Determinazione della numerosità campionaria

Prof.ssa Marta Di Nicola
marta.dinicola@unich.it

Contents lists available at ScienceDirect

International Journal of Antimicrobial Agents

journal homepage: www.elsevier.com/locate/ijantimicag

Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial ☆

2.10. Statistics

Assuming a 50% efficacy of hydroxychloroquine in reducing the viral load at day 7, an 85% power, a type I error rate of 5% and 10% loss to follow-up, we calculated that a total of 48 COVID-19 patients (ie, 24 cases in the hydroxychloroquine group and 24 in the control group) would be required for the analysis (Fleiss with CC). Statistical differences were evaluated by Pearson's chi-square or Fisher's exact tests as categorical variables, as appropriate. Means of quantitative data were compared using Student's t-test. Analyses were performed in Stata version 14.2.

La dimensione del campione

Stima (su considerazioni statistiche)

Giustificazione (su considerazioni economiche e biologiche)

Aggiustamento (in accordo a potenziali dropouts o effetti di covariate)

Perché stimare a priori la dimensione del campione da studiare?

Per assicurare che il numero di soggetti/pazienti coinvolti nello studio clinico permetta di rispondere adeguatamente al quesito di interesse, considerando che:

1. Uno studio di dimensioni limitate avrà un elevata probabilità di non riuscire a riconoscere un trattamento promettente;

2. Uno studio di dimensioni eccessive è economicamente oneroso e rischia di sottoporre un eccessivo numero di soggetti/pazienti ad un trattamento non efficace.

Patients I need = sample size

Patients I can get = power

Come si calcola la sample size?

Il calcolo della sample size varia a seconda di:

- Obiettivo dello studio (superiorità, inferiorità, etc);
- Disegno dello studio (gruppi paralleli, gruppi indipendenti, cross-over, etc), numero dei gruppi di trattamento da confrontare;
- End-point primario (percentuale di successo, probabilità di sopravvivenza libera da evento, livello medio di un parametro quantitativo, etc);
- Parametri clinici (entità dell'effetto, percentuale di drop-out, variabilità del fenomeno);
- Parametri statistici (tipo di test d'ipotesi, livello di significatività, potenza dello studio).

Power Analysis

Test for	Null Hypothesis	Alternative Hypothesis
Equality	$H_0 : \mu_T - \mu_S = 0$	$H_a : \mu_T - \mu_S \neq 0$
Non-inferiority	$H_0 : \mu_T - \mu_S \geq \delta$	$H_0 : \mu_T - \mu_S < \delta$
Superiority	$H_0 : \mu_T - \mu_S \leq \delta$	$H_0 : \mu_T - \mu_S > \delta$
Equivalence	$H_0 : \lvert\mu_T - \mu_S\rvert \geq \delta$	$H_0 : \lvert\mu_T - \mu_S\rvert < \delta$

Quali sono i fattori che entrano nel calcolo della sample size?

- Il potere del test (1-β) che si desidera ottenere (valori di riferimento standard sono l'80% e il 90%). Probabilità di rigettare l'ipotesi nulla quando esiste una reale differenza o associazione.

