

De zin en onzin van sample size berekeningen

Jos Twisk

Sample size berekeningen

- **Nodig voor subsidiegevers**
- **Nodig voor medisch etische commissies**
- **Nodig in verband met 'consort statement'**

Sample size berekening

- Te verwachten effect (klinisch relevant effect)
- Standaard deviatie van de uitkomstvariabele
- Power en significantie

Voor continue variabelen

$$N = \frac{(Z_{(1-\alpha/2)} + Z_{(1-\beta)})^2 \times \sigma^2 \times (1 + \frac{1}{r})}{v^2}$$

N = steekproefgrootte

$Z_{(1-\alpha/2)}$ = $(1 - \alpha/2)$ percentiel van standaard normale verdeling

$Z_{(1-\beta)}$ = $(1 - \beta)$ percentiel van standaard normale verdeling

Σ = standaard deviatie van de uitkomst variabele

r = verhouding tussen de aantallen in de groepen

v = verschil tussen de groepen

Voorbeeld

**Experimenteel onderzoek
naar de werking van een
nieuwe medicatie ter
verlaging van de bloeddruk**

Voorbeeld

- $\alpha = 5\%$
- power = 90%
- SD = 15
- klinisch relevant verschil = 3

$$N_1 = N_2 = 10.51 * (15)^2 * 2 / (3)^2 = 526$$

Voorbeeld

- $\alpha = 5\%$
- power = 80%
- SD = 15
- klinisch relevant verschil = 3

$$N_1 = N_2 = 7.85 * (15)^2 * 2 / (3)^2 = 393$$

Voorbeeld

- $\alpha = 5\%$
- power = 80%
- SD = 15
- klinisch relevant verschil = 5

$$N_1 = N_2 = 7.85 * (15)^2 * 2 / (5)^2 = 141$$

Voorbeeld

- $\alpha = 5\%$
- power = 80%
- SD = 10
- klinisch relevant verschil = 5

$$N_1 = N_2 = 7.85 * (10)^2 * 2 / (5)^2 = 62$$

Conclusie

Door een paar kleine (realistische) aanpassingen van de sample size berekening kan de benodigde sample size sterk variëren!

526 – 62 in iedere groep

Dichotome uitkomsten

Minder mogelijkheden

Standaard deviatie direct afhankelijk van
het klinisch relevante effect

Voor dichotome variabelen

$$N_1 = N_2 = \frac{(Z_{(1-\alpha/2)} + Z_{(1-\beta)})^2 \times \bar{p}(1-\bar{p}) \times 2}{(p_1 - p_2)^2}$$

Voorbeeld

Experimenteel onderzoek naar de werking van een nieuwe interventie voor patiënten met lage rugklachten

Uitkomstvariabele: Herstel

Voorbeeld

- $\alpha = 5\%$
- power = 90%
- $p_0 = 20\%$
- klinisch relevant verschil = 10%

$$N_1 = N_2 = 10.51 * 0.25 * 0.75 * 2 / (0.10)^2 = 394$$

Voorbeeld

- $\alpha = 5\%$
- power = 80%
- $p_0 = 20\%$
- klinisch relevant verschil = 10%

$$N_1 = N_2 = 7.85 * 0.25 * 0.75 * 2 / (0.10)^2 = 294$$

Voorbeeld

- $\alpha = 5\%$
- power = 80%
- $p_0 = 20\%$
- klinisch relevant verschil = 15%

$$N_1 = N_2 = 7.85 * 0.275 * 0.725 * 2 / (0.15)^2 = 139$$

**Hoe gaat het nu in de
praktijk?**

Stel: we doen een mooie sample size berekening en het blijkt dat we 80 patiënten nodig hebben. Nu blijkt dat we maar maximaal 60 patiënten in iedere groep kunnen includeren.

Wat te doen?

Onderzoek niet uitvoeren

Evidence based medicine

Beleid gebaseerd op de resultaten van 1
RCT?

Evidence based medicine

- **Meta analyses**
- **Vergroting sample size**
- **Ook kleine onderzoeken kunnen een wezenlijke bijdrage leveren**

Stel: we doen een mooie sample size berekening en het blijkt dat we 80 patiënten nodig hebben. Nu blijkt dat we zonder problemen 200 patiënten in iedere groep kunnen includeren.

Wat te doen?

Onderzoek uitvoeren bij maar 80 mensen?

**1 uitkomst uit de sample size
berekening:**

Zo groot mogelijk!

- **Sample size berekeningen hechten te veel belang aan significantie en power**
- **Sample size berekeningen zijn te veel afhankelijk van aanames en schattingen**
- **Kleinschalig onderzoek nauwelijks mogelijk**

**Andere overwegingen zouden de grootte
van de onderzoekspopulatie moeten
bepalen**

Overwegingen

- **Logistieke overwegingen**
- **Financiële overwegingen**
- **Ethische overwegingen**



- **Sample size berekeningen zijn heel erg gerelateerd aan de toetsingstheorie**
- **Wat is eigenlijk het belang van de toetsingstheorie?**



- **Moeten we wel uitspraken doen of een bepaald resultaat significant is of niet?**

Significantie is afhankelijk van sample size

Number of obs = 50

Weight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
intervention	-5.722	3.103	-1.844	0.072	-11.804	0.360
_cons	94.238	4.906	19.210	0.000	84.526	103.960

Number of obs = 50

Weight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
intervention	-5.722	3.103	-1.844	0.072	-11.804	0.360
_cons	94.238	4.906	19.210	0.000	84.526	103.960

Number of obs = 60

Weight	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
intervention	-5.722	2.822	-2.027	0.047	-11.253	-0.191
_cons	94.238	4.462	21.120	0.000	85.492	102.984



- **Moeten we wel uitspraken doen of een bepaald resultaat significant is of niet?**

- **..... Nee!**

Waarom nee?

- **Significantie zegt niets over het belang van het gevonden effect**
- **Significantie is heel erg afhankelijk van de grootte van de onderzoekspopulatie**
- **Significantie wordt door bijna iedereen verkeerd geïnterpreteerd!**

Significantie belangrijk

- **Psychologie**
- **Geneeskunde**
- **(Farmaceutische) industrie**
- **Reviewers**

Significantie belangrijk

- **Het is makkelijk**

Conclusie

- **Sample size berekeningen zijn eigenlijk onzinnig.**
- **Statistisch toetsen is (binnen ons vakgebied) eigenlijk onzinnig.**