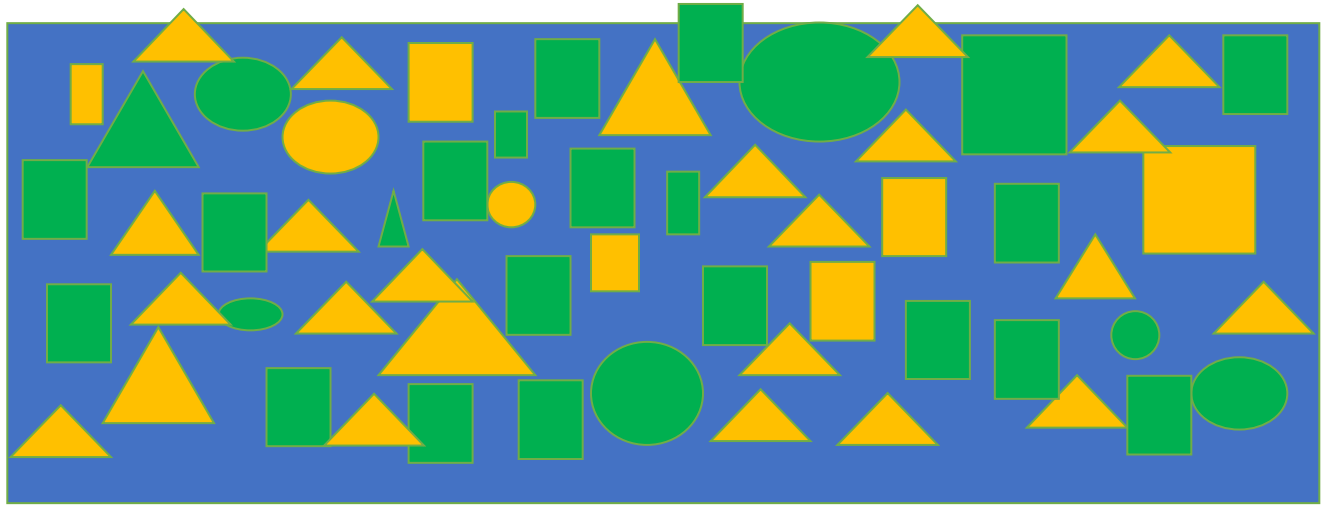
Randomized Clinical Trials



- With randomization – standard methods produce estimates of causal treatment effects