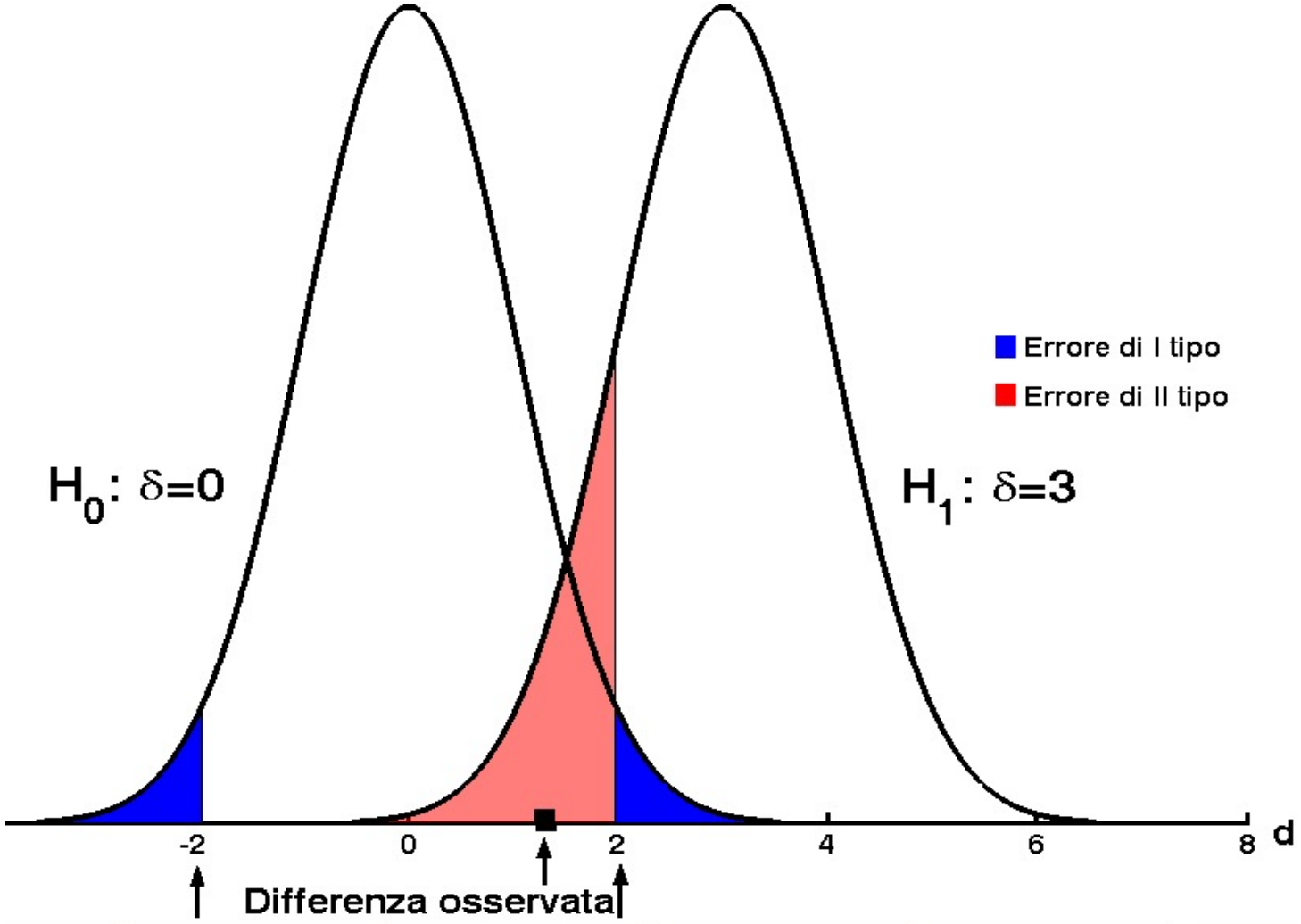
- La differenza minima tra i trattamenti che sia "clinicamente rilevante". Può essere espresso in generale come il grado di beneficio che il nuovo trattamento dovrebbe fornire rispetto a quello vecchio perché valga la pena di utilizzarlo. Negli studi di *superiorità* è la minima differenza clinicamente rilevante; negli studi di *non-inferiorità* massimo svantaggio clinicamente tollerabile

- Il livello di significatività (α), cioè la probabilità di ottenere una differenza "statisticamente significativa" quando di fatto differenza non c'è.

