Numbers needed for behavior change: Introducing a practical measure of effect size that connects research to practice

Stefan Gruijters, Department of work and social psychology - Maastricht University

Gjalt-Jorn Peters, Department of methodology and statistics – Open University

https://psyarxiv.com/2bau7/

Claims in this presentation:

- Measures of effect size are important

  - *The primary product of a research inquiry is one **or more** measures of effect size, not P values. -Jacob Cohen (1990)*

- 'Traditional' indices do a **poor job** in conveying information about effectiveness in **intuitive** and **practical** ways

- There are alternative effect size indices that *can* convey information in a way that gives practically relevant and intuitive information.

- We suggest – for *behavior change intervention evaluations* – one may have particular merit, the ***Numbers Needed to Treat*** index.

Traditional effect size indices:

- Standardized mean difference (e.g. estimated with *d, g*)

- Binning effect sizes (e.g. small [.2 ; .4], medium[.4 ; ,8], large [.8 ; ∞])

- How many std. dev. change do we need to call an intervention 'succesful'?

Cohen's d: 0.6
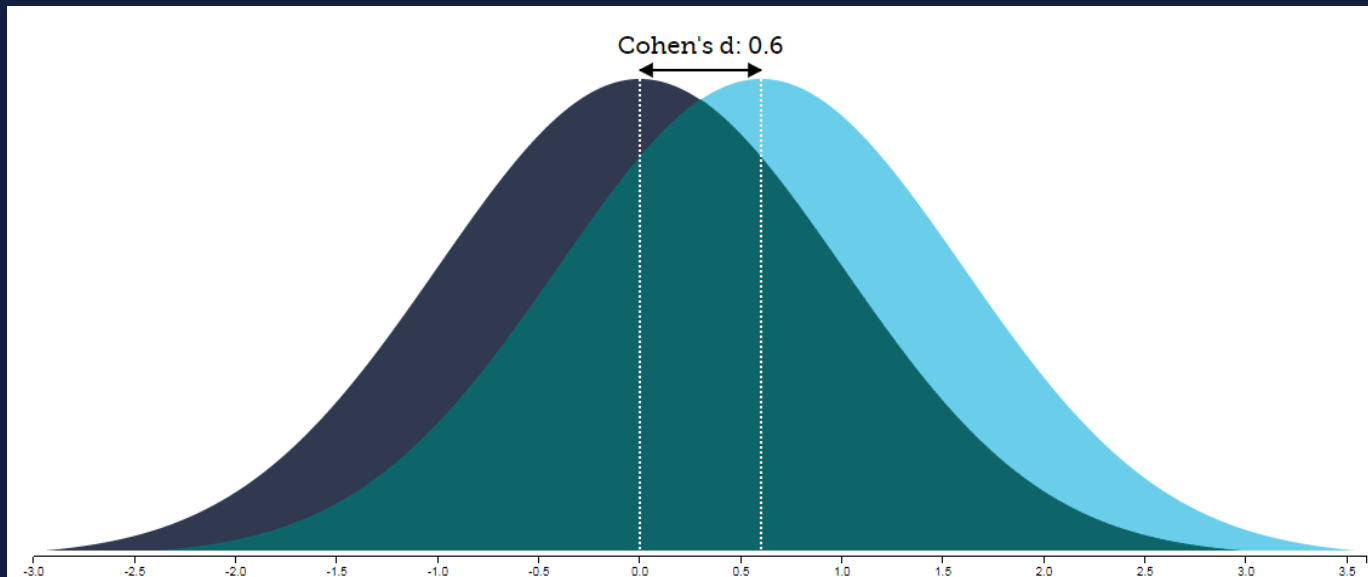
From: Kristoffer Magnusson http://rpsychologist.com/d3/cohend/

Traditional effect size indices:

- Proportion explained variance (e.g. R^2, eta^2, $omega^2$, etc.)

- "For most of the softer, wilder areas of psychology, squaring the correlation coefficient tends to make it go away – vanish into nothingsness as it were"

  - Rosenthal, R. (1990). How are we doing in soft psychology? American Psychologist, 45, 775–777. http://doi.org/10.1037/0003-066X.45.6.775

