The precision of the estimation of a population parameter

depend on

- the estimator
  - ex. mean is more precise then median
- the sample size
  - law of big numbers – bigger sample size – bigger precision
- variation of the data
  - less variation – bigger precision
- needed precision
  - bigger precision – wider interval

Example

- We want to estimate the frequency of parodontosis in population with age higher than 60 years old.
- In a study with 1000 participants, 300 had paradontosis

$$f = \frac{300}{1000} = 0.30 = 30\%$$

95% confidence intervals for the frequency of mouth cancer

$$[f - 1.96\sqrt{\frac{f(1-f)}{n}}; f + 1.96\sqrt{\frac{f(1-f)}{n}}]$$

$$\left[0.3 - 1.96\sqrt{\frac{0.3(1-0.3)}{1000}}; 0.3 + 1.96\sqrt{\frac{0.3(1-0.3)}{1000}}\right]$$

$$[0.3 - 0.028; 0.3 + 0.028]$$

[0.27; 0.33] = [27%; 33%]

Interpretation: with a 5% error we estimate that the population frequency of paradontosis

in the population over 60 years is between 27% and 33%

If we select a small sample size?

- We want to estimate the frequency of parodontosis in population with age higher than 60 years old.
- In a study with 100 participants, 30 had paradontosis

$$f = \frac{30}{100} = 0.30\text{=}30\%$$

$$f = \frac{30}{100} = 0.3$$

$$[0.3 - 1.96\sqrt{\frac{0.3(1-0.3)}{100}}; 0.3 + 1.96\sqrt{\frac{0.3(1-0.3)}{100}}]$$

$$[0.3 - 0.09; 0.3 + 0.09]$$

[0.21; 0.39] – interval for n=100

frequency of paradontosis between 21% și 39% with a 95% probability

[0.27; 0.33] – interval for n=10.000

frequency of paradontosis between 27% și 33% with a 95% probability

Answer: smaller the sample → wider the interval

an increasing n at the denominator, an opposite effect

If the sample increase → it increases the measurement precision by decreasing the interval required for estimation

If the probability decrease 95% → 80%

- We want to estimate the frequency of parodontosis in population with age higher than 60 years old.
- In a study with 100 participants, 30 had paradontosis

For 80% confidence, Z critic =1.29

$$f = \frac{30}{100} = 0.3$$

$$[0.3 - 1.29\sqrt{\frac{0.3(1-0.3)}{100}}; 0.3 + 1.29\sqrt{\frac{0.3(1-0.3)}{100}}]$$

$$[0.3 - 0.059; 0.3 + 0.059]$$

[0.24; 0.36] – interval for 80%

frequency of parodontosis between 24% și 36% with a 95% probability

[0.21; 0.39] – interval for 95%

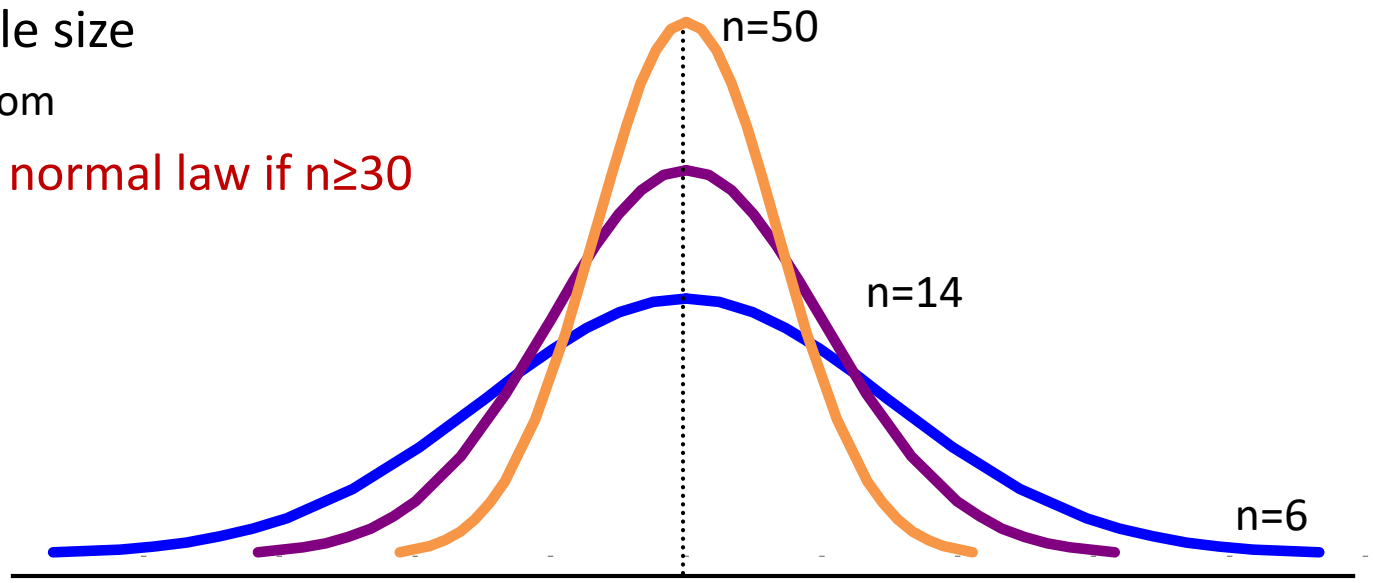
frequency of parodontosis between 21% și 39% with a 95% probability

