

# Propensity-Score Matching Workshop

*Holger Langkabel*



---

# Introduction: Confounding in Non-Randomized Settings

---

**Assessing Balance**

**The Propensity Score**

**Matching & Post-Matching Balance Assessment**

**Treatment-Effect Estimation**

**Wrap-Up**

# The Evaluation Problem

Non-interventional studies (NIS)/observational data:

- Due to the self-selection of the subjects into treatment and control group, some covariate that is influencing outcome might be unbalanced between the group (e.g. sicker people are more likely to join the treatment group).
- As a result, estimates of a naïve between-group comparison will be biased.
- This is referred to as the “evaluation problem”.

# Imbalances in Randomized Studies

Randomized studies:

- No self-selection: Randomization results in automatic balance between treatment and control group.
- BUT: If results of different randomized controlled trials are combined, there might be imbalances due to different inclusion/exclusion criteria (or due to chance).

# Statistical Background

## *How is Propensity-Score Matching solving the evaluation problem?*

The difference between treatment and control group is not calculated at gross but at individual level. To do this, two subjects (one of the treatment and one of the control group), which share equal (or at least similar) baseline characteristics, are matched.

If then a conditional independence assumption (CIA) holds (cannot be tested), it can be shown that

$$ATT = E(y_{1i}|T = 1, \mathbf{x}_i) - E(y_{0i}|T = 0, \mathbf{x}_i).$$

# General Procedure of Propensity-Score Matching

- Step 0: Check for imbalances between control and treatment group.
- Step 1: Estimate the propensity model.
- Step 2: Predict individual propensity score.
- Step 3: Match controls on treated subjects.
- Step 4: Check if propensity score matching reduced the imbalances. If necessary, repeat 1 through 3.
- Step 5: Estimate treatment effect.

# Introduction: Confounding in Non-Randomized Settings

---

## Assessing Balance

---

## The Propensity Score

## Matching & Post-Matching Balance Assessment

## Treatment-Effect Estimation

## Wrap-Up

## Step 0 / Step 4: Check for Imbalances

- Compare means and standard deviations
- Calculate standardized difference
- Compare medians and other quantiles
- Compare boxplots
- Compare proportions of subjects in different subgroups
- Look at histograms/kernel density plots
- ...

The use of statistical tests is generally discouraged.



# Some Useful R-Code

```
str(data)
```

```
summary(data[which(data$wearable==1),])
```

```
summary(data[which(data$wearable==0),])
```

```
tapply(data$age, list(data$wearable), <mean|sd>)
```

```
aggregate(data$age, list(data$wearable), summary)
```

```
boxplot(age ~ wearable, data=data)
```

```
tab <- CreateTableOne(  
  vars=c("sex", "age", "height", "weight", "bmi"),  
  strata="wearable", data=data, test=FALSE)  
print(tab, smd=TRUE)
```

# **Introduction: Confounding in Non-Randomized Settings**

## **Assessing Balance**

---

## **The Propensity Score**

---

## **Matching & Post-Matching Balance Assessment**

## **Treatment-Effect Estimation**

## **Wrap-Up**

# Dimensionality Problem

## Problem:

If too many variables are considered for matching, it will become very unlikely to find two subjects with similar characteristics (“curse of dimensionality”).