- Is an intervention that explains 8 % of the outcome variance 'succesful'?
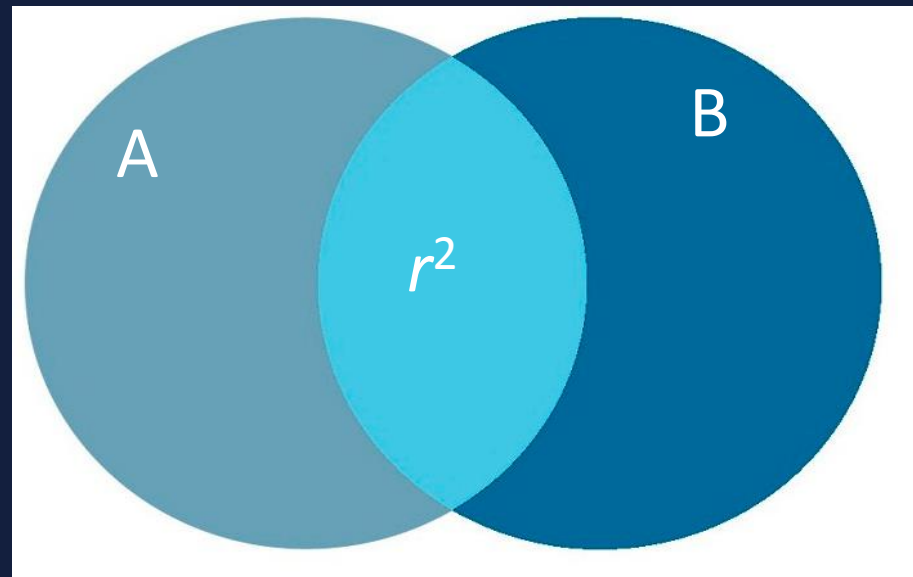
Alternative ways to express effect size magnitude intuitively

- **Probability of superiority (P[y2] > P[y1])**
  - Common language estimator (CL; Mcgraw & Wong, 1992)
  - The A estimator (Vargha & Dalaney, 2000), related to AUC
  - These estimate: **What is the probability that someone in the experimental group scores higher than someone in the control group**
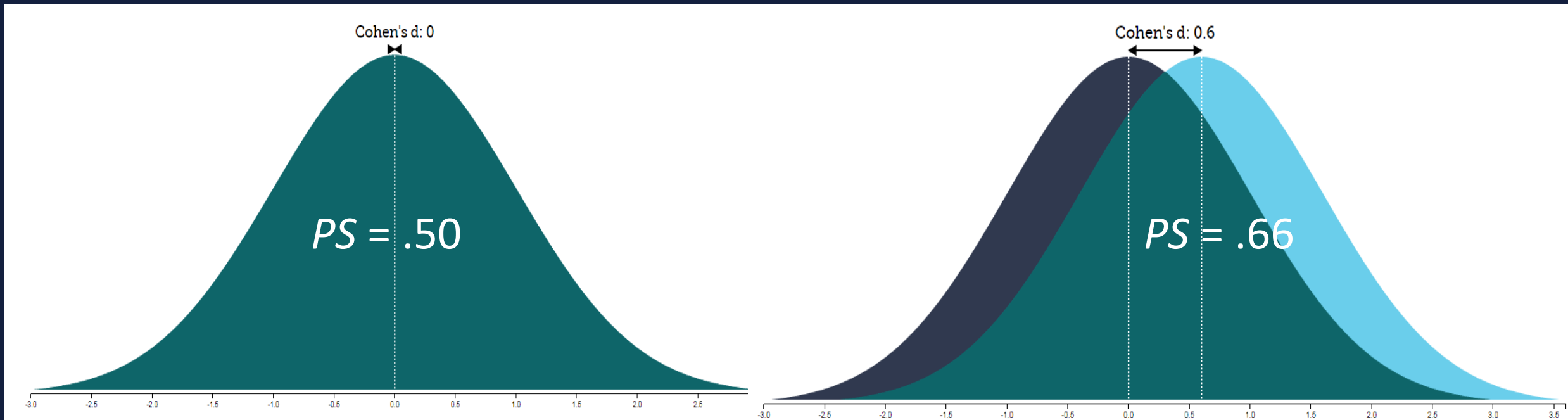
Why not use a simple frequency format?

- The Numbers Needed to Treat (NNT) index (e.g. Cook & Sackett, 1995; Laupacis et al., 1988)

- We prefer **Numbers Needed for Change (NNC)** in the health psychology context

- **How many people do we need to expose to the intervention to have one more desired outcome compared to the control condition**