Answer: the smaller the probability – the narrower the interval

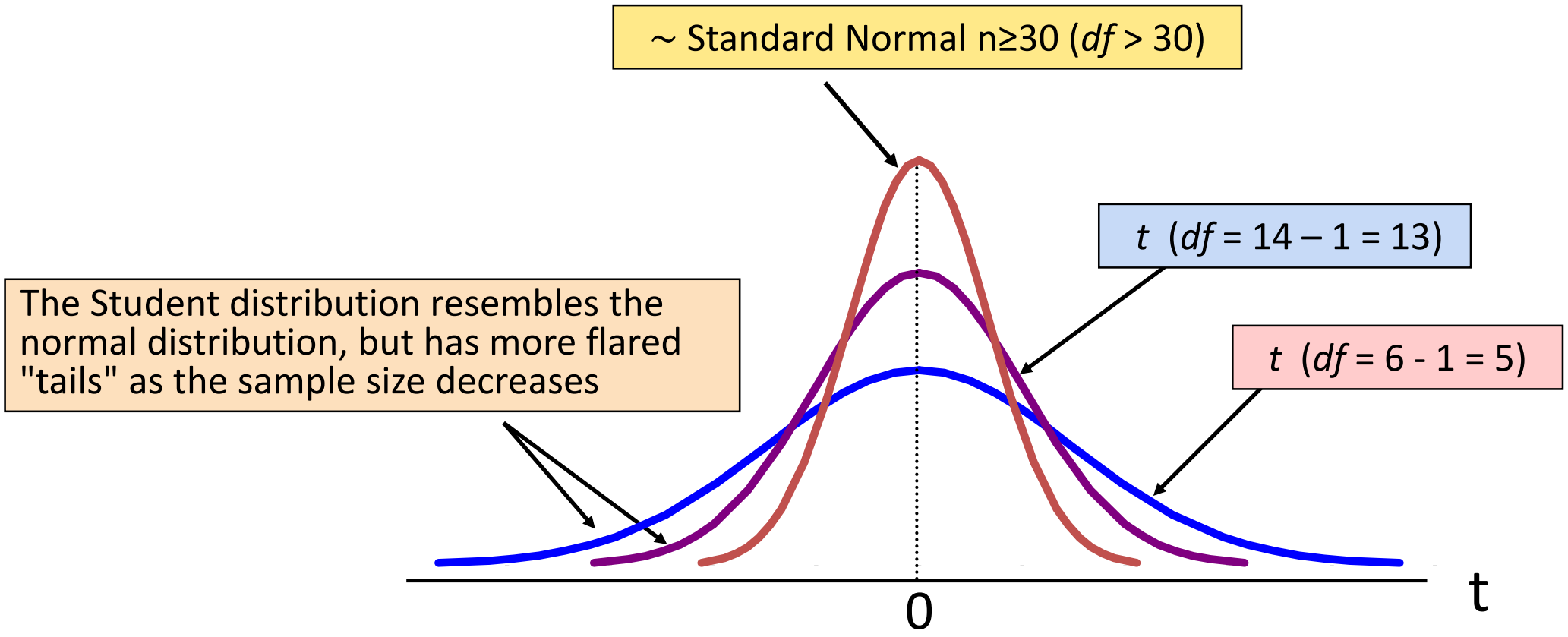
Z at numerator

Increasing precision → interval increase

- The sampling distribution of the mean is normally distributed when the population standard deviation σ is known
- Often σ is not known
  - it is estimated with the standard deviation obtained on the sample s
    - in this case the sampling distribution no longer follows the normal law
    - the sampling distribution follows Student's law – very similar
    - Student's law varies with sample size
      - depends on the degrees of freedom
      - Student's law is close to the normal law if $n \geq 30$