## Solution:

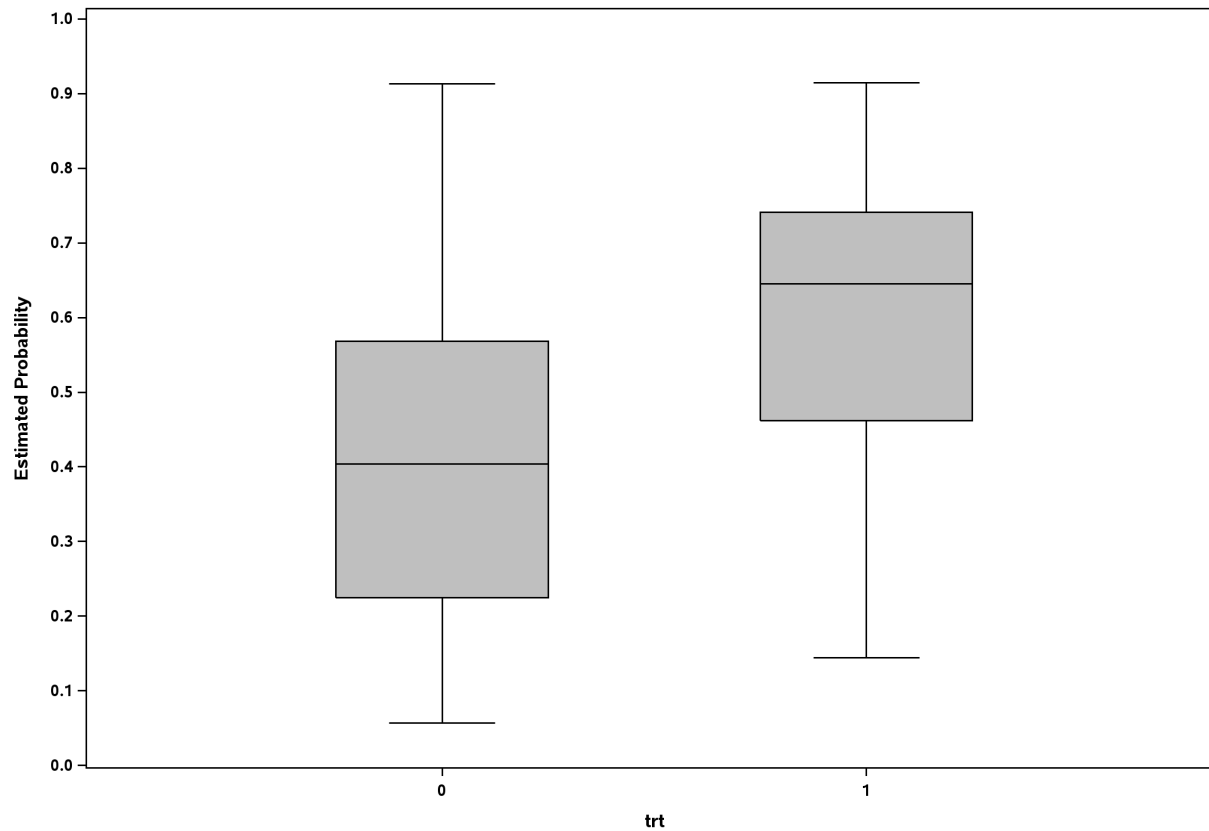
Matching is not done on the full vector of covariates but on the propensity score (i.e. the conditional probability of participation in the treatment group). This approach is still statistically valid.

# Common Support

- Common support is the condition that for all values of the propensity score there are both treated and untreated subjects.
- Common support is crucial to the successful use of matching and to the generalizability of the results.

# Example of Good Common Support

Boxplot to check for common support



Box: 25th to 75th percentile; horizontal line: median; vertical lines extend from the box as far as the data extends, including outliers.

# Step 1: Estimate Propensity Model

Since the propensity is the probability of participating in the treatment group, it can be modelled by any binary response model (e.g. probit or logit model) with treatment group as outcome.

## Step 2: Predict Propensity Scores

Easy from the programmer's perspective:

```
predict(model, type="response")
```

# Some Useful R-Code

```
model <- glm(wearable ~ age + bmi,  
             family=binomial(link="logit"), data=data)  
summary(model)
```

```
ps <- predict(model, type="response")  
summary(ps)  
boxplot(ps ~ wearable, data=data)
```

```
dataps <- data.frame(data, ps)  
sm.density.compare(dataps$ps, dataps$wearable)  
legend("topright", c("0", "1"), fill=2+(0:1))
```



# Introduction: Confounding in Non-Randomized Settings

## Assessing Balance

## The Propensity Score

---

## Matching & Post-Matching Balance Assessment

---

## Treatment-Effect Estimation

## Wrap-Up

## Step 3: Match Controls on Treated

Matching methods:

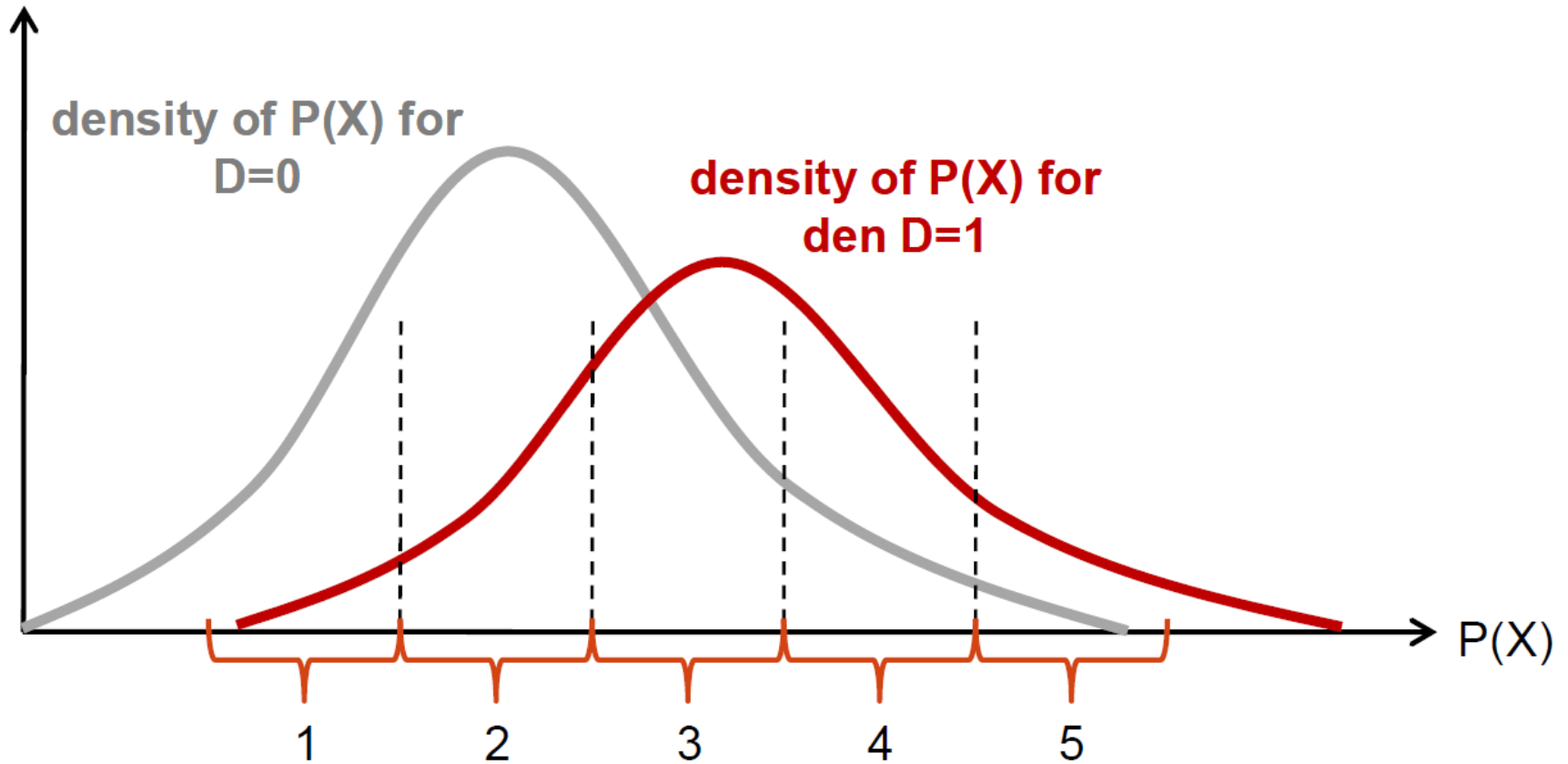
- ***k*-nearest-neighbor matching:** Use the  $k$  controls that are nearest to the treated subject in terms of the propensity score ( $k$  may be 1).
- **Caliper matching:** Use all controls that are within some specified radius (caliper).
- **Block-wise matching:** Stratify on the propensity score.
- **Kernel matching:** Weight the control group observations by their distance to the treated subject.

Matching can be done with or without replacement (i.e. one control can be matched to several treated subjects or only to one).