- E.g., NNC = 6
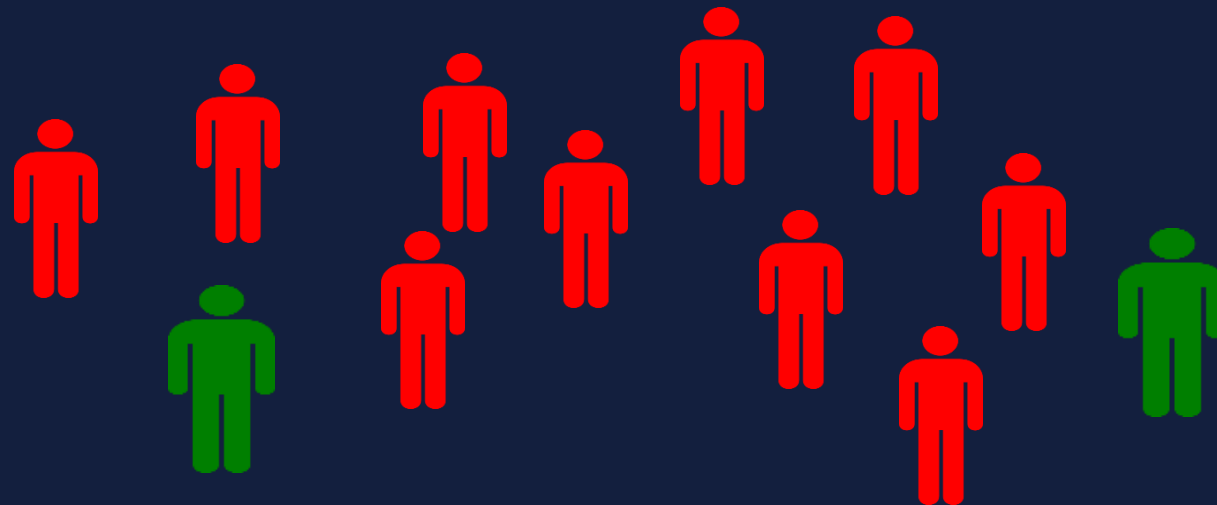
NNC for a binary outcome

- Example: physical excercise (desired/undesired levels)
- NNC = 1/(EER–CER) = 1 / success rate difference
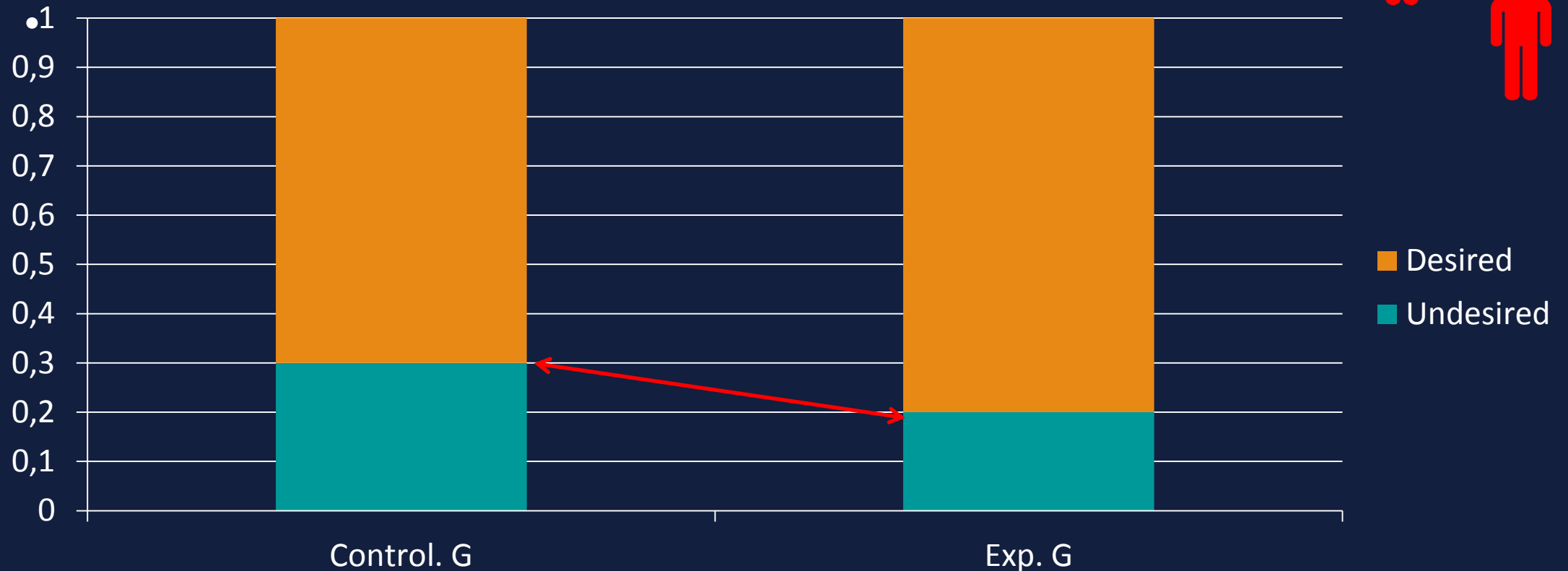
NNC can also be reported for continous outcomes

- We need to know **two** things (see Furukawa, 2011):

- How many *desireable* and *undesireable* events in the control group?
  - Aka. **Control Event Rate**