If n<30: we estimate with Student distribution



Student's law varies with sample size

depends on the degrees of freedom

Degree of freedom = df

The number of components that are "free" to vary in a data set

- df = n – 1

Ex: 5 measurements, the average = 6

- we know the first four values: 8, 9, 10, 11
- the fifth can be calculated

--> the data set has only 4 degrees of freedom

1-α% confidence interval for the average μ in the case of small samples n<30 with σ unknown

- using Student t distribution

$$\left[\bar{X} - t_{n-1,1-\frac{\alpha}{2}} \frac{s}{\sqrt{n-1}}, \bar{X} + t_{n-1,1-\frac{\alpha}{2}} \frac{s}{\sqrt{n-1}}\right]$$

where

$\bar{X}$ - sample arithmetic mean,

s – sample standard deviation,

n – sample size,

$t_{n-1,1-\frac{\alpha}{2}}$ critical t for n-1 degree of freedom

$1-\alpha$ level of confidence

Example

Study of mobility by extension of the lumbar spine in individuals aged between 30 and 39 years

n=17, mean=40° şi t_α=2.11, s=2.36°

$$\left[40 - 2.11 * \frac{2.36}{\sqrt{17-1}}; 40 - 2.11 * \frac{2.36}{\sqrt{17-1}}\right]$$

$$[40 - 1.24; 40 + 1.24]$$

[38.76°; 41.24°]

Answer: mobility of the lumbar spine in young people is between 38.76° and 41.24° with a 5% error

Author: PhD. MsC. Bondor Cosmina-Ioana

Hypothesis Testing I

A ALWAYS

S SEEK

K KNOWLEDGE

Aim

Compare two groups

Compare one group with the population

Compare multiple groups

} two approaches:
- with confidence intervals
- with statistical tests

Objective: to find difference δ (delta)

Examples of medical questions

- Treatments
  - Is Losartan more efficient than Furosemide in reducing systolic blood pressure?
- Risk factors
  - Is Borrelia burgdorferi virus a risk factor for neuropathy?
  - Is the weight related to oxidative stress ?

- Medical question
  - if the premature-born boy's weight during childhood is influenced by the condition of their birth
- we cannot respond exactly to the question
  - but we can compare the average weight of 2 years boys who were born prematurely with the 2 years boys who were not
    - we will compare something with something to see a difference δ
    - what to compare
    - condition of what we compare should be similar except the feature of interest

Objective: δ

Step 1: establishing the objective

Transform the medical question into statistical hypotheses

ex. Is the infection with Covid-19 virus the cause of depression? → Is the frequency of depression higher in people who tested positive for Covid-19 in the last year compared to those who did not?

Objective of the research: demonstrate that **there is a difference** δ between two groups **which will maintain in majority o**f the studies if we repeat the study on all possible samples

Observation: there are design of the studies where we compare one group with a value (ex. the known mean of the population), or with a historical group or compare multiple groups

Objective: δ – clinically important

Step 1: establishing the difference δ

- the difference δ should be clinically important
- What difference δ should be between two drugs to declare one is more efficient than the other for the treatment of the targeted disease?
- In how many people (frequency δ) the treatment with the drug should be succsesfull compare with the placebo that one can say that the drug is more efficient than the placebo?

Step 2: establish the design of the study based on:

Objective: δ – clinically important

Study design

- what you want to compare:
  - 2 frequencies
  - 2 means
  - 2 medians
  - 3 frequencies
  - 3 means
  - etc...

