

# Propensity scores in R: Managing multiple scenario analyses in a single clean script

Kevin Deighton, [kdeighton@deltahat.co.uk](mailto:kdeighton@deltahat.co.uk)

Delta Hat, Nottingham, UK



# Contents

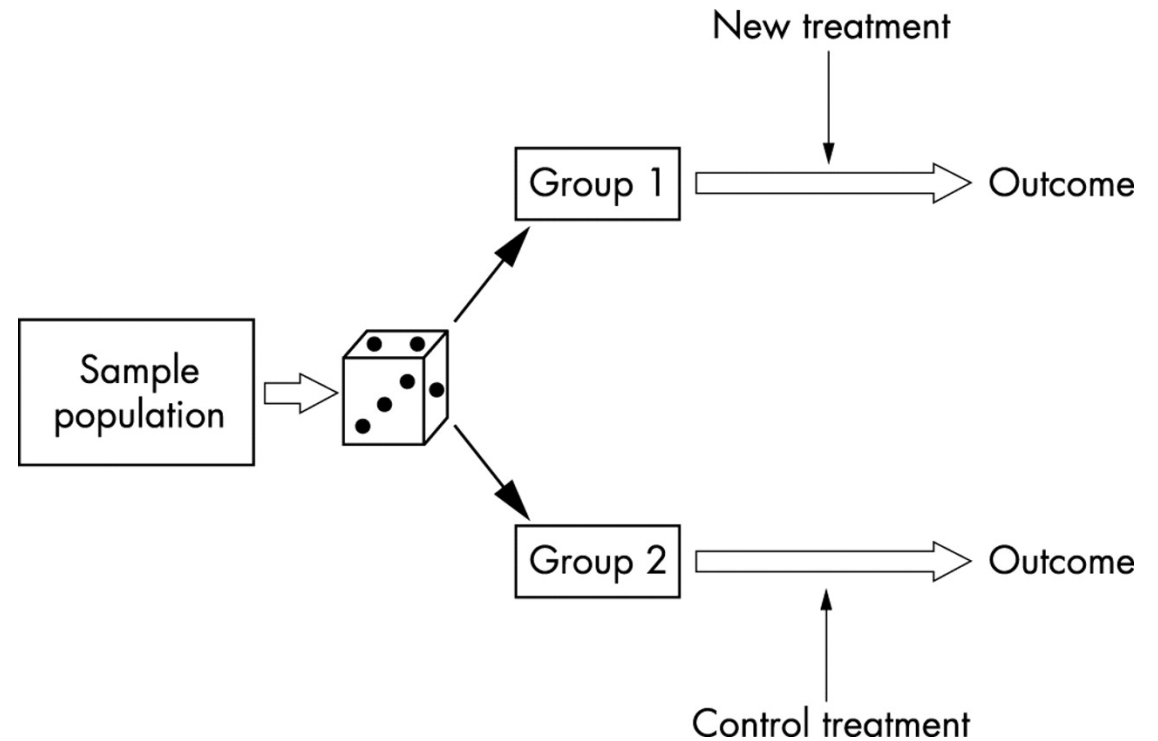


- Randomised controlled trials (RCTs)
- Single arm (uncontrolled) trials
- Propensity scores
- Stabilised Weights
- Example application
- Analysis workflow
- Coding demonstration

# Randomised controlled trials (RCTs)



- RCTs represent the ‘gold standard’ of study design
- Randomisation aims to ensure that factors affecting outcomes are evenly balanced between treatment and control groups
- Any differences in outcomes are therefore attributable to the intervention (i.e., the ‘treatment effect’)