- What is the effect of the intervention in terms of Cohen's *d*
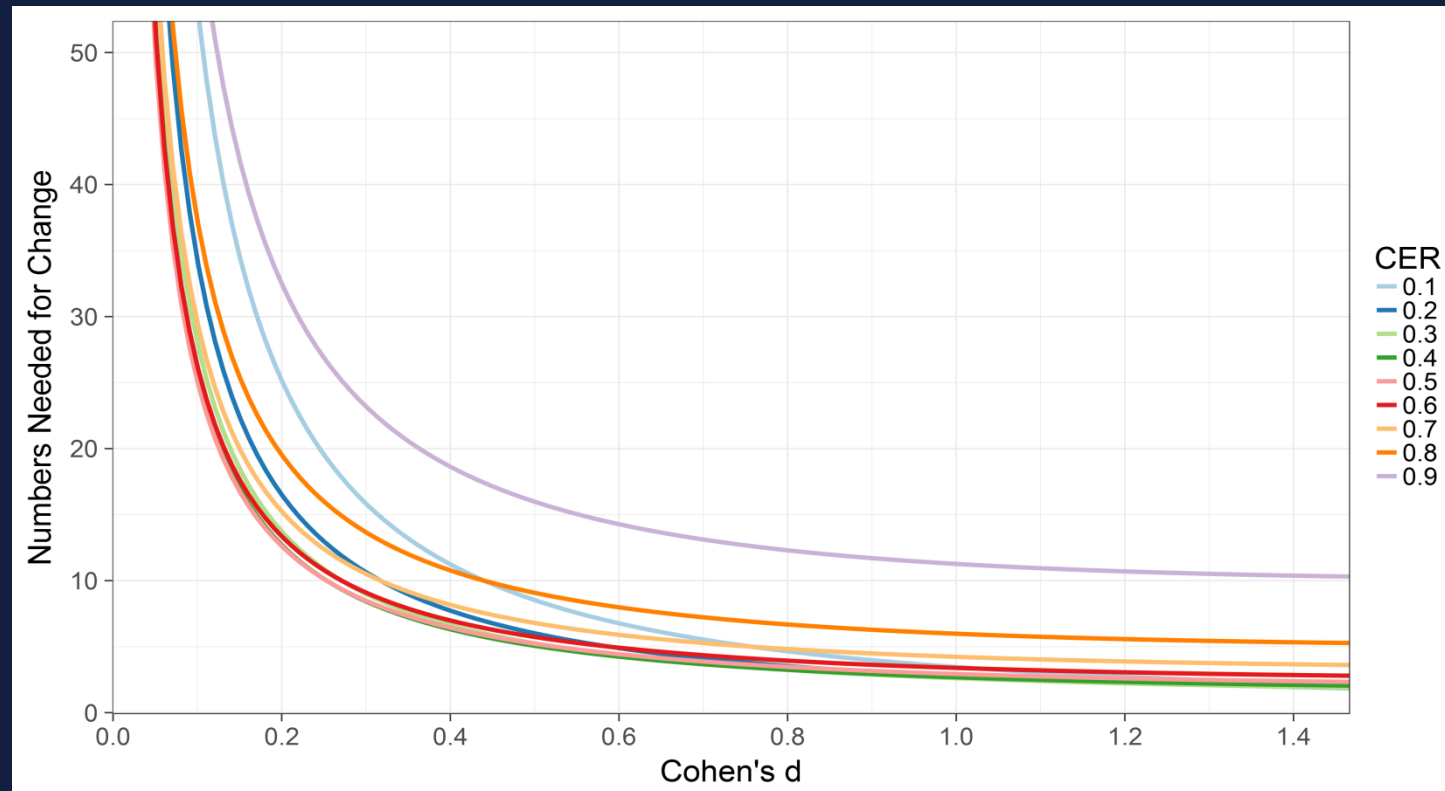
Estimating the control event rate (CER)

