Medical statistics, Part 1
Faculty of medicine, NTNU 2009

NTNU

Sample size and Power calculations

Geir Jacobsen, ISM

Medical statistics, Part 1

Faculty of medicine, NTNU 2009

NTNU

Sample size and Power calculations

The essential question in any trial/analysis:

"How many patients/persons/observations do I need?"

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Sample size (an example)

”Twenty patients (10 in Arm A and 10 in Arm B) will be included initially as a ”run in phase” of the study for the initial evaluation of feasibility and safety… The median survival in patients who are given (…) is taken to be 6.5 months. To detect an increase of at least 3 months in survival among patients given (…) (ie. to 9.5 months), the trial would need to recruit 50 patients; this with 70% power and a level of significance of 5% (two-sided)

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Sample size and Power calculations

The essential question in any trial/analysis:

”How many patients/persons/observations do I need?”

a. Did we get an answer to our question?
b. Are we satisfied with the answer?

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Sample size (an example)

”Twenty patients (10 in Arm A and 10 in Arm B) will be included initially as a ”run in phase” of the study for the initial evaluation of feasibility and safety… The median survival in patients who are given (…) is taken to be 6.5 months. To detect an increase of at least 3 months in survival among patients given (…) (ie. to 9.5 months), the trial would need to recruit 50 patients; this with 70% power and a level of significance of 5% (two-sided)

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Expect three questions in return (I)

- *How frequent is the condition you are interested in?*
  - ✓ Which relates to your knowledge about the *incidence* and *prevalence* of the disease under study (or any other relevant outcome measure)

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Expect three questions in return (II)

- *What is the size of the difference you would like to observe?*
  - ✓ Which relates to the magnitude of the effect you aim to uncover, - from your *clinical* og *biological* point of view
  - ✓ In other words: What is the minimum difference that is of clinical importance (significance) *to you*?

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Expect three questions in return (III)

- *How sure do you want to be (ie. once you draw your conclusion)?*
  - ✓ That your observed difference is a ”true” one
  - ✓ That the difference you look for, is not overlooked

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Sample size (an example)

”Twenty patients (10 in Arm A and 10 in Arm B) will be included initially as a ”run in phase” of the study for the initial evaluation of feasibility and safety… The median survival in patients who are given (…) is taken to be 6.5 months. To detect an increase of at least 3 months in survival among patients given (…) (ie. to 9.5 months), the trial would need to recruit 50 patients; this with 70% power and a level of significance of 5% (two-sided)

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

During this lecture we shall touch upon

- One sample test
- Two sample test

- One sided test
- Two sided test

- Continous variables
- Proportions

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Two quotations worth noticing:

”An hypothesis is a statement whose incorrect rejection one tries to avoid”

#

- ”Hypotheses can only be tested, - but never proven”

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Testing of hypothesis requires that you consider:

The *Null vis á vis* The *Alternative* hypothesis

#

H_0: $\mu = \mu_0$ vs. H_1: $\mu = \mu_1$

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

For instance: A randomised controlled trial

Ie. A comparison between Trt. A vs Trt. B

H_0: $\mu_A = \mu_B$ vs. H_1: $\mu_A \neq \mu_B$

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Given that H_0 = True

		Our conclusion	
		Yes = True	No = False
"The ABSOLUTE Truth"	Yes = True	ok	Type I error (α)
	No = False	Type II error (β)	ok

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

NTNU

Significance (statistical):

The probability to reject H_0 when H_0 is true. Expressed as our choice of the level of α

(= ”Our willingness to commit a Type I error”)

Power:

The ability of our study to reject H_0 når H_0 false. Expressed as our choice of the level of β (or really $1 - \beta$)

(= ”Our willingness to commit a Type II error”)

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

NTNU

Power calculations - Practical

Suppose that a case-control study is planned for assessing the relationship between smoking in pregnancy and low birth weight in the offspring. Cases would be women giving birth to babies weighing 2500 grams or below, and controls would be women giving birth to babies over 2500 grams. Because only a small minority of the population falls into the case group, the overall prevalence of smoking in the general population of pregnant women serves quite well as an estimate of p_0, which in a case-control study denotes the proportion of controls who have the exposure (in this case 25%). Suppose that an odds ratio of 1,8 is regarded as important to detect, that 175 cases are available for study, and that a case-to-control ratio *(r)* is planned.

What power has the study to detect an odds ratio = 1,8?