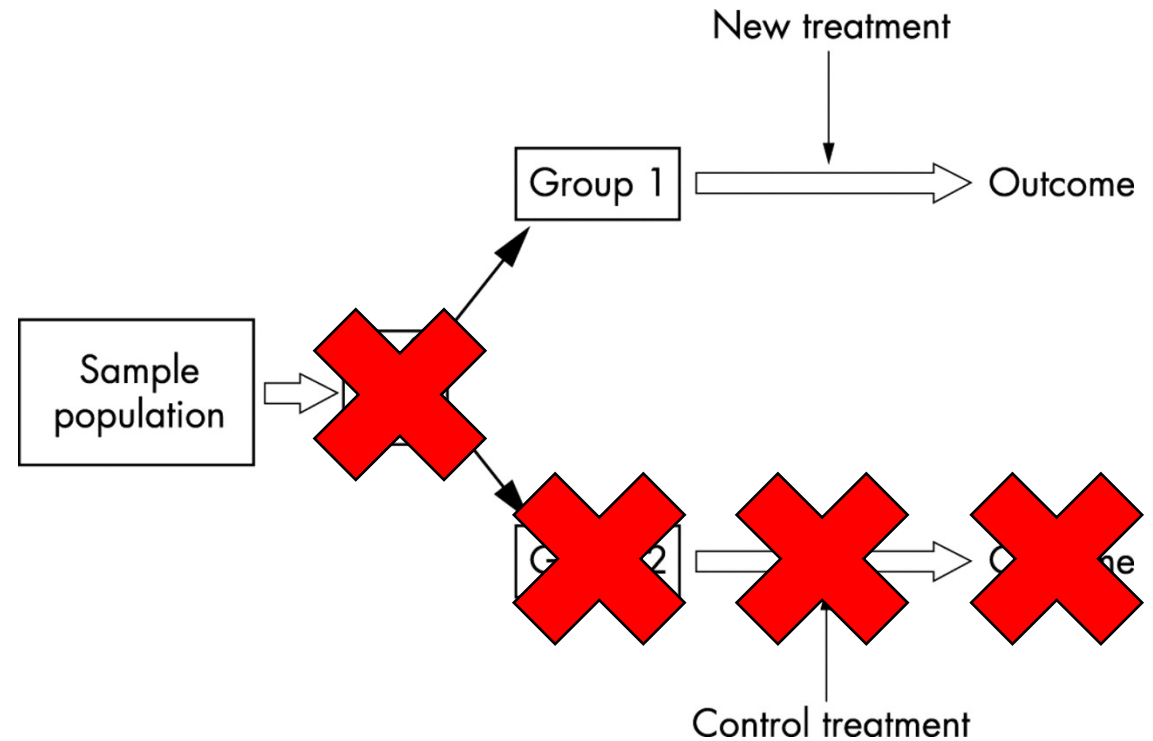


*Schematic extracted from Kendall (2003)*

# Single arm (uncontrolled) trials



- All participants are provided with the intervention, without the inclusion of a control arm
- Effectiveness of the intervention is determined using before-and-after comparisons
- Single arm trials may be required due to the lack of an acceptable control arm or ethical issues

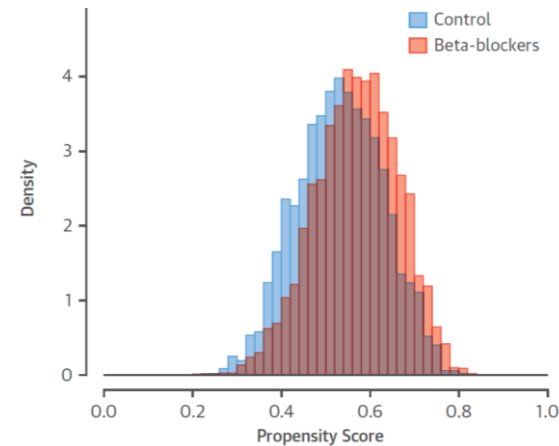


*Schematic adapted from Kendall (2003)*

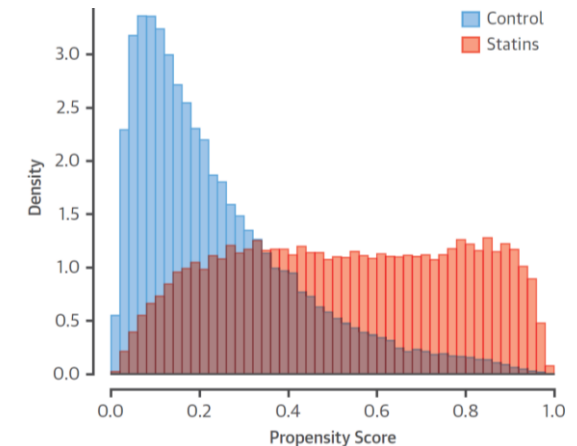
# Propensity scores



- Used to compare treatments when patient-level data are available for both comparators
- Represents the probability of treatment assignment as a function of a set of observable, prognostic covariates
- Estimated via logistic regression of exposure to the intervention (as a binary variable), given the selected set of covariates
- Aims to ensure ‘exchangeability’ between treatment groups



No extreme propensity scores, good overlap of treatment and control.



Some extreme propensity scores, poor overlap of treatment and control.