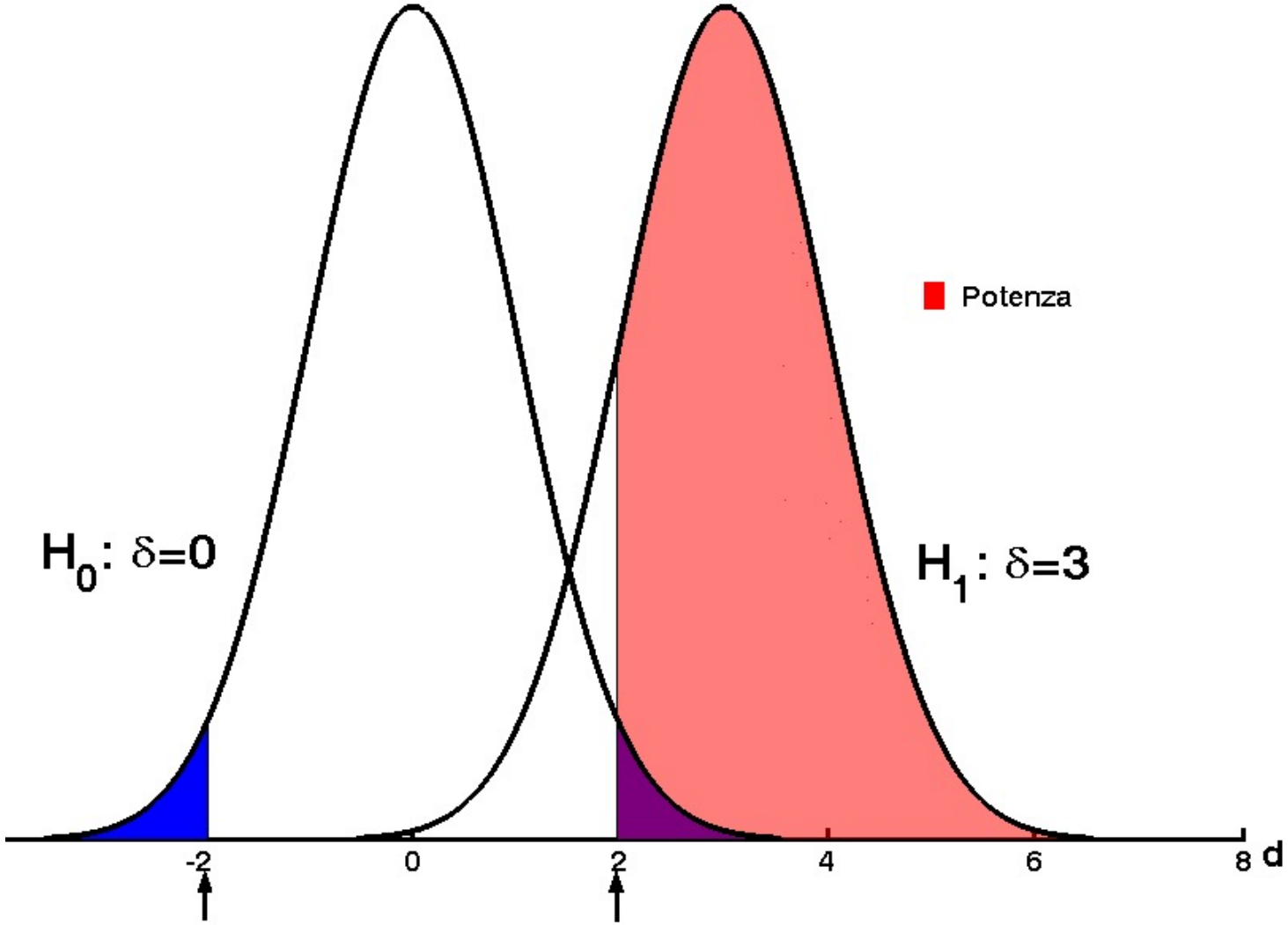
Errori di I e II tipo

Conclusione Test	Medie della Popolazione: Uguali (H_0)	Medie della Popolazione: Differenti (H_1)
Le medie *non sono* differenti	VN	FN II (β)
Le medie *sono* differenti	I (α) FP	potenza (1-β) VP

Errori di I e II tipo



Potenza del test (1- β)



Effetto della precisione sulla sample size

Per i valori più comuni di α e 1-β

	1-β		
α	0,80	0,90	0,95
0,10	6,2	8,6	10,8
0,05	7,9	10,5	13,0
0,01	11,7	13,0	17,8

$$(Z_{\beta} + Z_{\alpha/2})^2$$

Variabili quantitative

$$n = \frac{2\sigma^2}{(\mu_1 - \mu_2)^2}(Z_\beta + Z_{\alpha/2})^2$$

! A posteriori è bene comparare la varianza dei dati osservati con quella utilizzata nella stima della sample size!

Variabili quantitative

Esempio: Il farmaco di riferimento riduce la pressione sistolica di 25 mmHg, il nuovo farmaco per essere competitivo dovrebbe ridurre la pressione sistolica di almeno 30 mmHg (ovvero 5 mmHg in più). La deviazione standard della riduzione della pressione viene stimata in 10 mmHg da studi precedenti. Si adotta un alfa del 5% e una potenza del 90%, pertanto f(α,β)=10,5

```
. sampsi 25 30, sd1(10) sd2(10) power(.9)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                  and m2 is the mean in population 2
Assumptions:

         alpha =    0.0500  (two-sided)
         power =    0.9000
            m1 =        25
            m2 =        30
           sd1 =        10
           sd2 =        10
         n2/n1 =      1.00

Estimated required sample sizes:

            n1 =        85
            n2 =        85
```

Occorrono almeno 85 soggetti per gruppo.