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Table 10–11. Definitions of symbols used in equations for calculating power and required sample size

Symbol	Definition
d^*	Nonnull value of the difference in proportions or means (i.e., the magnitude of difference one wishes to detect)
n	In a cohort or cross-sectional study, the number of exposed individuals studied; in a case-control study, the number of cases
r	In a cohort or cross-sectional study, the ratio of the number of unexposed individuals studied to the number of exposed individuals studied; in a case-control study, the ratio of the number of controls studied to the number of cases studied
σ	Standard deviation in the population for a continuously distributed variable
p_1	In a cohort study (or a cross-sectional study), the proportion of exposed individuals who develop (or have) the disease; in a case-control study, the proportion of cases who are exposed
p_0	In a cohort study (or a cross-sectional study), the proportion of unexposed individuals who develop (or have) the disease; in a case-control study, the proportion of controls who are exposed
$\bar{p}$	$\frac{p_1 + rp_0}{1 + r}$ = weighted average of p_1 and p_0

NTNU

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Table 10–12. Formulas for use in calculating the power of a study to detect an association

Z_β for difference in means: $\frac{d^*}{\sigma}\sqrt{\frac{nr}{r+1}} - Z_{\alpha/2}$

Z_β for difference in proportions: $\left[\frac{n(d^*)^2 r}{(r+1)\bar{p}(1-\bar{p})}\right]^{1/2} - Z_{\alpha/2}$

Value of p_1 in terms of p_0 and a specified odds ratio (OR): $p_1 = \frac{p_0 OR}{1 + p_0(OR - 1)}$

Value of p_1 in terms of p_0 and a specified risk ratio (RR) (for use in cohort or cross-sectional studies only): $p_1 = p_0 RR$

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

NTNU

Table 10–15. Formulas for use in calculations of required sample size

Difference in means:

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2\, \sigma^2\, (r + 1)}{(d^*)^2 r}$$

Difference in proportions:

$$n = \frac{(Z_{\alpha/2} + Z_{\beta})^2\, \bar{p}(1 - \bar{p})(r + 1)}{(d^*)^2 r}$$

NTNU

Table 10–14. Conversion of Z_β to the percentage corresponding to the power for detecting an association

Z_β	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−1.3	9.7	9.5	9.3	9.2	9.0	8.9	8.7	8.5	8.4	8.2
−1.2	11.5	11.3	11.1	10.9	10.7	10.6	10.4	10.2	10.0	9.9
−1.1	13.6	13.3	13.1	12.9	12.7	12.5	12.3	12.1	11.9	11.7
−1.0	15.9	15.6	15.4	15.2	14.9	14.7	14.5	14.2	14.0	13.8
−0.9	18.4	18.1	17.9	17.6	17.4	17.1	16.9	16.6	16.4	16.1
−0.8	21.2	20.9	20.6	20.3	20.0	19.8	19.5	19.2	18.9	18.7
−0.7	24.2	23.9	23.6	23.3	23.0	22.7	22.4	22.1	21.8	21.5
−0.6	27.4	27.1	26.8	26.4	26.1	25.8	25.5	25.1	24.8	24.5
−0.5	30.9	30.5	30.2	29.8	29.5	29.1	28.8	28.4	28.1	27.8
−0.4	34.5	34.1	33.7	33.4	33.0	32.6	32.3	31.9	31.6	31.2
−0.3	38.2	37.8	37.4	37.1	36.7	36.3	35.9	35.6	35.2	34.8
−0.2	42.1	41.7	41.3	40.9	40.5	40.1	39.7	39.4	39.0	38.6
−0.1	46.0	45.6	45.2	44.8	44.4	44.0	43.6	43.3	42.9	42.5
−0.0	50.0	49.6	49.2	48.8	48.4	48.0	47.6	47.2	46.8	46.4
0.0	50.0	50.4	50.8	51.2	51.6	52.0	52.4	52.8	53.2	53.6
0.1	54.0	54.4	54.8	55.2	55.6	56.0	56.4	56.7	57.1	57.5
0.2	57.9	58.3	58.7	59.1	59.5	59.9	60.3	60.6	61.0	61.4
0.3	61.8	62.2	62.6	62.9	63.3	63.7	64.1	64.4	64.8	65.2
0.4	65.5	65.9	66.3	66.6	67.0	67.4	67.7	68.1	68.4	68.8
0.5	69.1	69.5	69.8	70.2	70.5	70.9	71.2	71.6	71.9	72.2
0.6	72.6	72.9	73.2	73.6	73.9	74.2	74.5	74.9	75.2	75.5
0.7	75.8	76.1	76.4	76.7	77.0	77.3	77.6	77.9	78.2	78.5
0.8	78.8	79.1	79.4	79.7	80.0	80.2	80.5	80.8	81.1	81.3
0.9	81.6	81.9	82.1	82.4	82.6	82.9	83.1	83.4	83.6	83.9
1.0	84.1	84.4	84.6	84.8	85.1	85.3	85.5	85.8	86.0	86.2
1.1	86.4	86.7	86.9	87.1	87.3	87.5	87.7	87.9	88.1	88.3
1.2	88.5	88.7	88.9	89.1	89.3	89.4	89.6	89.8	90.0	90.1
1.3	90.3	90.5	90.7	90.8	91.0	91.1	91.3	91.5	91.6	91.8
1.4	91.9	92.1	92.2	92.4	92.5	92.6	92.8	92.9	93.1	93.2
1.5	93.3	93.4	93.6	93.7	93.8	93.9	94.1	94.2	94.3	94.4
1.6	94.5	94.6	94.7	94.8	94.9	95.1	95.2	95.3	95.4	95.4
1.7	95.5	95.6	95.7	95.8	95.9	96.0	96.1	96.2	96.2	96.3
1.8	96.4	96.5	96.6	96.6	96.7	96.8	96.9	96.9	97.0	97.1
1.9	97.1	97.2	97.3	97.3	97.4	97.4	97.5	97.6	97.6	97.7
2.0	97.7	97.8	97.8	97.9	97.9	98.0	98.0	98.1	98.1	98.2
2.1	98.2	98.3	98.3	98.3	98.4	98.4	98.5	98.5	98.5	98.6
2.2	98.6	98.6	98.7	98.7	98.7	98.8	98.8	98.8	98.9	98.9
2.3	98.9	99.0	99.0	99.0	99.0	99.1	99.1	99.1	99.1	99.2
2.4	99.2	99.2	99.2	99.2	99.3	99.3	99.3	99.3	99.3	99.4