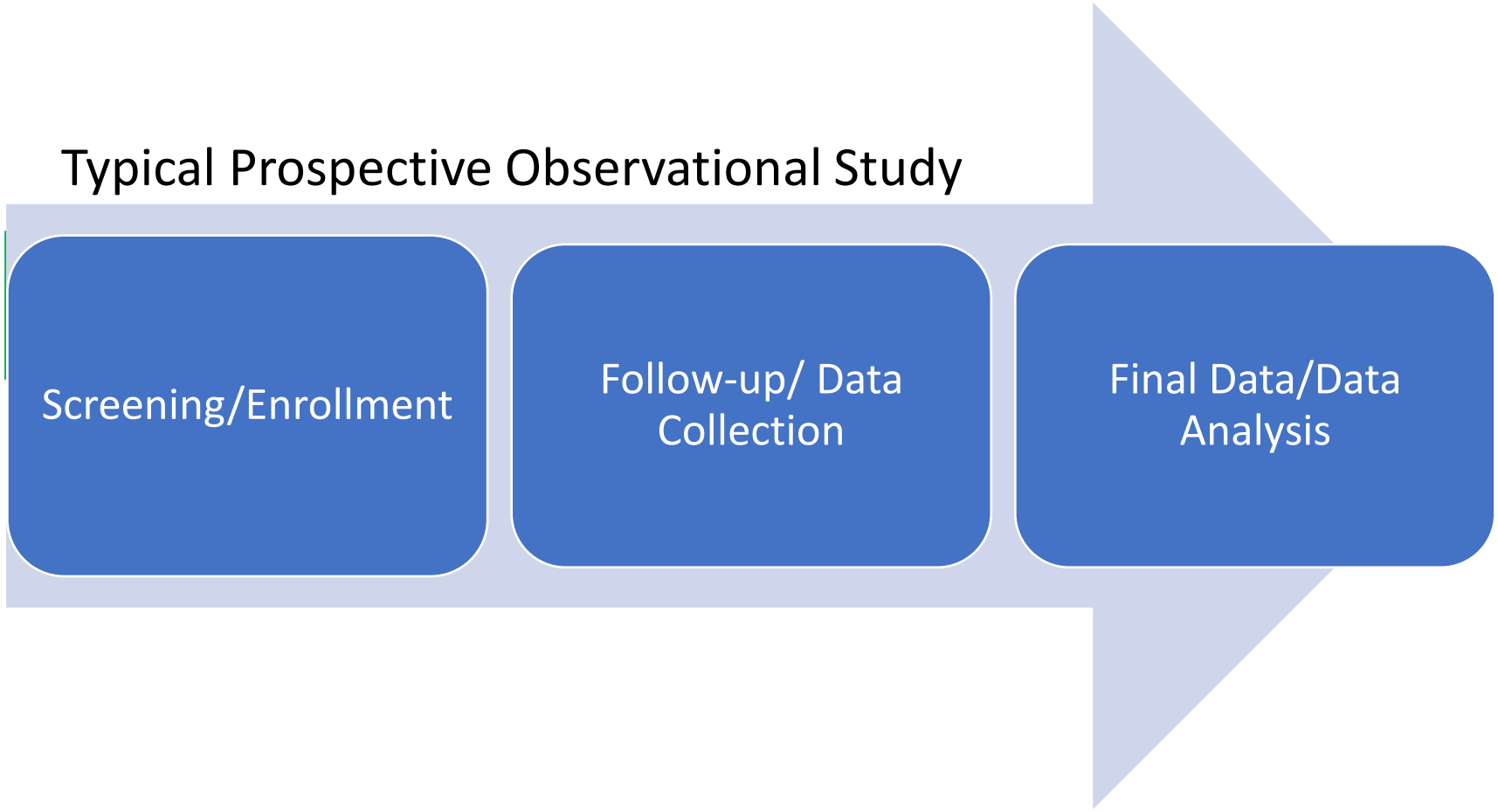
Randomized Clinical Trial vs. Observational Study (continued)

Observational Study



- <u>Without randomization</u> – standard methods produce only ‘associations’ …. Treatment groups are NOT comparable prior to drug/intervention initiation, thus comparisons are BIASED

