



AN INTRODUCTION TO PROPENSITY SCORE ANALYSIS FOR NON-RANDOMIZED CLINICAL TRIALS

RAJESH KUMAR GAJENDRAN

SIDDHA CENTRAL RESEARCH INSTITUTE, MINISTRY OF AYUSH, GOVT. OF INDIA

Disclaimer: This presentation is only for the educational/informational purpose and no official support or endorsement by the CCRS or SCRI of this presentation is intended or should be inferred.

- ❖ Randomized controlled clinical trials are the “**standard rule**” for clinical trial studies.
- ❖ But there are situation where the randomization is not possible.
 - ❖ **Observational studies.**
 - ❖ **Device clinical trial (due to ethical problem)**
 - ❖ **Phase IV post marketing trials.**
- ❖ Due to **non-randomized** there arise a bias in statistical analyses, potentially due to **confounding covariates** that needs to be addressed.
- ❖ This presentation provides an introduction to a methodology that address the bias due to **heterogeneity** and imbalance in comparative clinical studies.

EFFICACY (CLINICAL RESEARCH)

EFFICACY JUDGED WITH CONTROLLED ENVIRONMENT

EFFECTIVENESS (CLINICAL PRACTICE)

EFFECTIVENESS IS THE REAL TEST OF A DRUG IN LARGER POPULATION

Randomized trial

Patient with single disease

Compliance ensured

Strict inclusion and exclusion criteria

Pros - Homogeneity

Cons - Generalizability

Bias: addressed through proper planning, randomization, ensuring controlled environments.

Non-Randomized trial

Patient with varied disease, organ system function, Multiple concomitant drug usage etc.

Compliance are not always ensured

No strict inclusion or exclusion criteria

Cons - Heterogeneity

Pros - Generalizability

Bias ?

Say in a village Peoples are classified into two Ethnic groups and due to particular disease

Ethnic Type	Died	Total Population	Death Percentage
Ethnic - A	10	100	10%
Ethnic - B	6	30	20%

What can we conclude from above? Does Ethnic-A are more prone to die due that particular disease?

Ok Fine now New Drug is given to **Ethnic-A** and Active Control have been given to **Ethnic-B** to overcome the above disease and the results are given below, what can we infer about the drugs?

Treatment Group	Died	Total Population	Death Percentage
Ethnic-A (New Drug)	5	90	5.55%
Ethnic-B (Active control)	5	24	20.83%

We cant able to conclude only the Drug effect here, since the all the **Ethnic-B** peoples in the village are very old in age compared to **Ethnic-A** and more likely to die **even if there is no disease**. So we need to match with the same age group of **Ethnic-A with Ethnic-B to get a valid conclusion**.

Propensity Score (PS)

When under uncontrolled environment and where proper randomizations and are not in place for a trial, Propensity Score Analysis can help researchers to minimize bias and draw valid conclusions.