Variabili qualitative

$$n = \frac{p_1(100 - p_1) + p_2(100 - p_2)}{(p_1 - p_2)^2}(Z_\beta + Z_{\alpha/2})^2$$

Variabili qualitative

Esempio: Nei pazienti affetti da tumore X in stadio avanzato, la sopravvivenza a 5 anni è del 30% con il trattamento standard. Dati preliminari suggeriscono che nei pazienti sottoposti ad un nuovo trattamento la sopravvivenza salga al 40%. Si adotta un alfa del 5% e una potenza dell' 80%, pertanto f(α,β)=7,9

```
. sampsi 0.30 0.40, power(.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                  and p2 is the proportion in population 2
Assumptions:

         alpha =    0.0500  (two-sided)
         power =    0.8000
            p1 =    0.3000
            p2 =    0.4000
         n2/n1 =    1.00

Estimated required sample sizes:

            n1 =       376
            n2 =       376
```

Occorrono almeno 376 soggetti per gruppo.

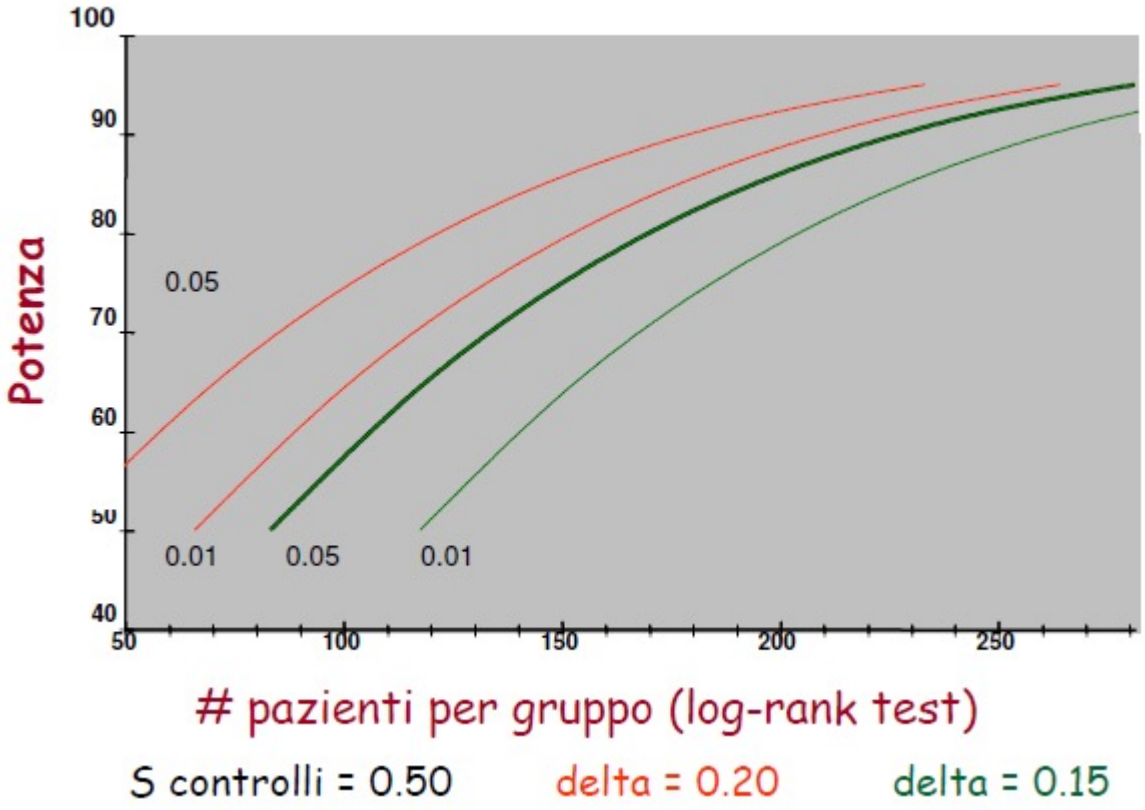
Variabili quantitative-Paired data

$$n = \frac{\sigma_D^{\,2}(Z_\beta + Z_{\alpha/2})^2}{\overline{D}^2}$$

```
. sampsi 498 485, sd1(20.2) sd2(19.5) method(change) pre(1) post(3) r1(.7)

Estimated sample size for two samples with repeated measures
Assumptions:
                                        alpha =    0.0500  (two-sided)
                                        power =    0.9000
                                           m1 =       498
                                           m2 =       485
                                          sd1 =      20.2
                                          sd2 =      19.5
                                        n2/n1 =      1.00
                  number of follow-up measurements =         3
 correlation between follow-up measurements =     0.700
                   number of baseline measurements =         1
   correlation between baseline & follow-up =     0.700

Method: CHANGE
 relative efficiency =     2.500
    adjustment to sd =     0.632
        adjusted sd1 =    12.776
        adjusted sd2 =    12.333

 Estimated required sample sizes:
                    n1 =        20
                    n2 =        20
```