- We need to know how many people are already showing desired behavior
  - Dichotomize the scores on 'desired' versus 'undesired'
- The CER depends on our definition of 'desired' versus 'undesired'
- For example, control group mean = 90, sd = 30 on minutes of physical excercise per week
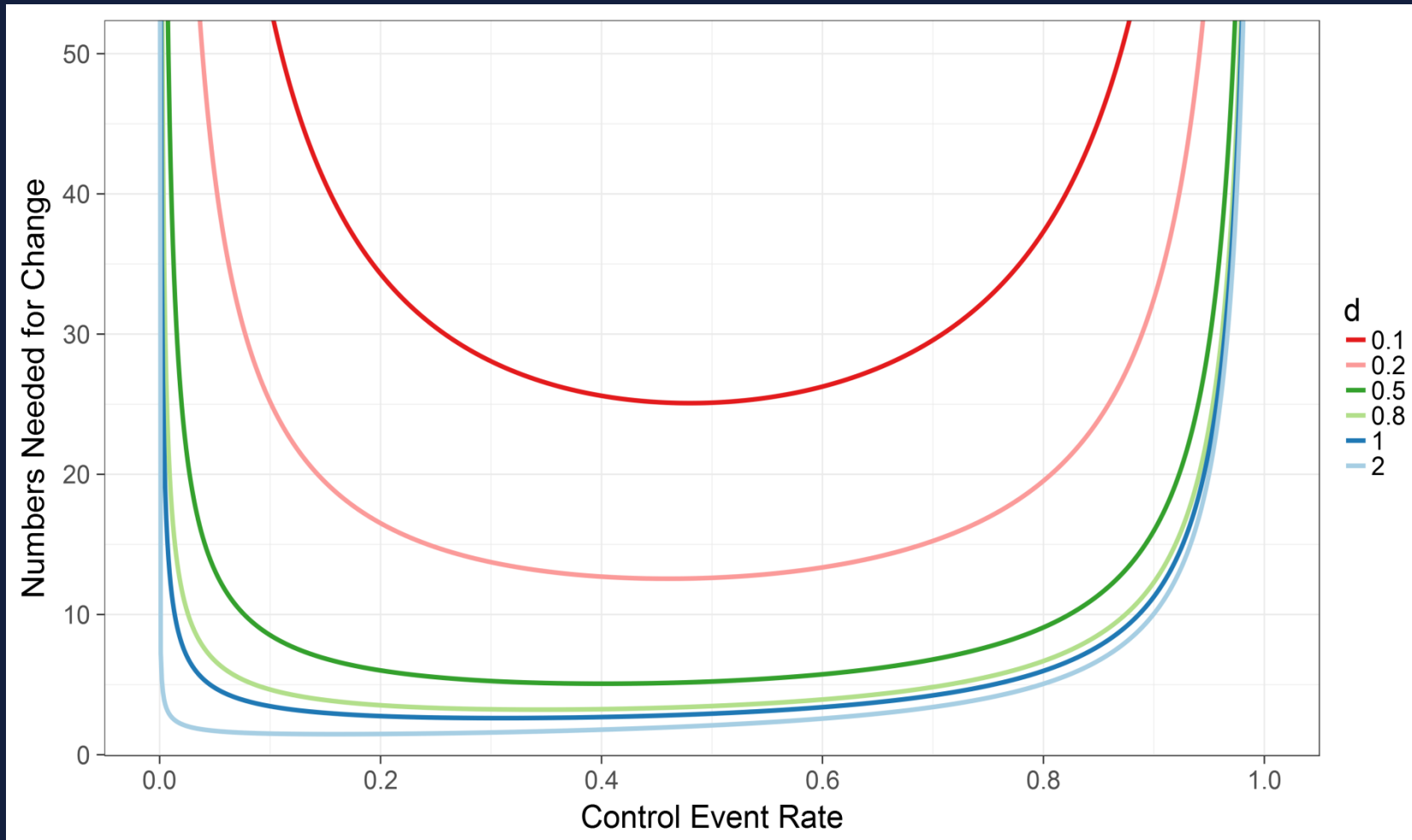

CER = .50
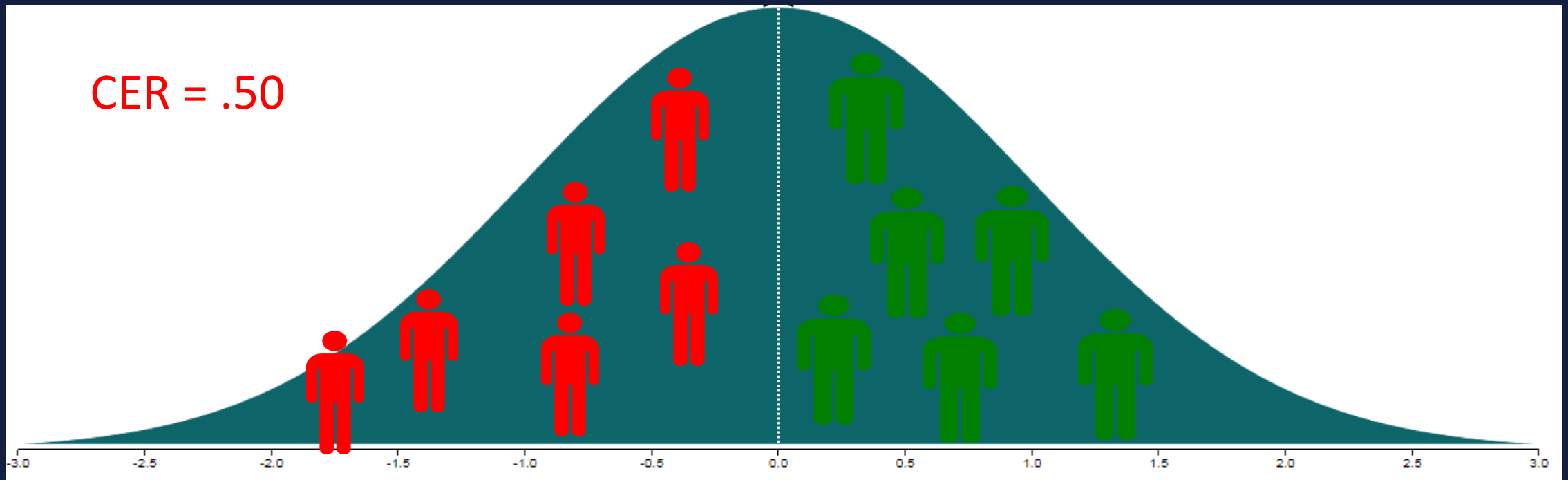
Estimating the control event rate (CER)