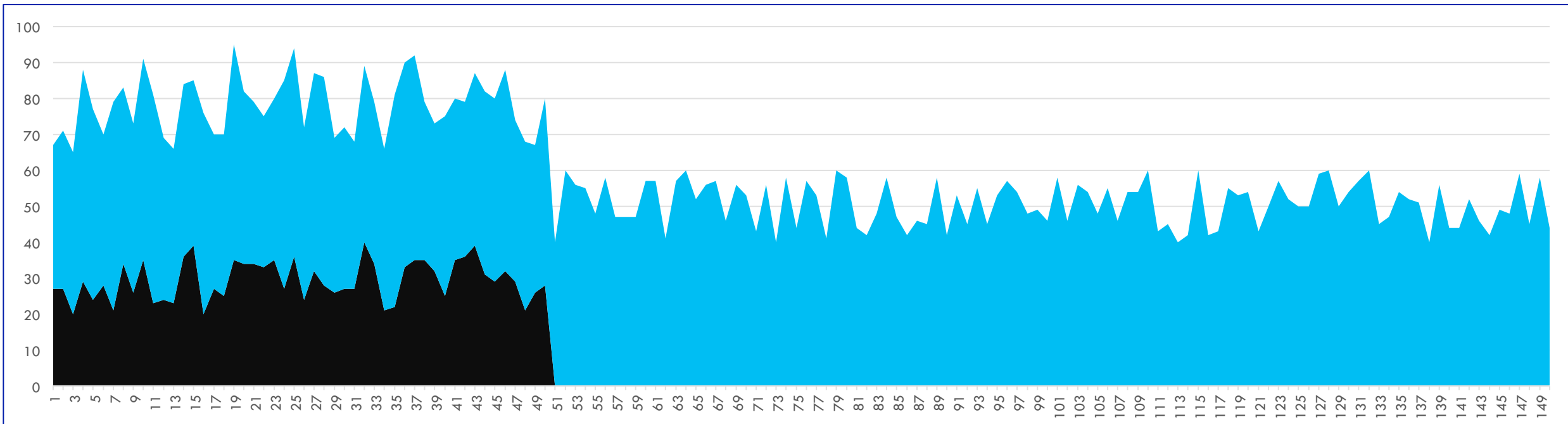
- PROPENSITY SCORES CAN PRODUCE SIMILAR CONTROL GROUP TO COMPARE WITH TREATMENT GROUP EVEN IF NOT RANDOMIZED.
- THE SIGNIFICANT COVARIATE CAN BE IDENTIFIED AND THE CONFOUNDING EFFECT CAN BE ADJUSTED TO ESTIMATE ONLY THE TREATMENT EFFECT.
- THE STUDY DESIGN CAN BE ADJUSTED TO REDUCE NON-RANDOMIZED BIAS.

A **propensity score** is the probability of a unit (e.g., Patient/subject) being assigned to a particular treatment given a set of observed covariates.

PS Definition e.g.

- ✓ Propensity score methods identify data-driven matches – control groups – for treatment groups. Statistically-similar control groups are identified for each treatment group using the observable characteristics of the treatment groups.

New Drug (Post Marketing)	Active Control	Total
50	150	200



- ❖ As per the Propensity Score definition we need to calculate the probability of each subjects being assigned to a treatment or not.
- ❖ To calculate this we need to Identify the possible covariates for each subject, including gender, age, race, BMI, height, weight, blood pressure etc.
- ❖ We can use Odds to calculate the probability of each subject being assigned to a treatment.

Where Odds are known to be simple functions of probabilities, defined as:

$$\text{Odds of an Event} = \frac{\text{Probability(Treated)}}{\text{Probability(Not Treated)}}$$

- ❖ Logistic Regression helps us to calculate the probability score of a subject.

$$\text{Model: Treatment (0,1)} = \langle \text{Covariates} \rangle$$

Simple-SAS Code

The Proc Logistic can be used to calculate the Propensity Score

```
Proc logistic data=Prop;  
Class trt gender Ethnicity;  
Model trt = gender Ethnicity age BMI  
          height weight.....etc.;  
Output=ps predicted=pred;  *Propensity Score;  
run;
```

Propensity score matches the control patient with treated patient using the subjects with similar probability, which means a quasi-randomized design.