Factors Affecting sample size

- I livelli di errore
- La misura dell'effetto (delta)
- La deviazione standard del fenomeno



Design	Hypothesis	Hypotheses and Sample Size Rules		
		H_0	H_a	Basic Rule
One-sample	Equality	$\mu - \mu_0 = 0$	$\mu - \mu_0 \neq 0$	$n = \frac{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 \sigma^2}{(\mu - \mu_0)^2}$
	Superiority	$\mu - \mu_0 \leq \delta$	$\mu - \mu_0 > \delta$	$n = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{(\mu - \mu_0 - \delta)^2}$
	Equivalence	$\lvert\mu - \mu_0\rvert \geq \delta$	$\lvert\mu - \mu_0\rvert < \delta$	$n = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{(\lvert\mu - \mu_0\rvert - \delta)^2}$
Two-sample Parallel	Equality	$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 \neq 0$	$n_i = \frac{2\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 \sigma^2}{(\mu_1 - \mu_2)^2}$
	Non-inferiority	$\mu_1 - \mu_2 \geq \delta$	$\mu_1 - \mu_2 < \delta$	$n_i = \frac{2\left(z_\alpha + z_\beta\right)^2 \sigma^2}{(\mu_1 - \mu_2 - \delta)^2}$
	Superiority	$\mu_1 - \mu_2 \leq \delta$	$\mu_1 - \mu_2 > \delta$	$n_i = \frac{2\left(z_\alpha + z_\beta\right)^2 \sigma^2}{(\mu_1 - \mu_2 - \delta)^2}$
	Equivalence	$\lvert\mu_1 - \mu_2\rvert \geq \delta$	$\lvert\mu_1 - \mu_2\rvert < \delta$	$n_i = \frac{2\left(z_\alpha + z_\beta\right)^2 \sigma^2}{(\lvert\mu_1 - \mu_2\rvert - \delta)^2}$
Two-sample Crossover	Equality	$\mu_1 - \mu_2 = 0$	$\mu_1 - \mu_2 \neq 0$	$n_i = \frac{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 \sigma^2}{2(\mu_1 - \mu_2)^2}$
	Non-inferiority	$\mu_1 - \mu_2 \geq \delta$	$\mu_1 - \mu_2 < \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{2(\mu_1 - \mu_2 - \delta)^2}$
	Superiority	$\mu_1 - \mu_2 \leq \delta$	$\mu_1 - \mu_2 > \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{2(\mu_1 - \mu_2 - \delta)^2}$
	Equivalence	$\lvert\mu_1 - \mu_2\rvert \geq \delta$	$\lvert\mu_1 - \mu_2\rvert < \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma^2}{2(\lvert\mu_1 - \mu_2\rvert - \delta)^2}$