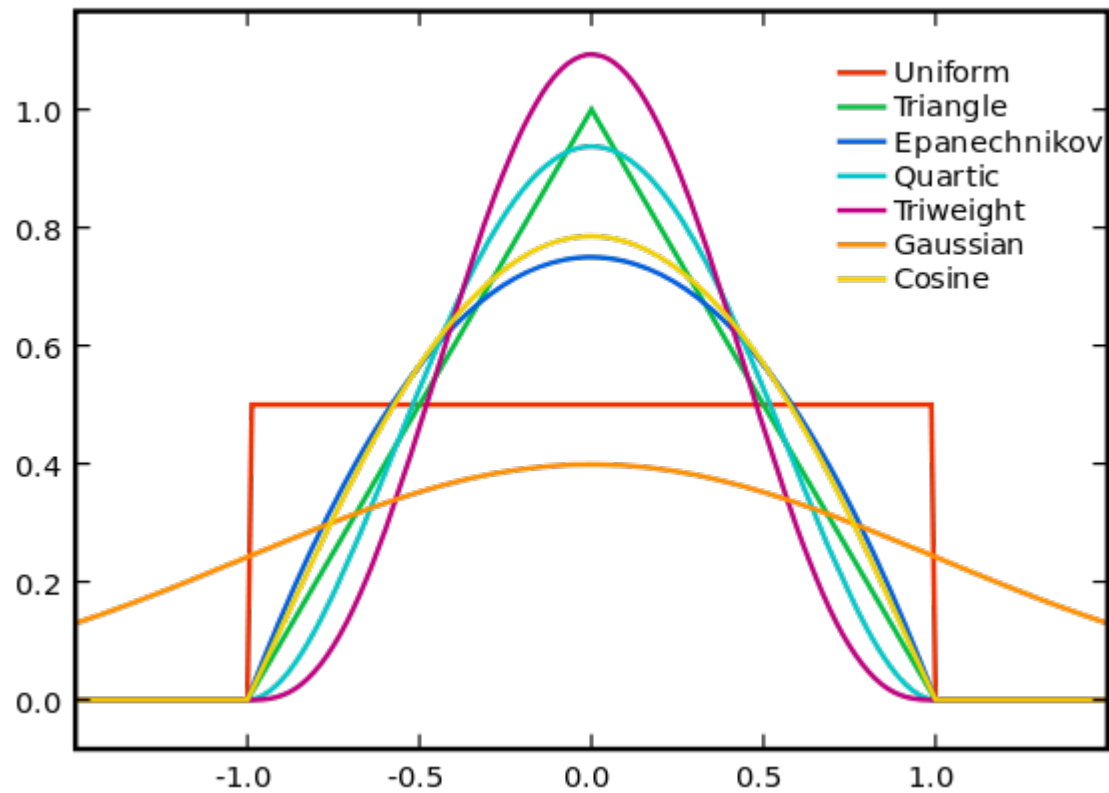
# Algorithm for 1-Nearest-Neighbor Matching Without Replacement

1. Repeat steps 2 through 4 for all treated subjects in random order of subjects:
2. Calculate the absolute distance of all control subjects in terms of the propensity score.
3. Keep all control subjects with absolute distance less than the specified caliper.
4.
  - If step 3 returns no matches, proceed to next treated subject.
  - If step 3 returns exactly one match, save matched pair and remove control subject from further matching.
  - If step 3 returns more than one match, choose one control subject at random and remove this from further matching.

# Block-wise Matching



# Some Kernels



# Some Useful R-Code

```
matching <- matchit(wearable ~ age + bmi,  
                    distance="logit",  
                    method="nearest", data=data)  
datam <- match.data(matching)  
  
summary(matching)
```

# Introduction: Confounding in Non-Randomized Settings

## Assessing Balance

## The Propensity Score

## Matching & Post-Matching Balance Assessment

---

## Treatment-Effect Estimation

---

## Wrap-Up

## Step 5: Estimate the Treatment Effect

This will depend on the matching method applied. Formulas might be very easy or very complicated. Ask your statistician.

For 1-nearest-neighbor matching it is simply:

```
te <- lm(sixmwt ~ wearable, data=datam)
summary(te)
```



# **Introduction: Confounding in Non-Randomized Settings**

## **Assessing Balance**

## **The Propensity Score**

## **Matching & Post-Matching Balance Assessment**

## **Treatment-Effect Estimation**

---

## **Wrap-Up**

---

*Doing now what patients need next*