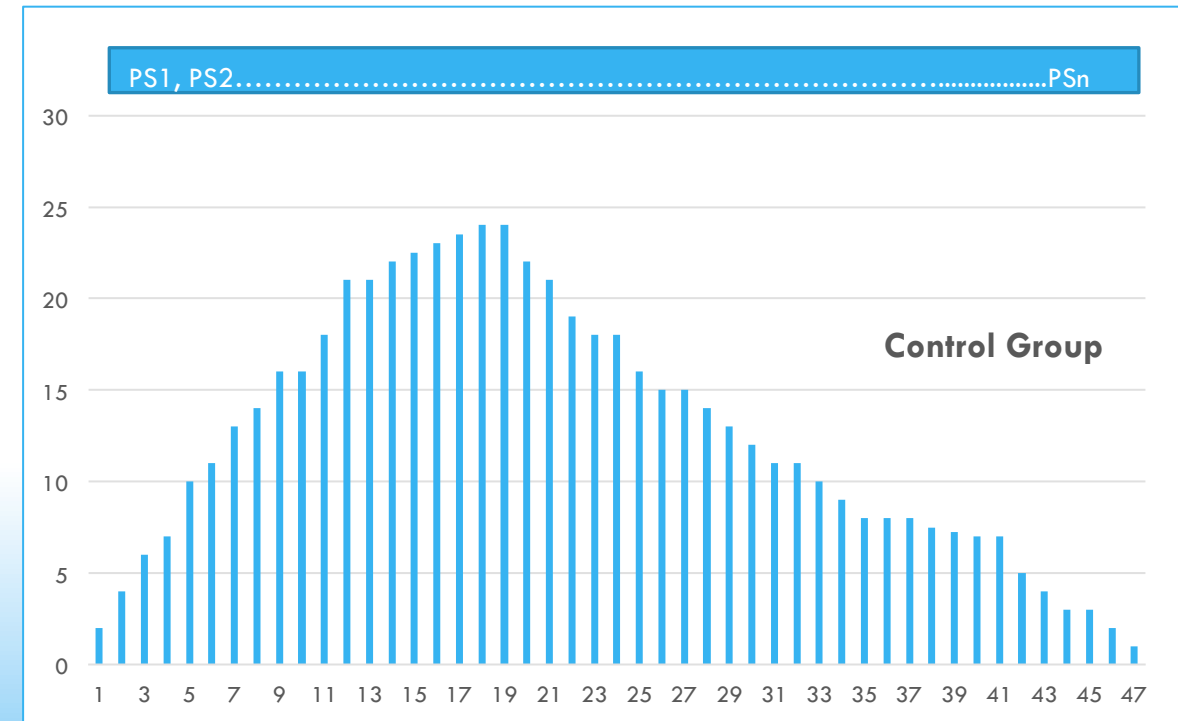
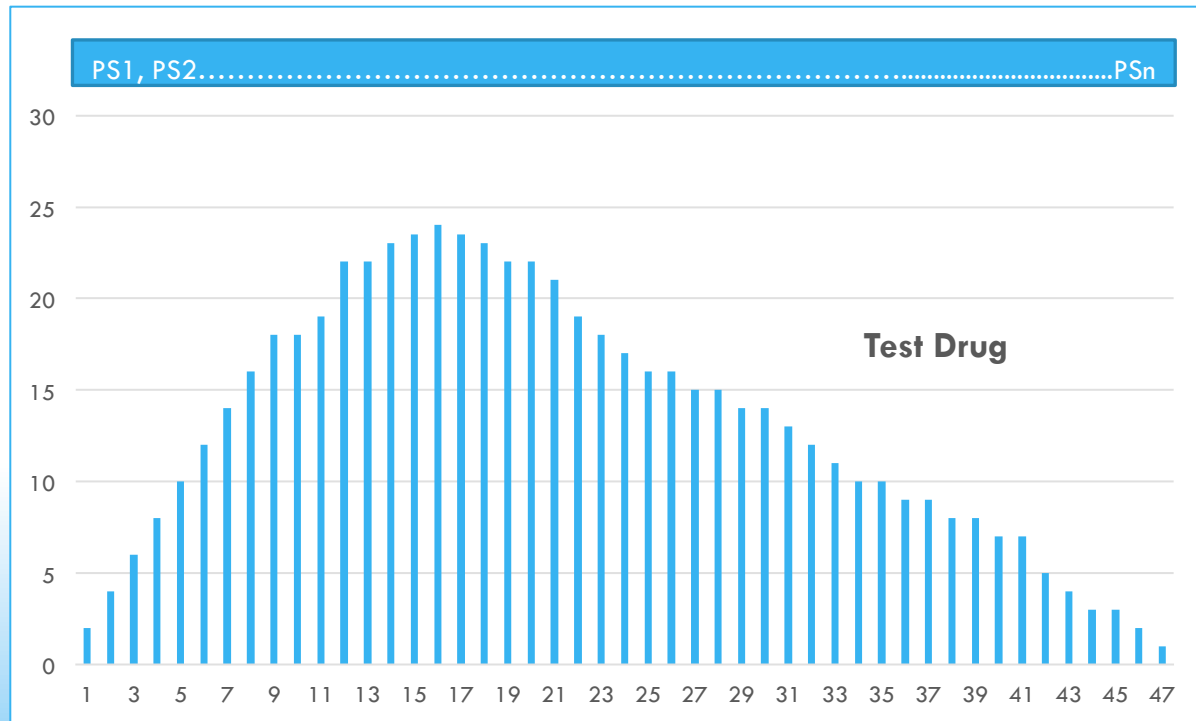
Matching methods:

- ✓ Find similar control patients with matching propensity score as in treated patients.
- ✓ Stratification
- ✓ Regression

Methodology

General Matching Method:

Matching method using similar propensity score for test and control group



Cons:

May need to exclude the unmatched patients.

The control group patient should be enough large to get a match with test drug group.

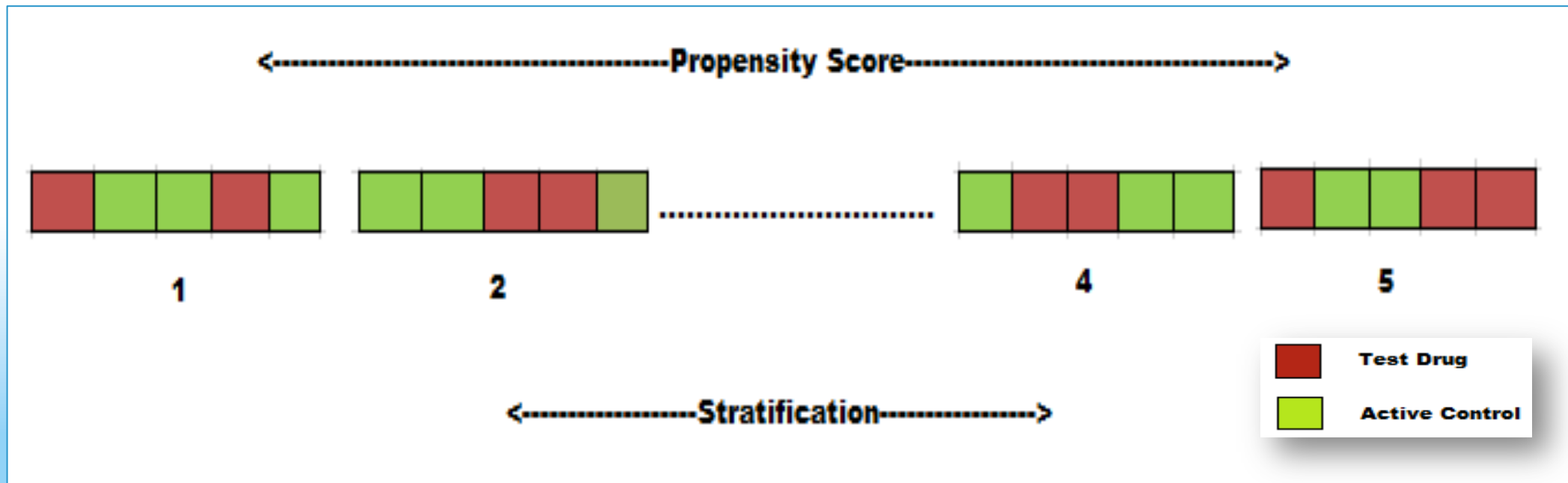
Methodology

Stratified Matching Method:

Propensity score are sorted and stratified based on Propensity quintile.