- But it is not required to assume that > control group mean (90 minutes per week sports) counts as 'desired'.

- For example, we could define > 120 minutes per week as 'desired'
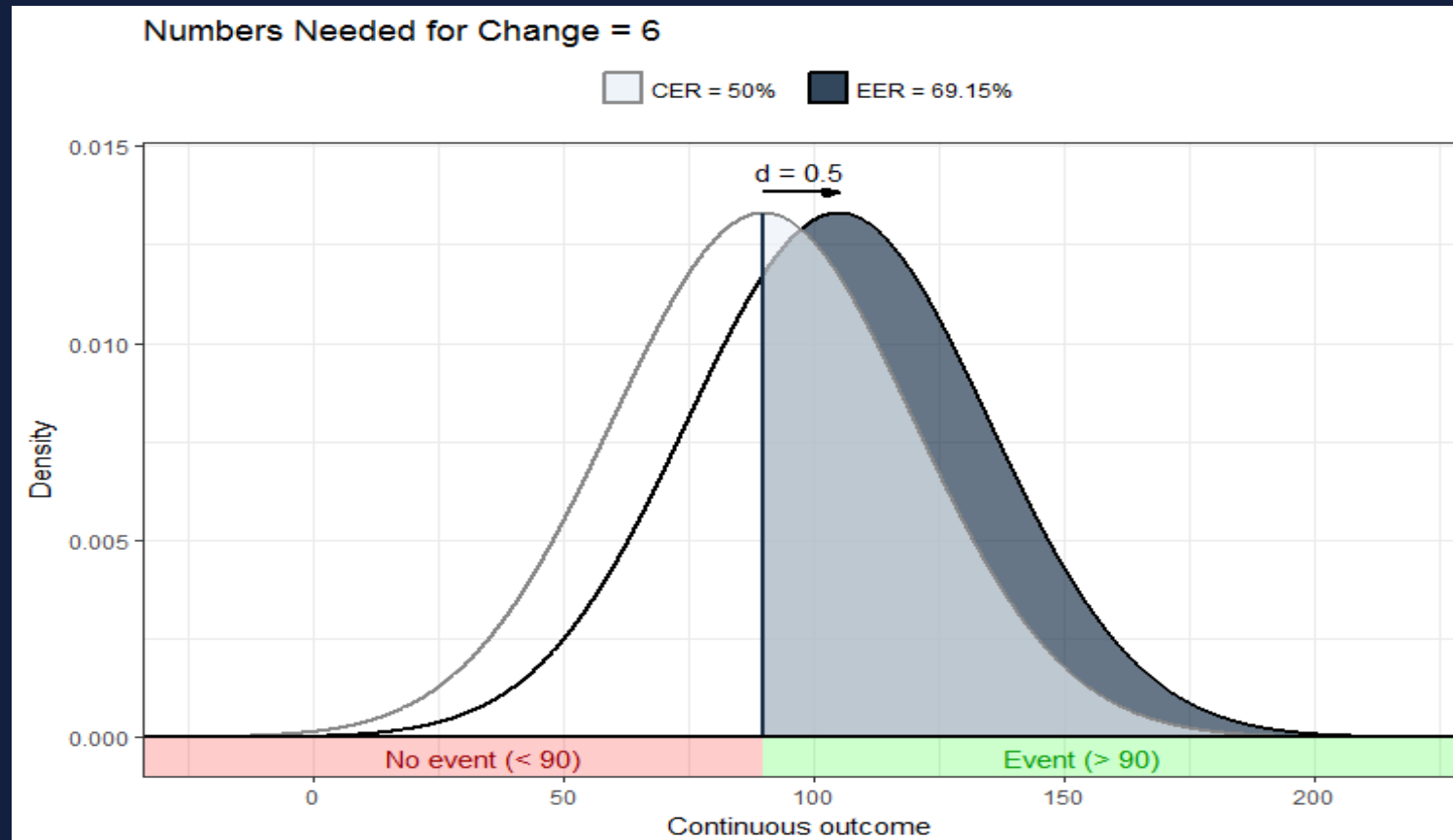
CER = .15
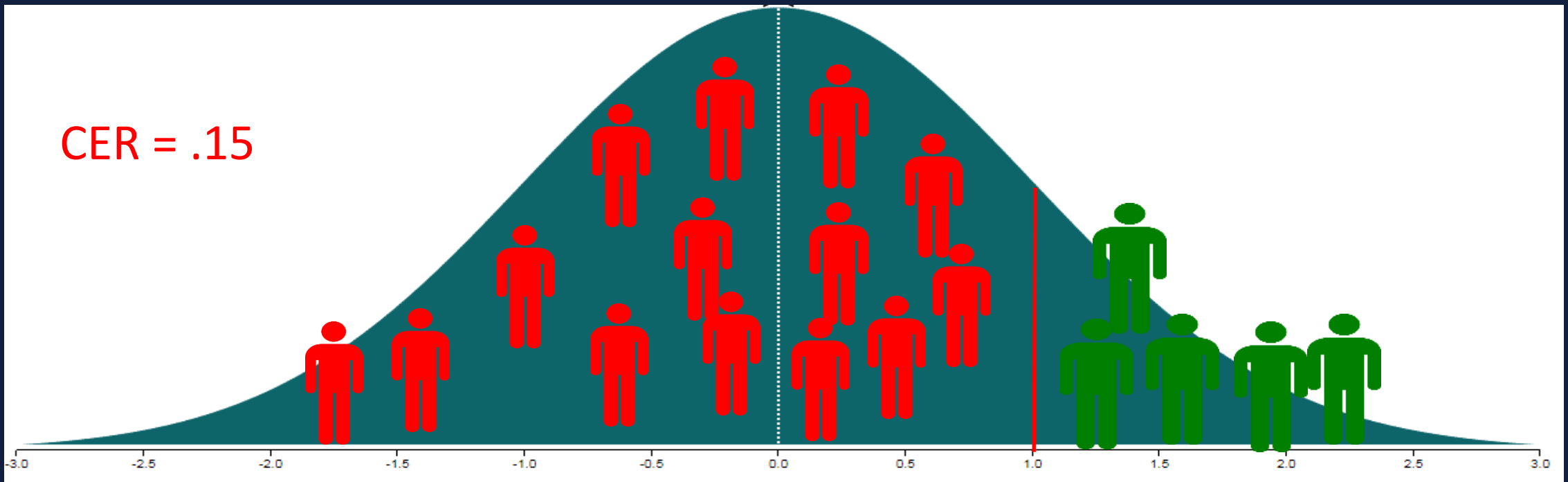
Estimating the control event rate (CER)