Typical Prospective Observational Study



Typical Prospective Observational Study

Screening/Enrollment

Follow-up/ Data Collection

Final Data/Data Analysis

TRT A

TRT B

Typical Prospective Observational Study

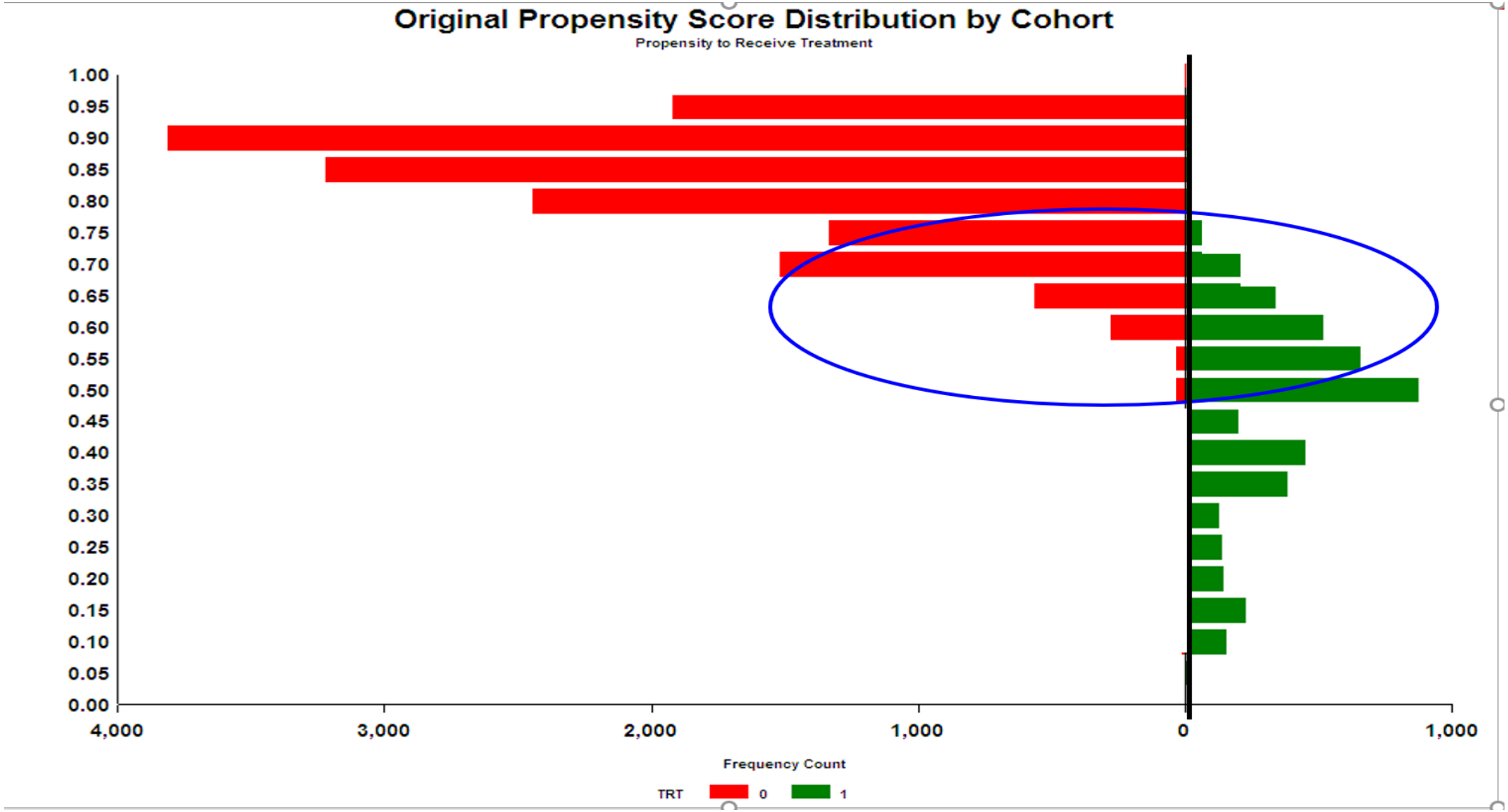
Screening/Enrollment

Follow-up/ Data Collection

Final Data/Data Analysis

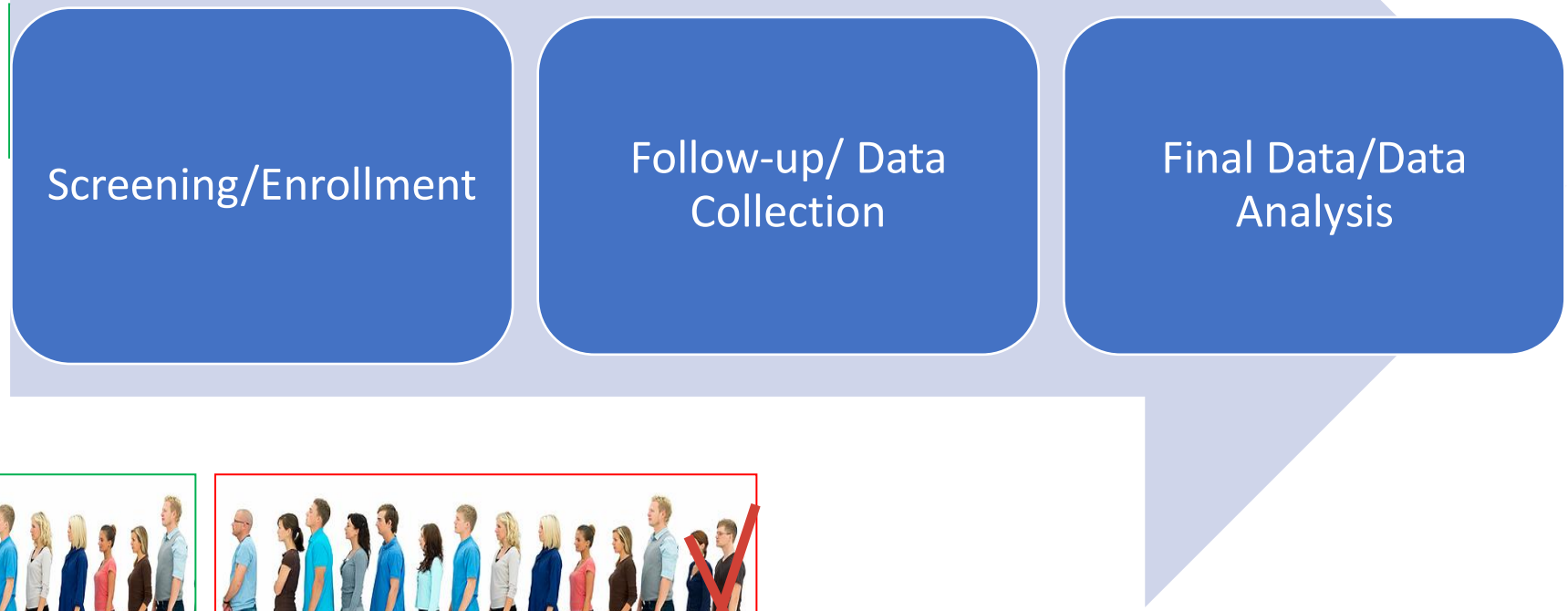
TRT A

TRT B


Original Propensity Score Distribution by Cohort
Propensity to Receive Treatment
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4,000
3,000
2,000
1,000
0
1,000
Frequency Count
TRT
0
1

Up-Front Matching

Typical Prospective Observational Study



TRT A

TRT B

Up-Front Matching Methodology for Selected Covariates

- Objective is to assess the comparative effectiveness of two drugs, say Drug A and Drug B where the patient populations are noticeably different but there is considerable overlap.

- To keep the description relatively concrete suppose we want to assess the average treatment effect in a patient population with characteristics of patients being treated with Drug A, i.e., the average treatment effect in the treated (in this case those treated with Drug A).

Up-Front Matching Methodology for Selected Covariates

- The key to up-front matching is to use of **readily available/accessible (inexpensive) covariates.**

- The goal of up-front matching is to create enrolled populations in the prospective observational study with (1) a higher percentage of patients **in the common support** as determined by the propensity score based on all baseline covariates in the enrolled populations at the end of the study; and (2) **balance across the inexpensive covariates** in the final enrolled populations. Note that the propensity score is used here for its balancing properties (Rosenbaum and Rubin 1983) and not as means to create treatment ignorability as the basis for causal inference; study outcomes play no role in the up-front matching method.

Up-Front Matching Methodology for Selected Covariates

Up-front matching is performed as follows:

a-) Determine the common support based on the propensity score ($\pi(s)$) distributions of the two treatment groups in the claims database.

b-) Determine the deciles of the propensity score distribution of Drug A; we are interested in enrolling patients in both groups who have pretreatment characteristics like those who were treated with Drug A.

c-) As patients are considered for enrollment, enroll only those whose propensity score is in the common support.