*Figures extracted from Elze et al. (2017)*

# Stabilised Weights



- Stabilised inverse probability of treatment weighting (‘Stabilised Weights’) can be used to produce balanced datasets for comparison
- Greater weighting is given to the observations which appear on one group but have a small probability of being found in that group

For the ‘treated’ group:

$$SW_i = \frac{p}{\pi_i}$$

For the ‘control’ group:

$$SW_i = \frac{(1 - p)}{(1 - \pi_i)}$$

The stabilised weight,  $SW$ , assigned for each individual,  $i$ , based on the proportion of treated patients,  $p$ , and the propensity score,  $\pi$

*Xu et al. (2010)*

# Example application



## Task overview

- Determine the comparative efficacy of a new drug (Drug B) to the standard of care (Drug A)
- Each drug was approved on the basis of single-arm clinical trials
- Similar inclusion/exclusion criteria were used for each drug trial
- Baseline and outcome measurements were the same between drug trials
- Required to present comparisons before and after propensity score adjustments for each analysis

## Analysis requirements

### Scenario/sensitivity analyses:

- Datasets (n = 3)
- Subsets of the primary dataset (n = 8)

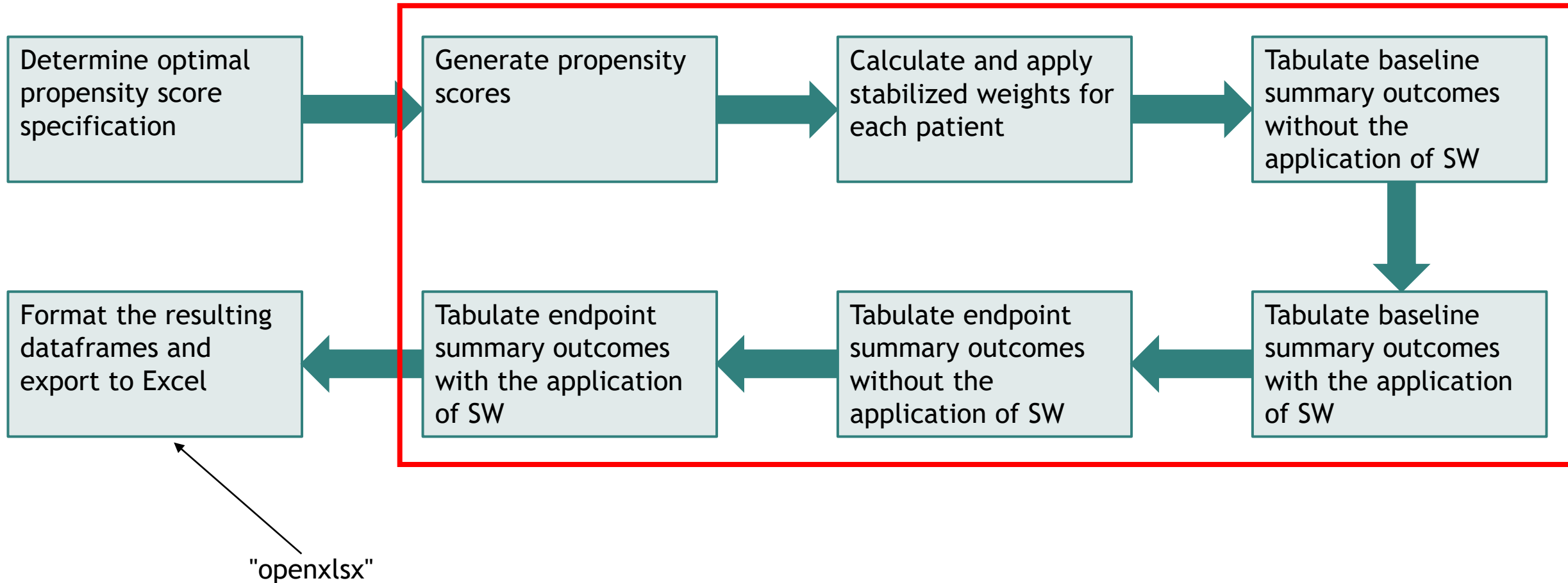
### Variables:

- Baseline variables (n = 18)
- Endpoint variables (n = 32)

### Summary outcomes:

- Mean
- Standard deviation
- Median
- Range
- N
- P-value

# Analysis workflow





# References



- Published studies:
  - Elze et al. (2017). Comparison of propensity score methods and covariate adjustment: evaluation in 4 cardiovascular studies. *Journal of the American College of Cardiology*, 69: 345-57
  - Kendall (2003). Designing a research project: randomised controlled trials and their principles. *Emergency Medicine Journal*, 20: 164-8
  - Xu et al. (2010). Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value in Health*, 13: 273-7
- R packages:
  - install.load: <https://cran.r-project.org/web/packages/install.load/install.load.pdf>
  - dplyr: <https://cran.r-project.org/web/packages/dplyr/dplyr.pdf>
  - weights: <https://cran.r-project.org/web/packages/weights/weights.pdf>