- The practical implications of a given Cohen's d depends on the CER
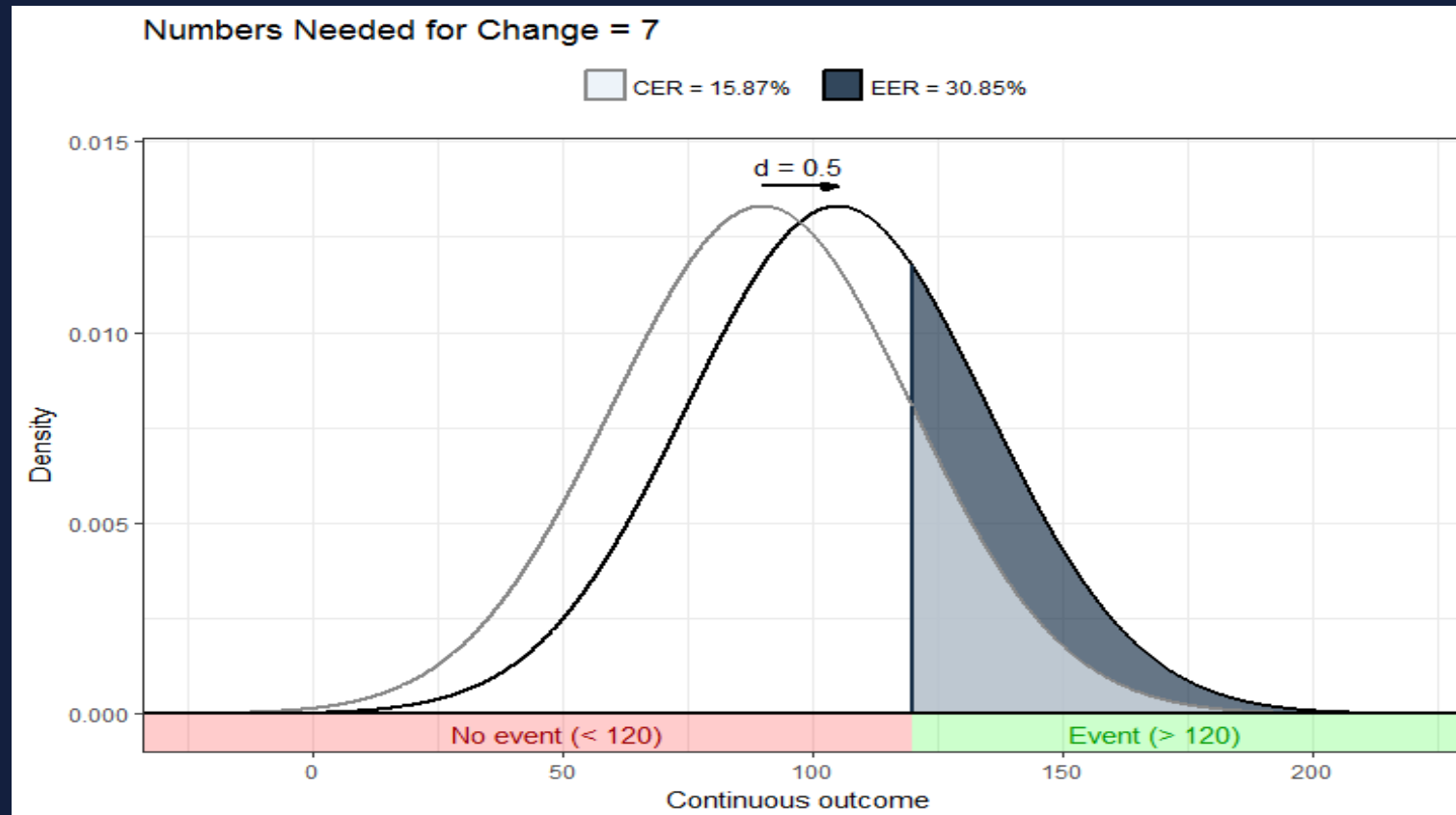
NNC can also be reported for continous outcomes