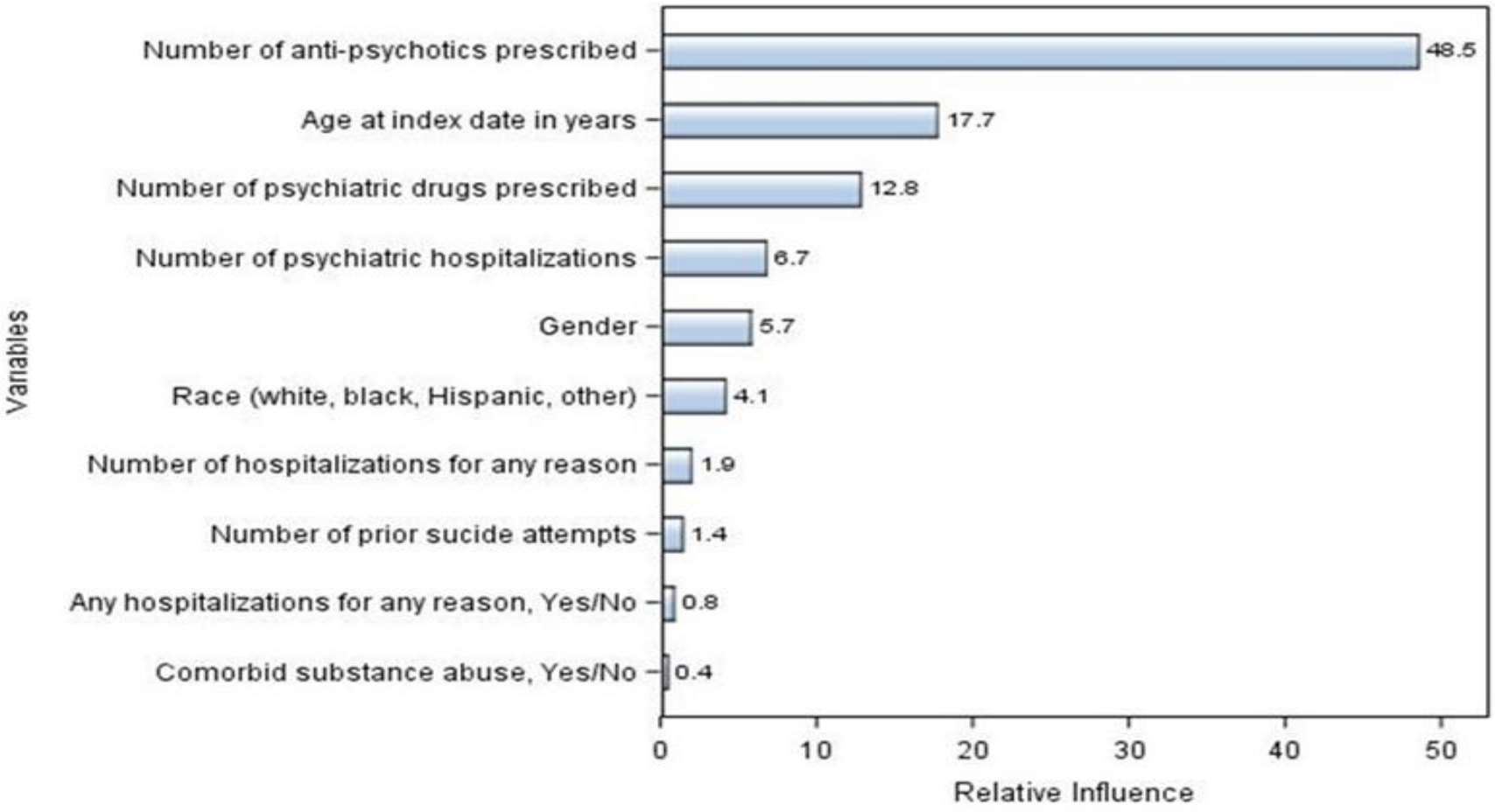
d-) Specify a quota of patients to be enrolled for each treatment group in each decile. For our illustration we take this quota to be the same for each treatment-decile combination – this is stratified matching (on the propensity score), a form of statistical sampling that is alternatively known as frequency matching.

e-) For each treatment group once the quota in a treatment-decile group has been filled that treatment-decile is closed to further enrollment.

f-) For each treatment group continue enrolling patients until all treatment-decile quotas for that treatment group have been met.