5 equal stratified blocks (Stratified based on the quintiles of a continuous confounding variable eliminated approximately 90% of the bias due to that variable - Rosenbaum and Rubin (1984) Cochran (1968))

Both the treatments are compared within each block and overall treatment effect is estimated using weighted average method.



Cons:

Some stratified class may have comparatively less patient.

Regression Method (Adjustment of Covariate)

Propensity scores are calculated using the logistic regression with treatment as dependent and all other co-variables are independent.

Then the estimated propensity score (probability) is incorporated into the treatment effect estimate model as one of the independent variable.

Dependent variable = independent variables, covariates, **Propensity score.**

After adjusting for propensity score the significance of co-variates should be checked

❖ Obtain the p-value for each co-variate as below

➤ For continuous variable like age, BMI ANOVA model can be used to obtain the p-value:

$$\text{AGE (DEPENDENT)} = \text{TREATMENT (INDEPENDENT)} + \text{<(PROPENSITY SCORE)>}$$

➤ For categorical variables like Ethnicity, Gender Logistic model or Multinomial Logistic can be used to obtain the p-value:

$$\text{ETHNICITY (DEPENDENT)} = \text{TREATMENT (INDEPENDENT)} + \text{<(PROPENSITY SCORE)>}$$

Post PS Analysis

Example

Co-variate	Age	Ethnicity	Gender	BMI
Pre-PS	0.019	0.036	0.025	0.010
Post-PS	0.59	0.45	0.85	0.016

Age, Ethnicity & Gender p-value: The 1st row p-value for Age, Ethnicity & Gender, indicates that there is a significant difference between the two treatments before adjusting for PS, but after incorporating the propensity score the p-value turns to be insignificant. Hence the Age, Ethnicity & Gender is balanced between the two treatments.

BMI p-value: But for BMI we still see that it is significant even after incorporating Propensity Score in the model; which means incorporating the Propensity Score in the model is still not enough to prevent bias in the conclusion. So we have to carefully attempt to interpret our end point conclusion keeping in mind that BMI is a hidden confounder variable in our study.

- The propensity score methods only allow one to account for measured baseline variables. Estimates using each of the estimates of treatment effect may be susceptible to bias due to unmeasured confounding variables.
- Splitting propensity score into quintiles, may reduce the sample size within each quintile if our overall sample size isn't very large.
- A high degree of statistical expertise is required in handling issues like pre-specification of clinically relevant covariates to be measured, suitable patient populations, planning of sample size in the context of propensity score methodology, handling missing covariates in generating propensity scores, and assessing the success of the propensity score method by evaluating treatment group overlap in terms of the distributions of propensity scores.
- Using propensity scores analysis methods may be helpful, but one can usually not be certain that we are accounting for all biases due to variables that we are not aware of.

Reference*

1. **Tutorial and case study in propensity score analysis: an application to estimating the effect of in-hospital smoking cessation counselling on mortality**
2. **Medical Device Clinical Trials: What We Should Know, Karen Rosales, Ph.D**
3. **Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group ralph b. D'AGOSTINO, jr.**
4. **Clinical investigations of medical devices – statistical considerations MHRA**
5. **Methodology for Non-Randomized Clinical Trials: Propensity Score Analysis Dan Conroy, Ph.D., inVentiv Health, Burlington, MA**
6. **Propensity score analysis and assessment of propensity score approaches using sas® Procedures rheta E. Lanehart, patricia rodriguez de gil, eun sook kim, aarti P. Bellara, jeffrey D. Kromrey, and reginald S. Lee, university of south florida.**

*** Disclaimer: The use of copyrighted material in this presentation, if any, is fair use for non-profit educational/information purpose only.**



QUESTIONS/DISCUSSION



THANK
YOU ALL