- Intervention to improve minutes of physical activity per week (**mean = 90, sd = 30)**

- **nnc(CER = .50, d = .5);**

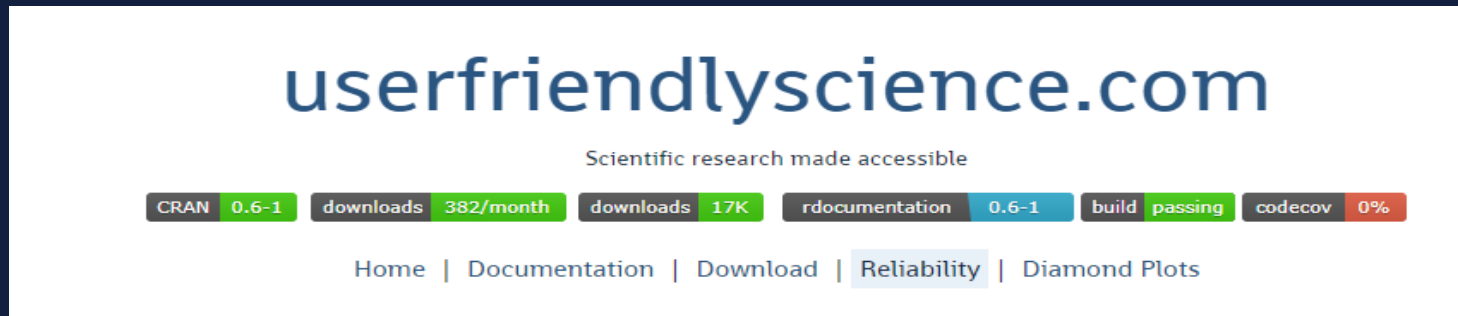# NNC can also be reported for continous outcomes

- Intervention to improve minutes of physical activity per week

- nnc(CER = .15, d = .5);

Conclusions

- Interventions with a given Cohen's d can have **different practical implications**

- Ultimately, cost-benefit ratio determines the practical utility

- NNC can be used as an **additional measure** of effect size, to express practical implications of interventions in a more tangible way

- See http: **stefangruijters.nl** for more info and links to materials

- It's important to analyze effectiveness from different perspectives, let's not just stick with traditional indices but use NNC additionally.

Stefangruijters@gmail.com or Gjalt-jorn@behaviorchange.eu for info, questions or collaboration