Note: For values of Z_β less than −1.39, the power is less than 8.2%; for values of Z_β greater than 2.49, the power is greater than 99.4%.

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Power calculations - Practical

Suppose that a case-control study is planned for assessing the relationship between smoking in pregnancy and low birth weight in the offspring. Cases would be women giving birth to babies weighing 2500 grams or below, and controls would be women giving birth to babies over 2500 grams. Because only a small minority of the population falls into the case group, the overall prevalence of smoking in the general population of pregnant women serves quite well as an estimate of p_0, which in a case-control study denotes the proportion of controls who have the exposure (in this case 25%). Suppose that an odds ratio of 1,8 is regarded as important to detect, that 175 cases are available for study, and that a case-to-control ratio *(r)* is planned.

What power has the study to detect an odds ratio = 1,8?

Power Calculation: Practical

p_0: Proportion of exposed controls
OR of importance to identify = 1.8
Number ♀ available for study: n = 175
Control - to - case ratio = 2

$$p_1 = \frac{(0.25) \times (1.8)}{[1+(0.25) \times (1.8 - 1)]} = \underline{0.375}$$

$$d^* = p_1 - p_0 = 0.375 - 0.250 = \underline{0.125}$$

$$\bar{p} = \frac{[(0.375) + (2) \times (0.25)]}{(1 + 2)} = \underline{0.292}$$

$$Z_\vartheta = \left[\frac{(175) \times (0.125)^2 \times (2)}{(2+1) \times 0.292 \times 0.708}\right]^{1/2} - 1.96 = \underline{1.01}$$

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Factors affecting the power ($1 - \beta$)

- If the significance level is made smaller (α decreases), z_α decreases and hence the power decreases
- If the alternative mean is shifted further away from the null mean ($|\mu_1 - \mu_0|$ increases), then the power increases
- If the standard deviation of the distribution of individual observations increases (σ increases), then the power decreases
- If the sample size increases (n increases), then the power increases

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

Factors affecting the sample size (n)

- The sample size increases as σ^2
- The sample size increasesas the significance level is made smaller (α decreases)
- The sample size increases as the required power increases ($1 - \beta$ increases)
- The sample size decreases as the absolute value of the distance between the null and alternative means ($|\mu_1 - \mu_0|$) increases

Medical statistics, Part 1
Faculty of medicine, NTNU 2009

What you need to know and decide

1. α Level and $Z_{\alpha/2}$
2. β level and Z_{β}
3. δ level = difference (in prevalence, incidence, or any outcome variable) between the groups you want to observe

Then – and only then – can you calculate the number needed in your study (ie. *n* in each arm)