Simulation Results

Matching Variables and Their Relative Influence



Note: All non-demographic variables are based on the 1-year follow-back period.

The baseline characteristics deemed to be of interest and ascertainable in both the claims database and in patients who are to be considered for enrollment at investigative sites are given in the Figure along with their relative importance's based on the preliminary propensity score model.

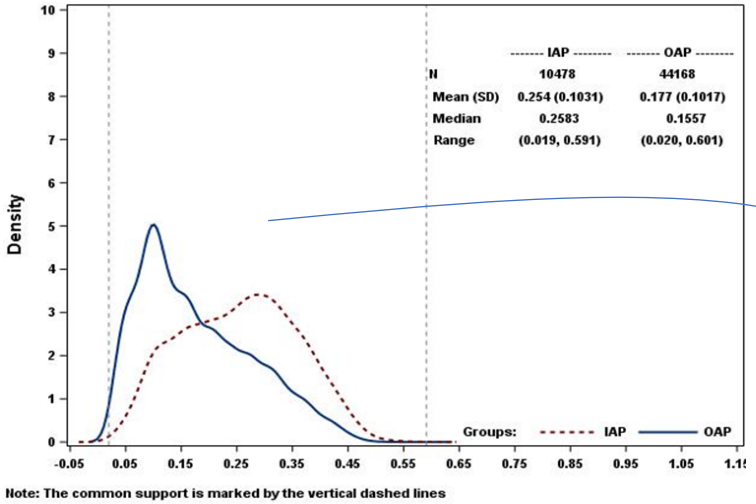
Simulation Results

Table 1 - Summary of Propensity Score Covariates in Claims Database

Variables	IAP (n=10478)		OAP (n=44168)		Standardized Difference in Means[1]
	Mean	Variance	Mean	Variance	
Number of anti-psychotics prescribed (*n.antipsychotics*)	1.74	1.69	1.18	1.40	0.457
Age at index date, years (*age*)	37.79	171.19	38.19	193.71	-0.030
Number of psychiatric hospitalizations (*num.psych.hosp*)	0.78	1.91	0.53	1.35	0.200
Number of psychiatric drugs prescribed (*n.all.psychiatric.drugs*)	3.97	8.64	3.66	8.57	0.108
Gender (*male*)	0.41	0.24	0.51	0.25	-0.194
Race (*white*)	0.36	0.23	0.43	0.24	-0.138
Race (*black*)	0.48	0.25	0.41	0.24	0.145
Race (*Hispanic*)	0.02	0.02	0.02	0.01	0.003
Race (*other*)	0.15	0.12	0.15	0.13	-0.013

IAP: injectable atypical anti-psychotics; OAP: oral atypical anti-psychotics. Variable names are listed within the parentheses.

(1) The difference in the means (IAP minus OAP) divided by the square root of the average of the variance in the 2 groups.