Design	Hypothesis	Hypotheses and Sample Size rules	
		H_0	Basic Rule
One-sample	Equality	$\pi - \pi_0 = 0$	$n = \frac{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 \pi(1-\pi)}{(\pi-\pi_0)^2}$
	Superiority	$\pi - \pi_0 \leq \delta$	$n = \frac{\left(z_\alpha + z_\beta\right)^2 \pi(1-\pi)}{(\pi-\pi_0-\delta)^2}$
	Equivalence	$\lvert\pi - \pi_0\rvert \geq \delta$	$n = \frac{\left(z_\alpha + z_\beta\right)^2 \pi(1-\pi)}{(\lvert\pi-\pi_0\rvert-\delta)^2}$
Two-sample Parallel	Equality	$\pi_1 - \pi_2 = 0$	$n_i = \frac{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 (\pi_1(1-\pi_2)+\pi_2(1-\pi_2))}{(\pi_1-\pi_2)^2}$
	Non-inferiority	$\pi_1 - \pi_2 \geq \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 (\pi_1(1-\pi_2)+\pi_2(1-\pi_2))}{(\pi_1-\pi_2-\delta)^2}$
	Superiority	$\pi_1 - \pi_2 \leq \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 (\pi_1(1-\pi_2)+\pi_2(1-\pi_2))}{(\pi_1-\pi_2-\delta)^2}$
	Equivalence	$\lvert\pi_1 - \pi_2\rvert \geq \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 (\pi_1(1-\pi_2)+\pi_2(1-\pi_2))}{(\lvert\pi_1-\pi_2\rvert-\delta)^2}$
Two-sample Crossover	Equality	$\pi_1 - \pi_2 = 0$	$n_i = \frac{\left(z_{\frac{\alpha}{2}} + z_\beta\right)^2 \sigma_d^2}{2(\pi_1-\pi_2)^2}$
	Non-inferiority	$\pi_1 - \pi_2 \geq \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma_d^2}{2(\pi_1-\pi_2-\delta)^2}$
	Superiority	$\pi_1 - \pi_2 \leq \delta$	$n_i = \frac{\left(z_\alpha + z_\beta\right)^2 \sigma_d^2}{2(\pi_1-\pi_2-\delta)^2}$
	Equivalence	$\lvert\pi_1 - \pi_2\rvert \geq \delta$	$n_i = \frac{\left(z_\alpha + z_{\beta/2}\right)^2 \sigma_d^2}{2(\lvert\pi_1-\pi_2\rvert-\delta)^2}$

Sample size adjustments

- Separate sample size calculation should be done for each important outcome & then use the max. estimate
- When two variables are correlated with a factor *p*, then sample size can be reduced by a factor of $1-p^2$
- Another option is to use Bonferroni correction for multiple outcomes

Sample size adjustments

- Allowing for response rates & other losses to the sample
  - The expected response rate
  - Loss to f/u
  - Lack of compliance
  - Other losses

$n_{new}=n/(1-L)$ when L is the loss to f/u rate

Sample size adjustments

Adjustment for unequal group size

- Assuming $n_1/n_2=k$
- Calculate n assuming equal
- Then

$n_2=0.5n(1+1/k)$

$n_1=0.5n(1+k)$

Reporting Sample Size Calculations

- Clear statement of the primary objective
- The desired level of significance
- The desired power
- The statistics methods that will be used for analysis
- Whether the test would be one or two-tailed
- The smallest difference
  - Smallest clinically important difference
  - The difference that investigators think is worth detecting
  - The difference that the investigators think is likely to be detected

Reporting Sample Size Calculations

- Justification for prior estimates used in calculations
- Clear statements about the assumptions made about the distribution or variability of the outcomes
- Clear statement about the scheduled duration of the study
- Statement about how the sample size was adjusted
- The software or formulae that was used

Strategies For Maximizing Power & Minimizing the Sample Size

- Use common outcomes (the power is driven more by the number of events than the total sample size)
- Use paired design (such as cross-over trial)
- Use continuous variables
- Choose the timing of the assessments of primary outcomes to be when the difference is most likely to be optimal

Online free software available

http://www.stat.ucla.edu/~jbond/HTMLPOWER/index.html
http://www.health.ucalgary.ca/~rollin/stats/ssize/
http://www.stat.uiowa.edu/%7Erlenth/Power/index.html
http://www.dssresearch.com/SampleSize/
http://www.stat.ucla.edu/calculators/powercalc/
http://hedwig.mgh.harvard.edu/sample_size/size.html
http://www.bobwheeler.com/stat/SSize/ssize.html
http://www.math.yorku.ca/SCS/Online/power/
http://www.surveysystem.com/sscalc.htm
http://www.researchinfo.com/docs/calculators/samplesize.cfm

Sample Sizes for Time-to-Event Studies

Outcome Measure	Examples of Associated Time-to-Event Measure
Death	Survival Time
Response (Tumor shrinkage)	Duration of response
Recurrence of disease	Time to recurrence
Relief of symptoms	Time to relief of symptoms/without symptoms
Quality of life 'scores'	Time to improvement/deterioration in scores
Toxicity	Time with/without toxicity

M_t = mean survival time in Treatment Group

M_c = mean survival time in Control Group

$$n = \frac{2\left(Z_{\alpha/2} + Z_{\beta}\right)^2}{\left(\ln(M_t/M_c)\right)^2}.$$

- Most software require the use of event-free rates (survival) and not event rates (death), because the log rank test is based on event-free rates
- Beware if your software is giving you total number of subjects or events.