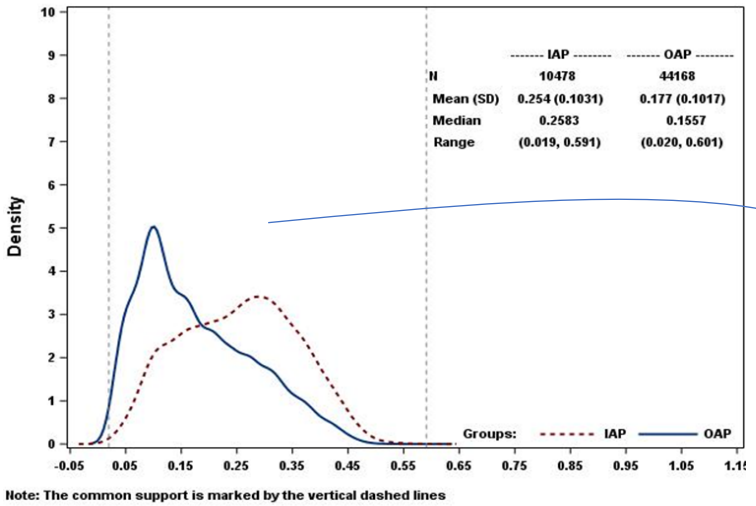
Note: The common support is marked by the vertical dashed lines

- Stratum 1, n=20
- Stratum 2, n=20
- Stratum 3, n=20
- Stratum 4, n=20
- Stratum 5, n=20
- Stratum 6, n=20
- Stratum 7, n=20
- Stratum 8, n=20
- Stratum 9, n=20
- Stratum 10, n=20

IAP, n=200

- Stratum 1, n=20
- Stratum 2, n=20
- Stratum 3, n=20
- Stratum 4, n=20
- Stratum 5, n=20
- Stratum 6, n=20
- Stratum 7, n=20
- Stratum 8, n=20
- Stratum 9, n=20
- Stratum 10, n=20

OAP, n=200



Stratum 1, n=20
Stratum 2, n=20
Stratum 3, n=20
Stratum 4, n=20
Stratum 5, n=20
Stratum 6, n=20
Stratum 7, n=20
Stratum 8, n=20
Stratum 9, n=20
Stratum 10, n=20

IAP, n=200

Stratum 1, n=20
Stratum 2, n=20
Stratum 3, n=20
Stratum 4, n=20
Stratum 5, n=20
Stratum 6, n=20
Stratum 7, n=20
Stratum 8, n=20
Stratum 9, n=20
Stratum 10, n=20

OAP, n=200

IAP, Π(s)

IAP, Π(s)

Simulation Results

Table 2 - Mean and Standard Deviation of SMD_{200}[1] Under No Up-Front Matching and Up-Front Matching

Variables	No Up-Front Matching		Up-Front Matching	
	Mean SMD_{200}	Standard Deviation of SMD_{200}	Mean SMD_{200}	Standard Deviation of SMD_{200}
Number of anti-psychotics prescribed (*n.antipsychotics*)	0.459	0.102	0.002	0.079
Age at index date, years (*age*)	-0.030	0.102	0.012	0.099
Number of psychiatric hospitalizations (*num.psych.hosp*)	0.206	0.097	0.001	0.096
Number of psychiatric drugs prescribed (*n.all.psychiatric.drugs*)	0.109	0.097	-0.009	0.098
Gender (*male*)	-0.198	0.105	0.009	0.093
Race (*white*)	-0.137	0.098	0.002	0.095
Race (*black*)	0.139	0.098	0.000	0.101
Race (*Hispanic*)	0.008	0.099	-0.002	0.105
Race (*other*)	-0.010	0.106	-0.001	0.101

(1) Standardized mean difference based on samples of size 200 across 500 studies.

Conclusions and Remarks

- Our simulation illustrates a major benefit of up-front matching: it creates populations of patients whose balance on the covariates for which matching was implemented is comparable to what would be achieved with randomization.

- Although up-front matching is based on only a subset of covariates, it is anticipated that it will provide a database that enables more robust and efficient estimates of treatment effect than using no matching at enrollment.

- These benefits are desirable even in POS's not intended for regulatory purposes but will be even more valuable for studies whose results become part of the evidence for regulatory decision making – their results will be more credible and there is the potential for significant cost efficiency in generating the data.

- In addition to potential efficiency gains based on balance there is the real possibility that the percentage of patients not in the common support will be relatively substantial, and the cost savings in not following such patients in a POS could be substantial.