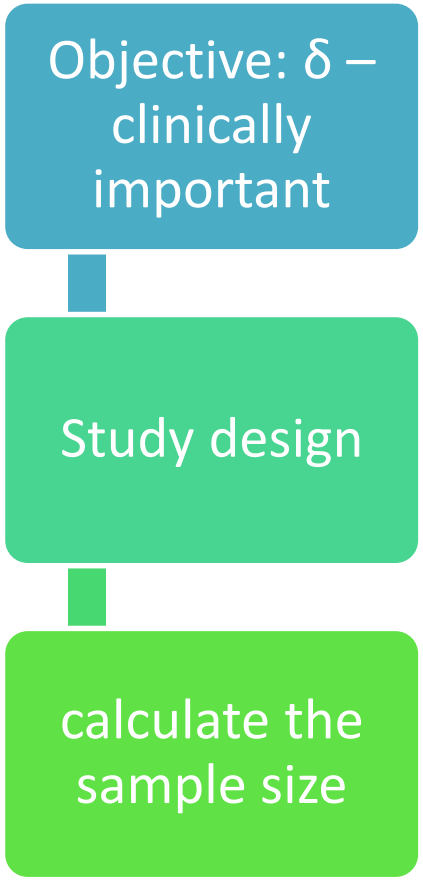
- the groups
  - dependent
  - independent

meaning: finding the right statistical test

Independent or dependent samples

- **Independent samples**
  - there is no connection between subjects in the two samples
- **Dependent samples**
  - subjects are related somehow
    - subjects are paired,
    - subjects are matched
      - age,
      - gender
      - etc.,
    - the same subjects tested two/multiple times
      - after the treatment
      - before the treatment

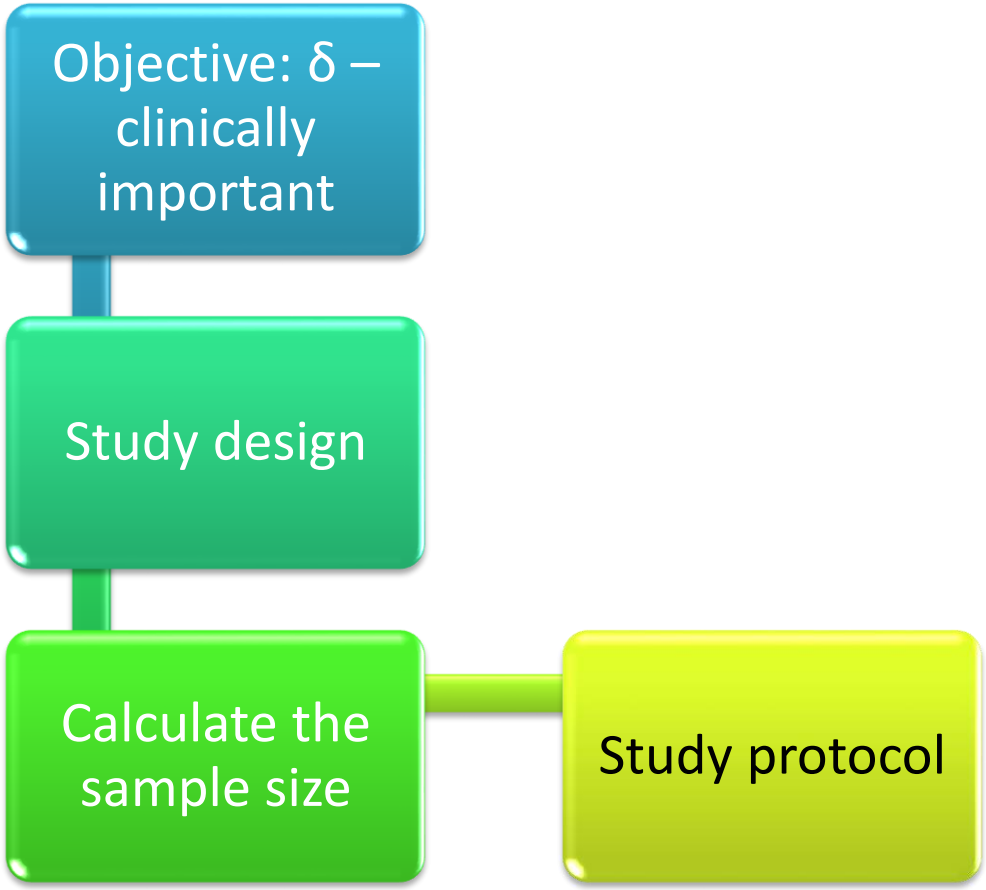
Step 3: compute the needed samples size



- different formulas for different designs
- a lot of possible design: you will need a statistician

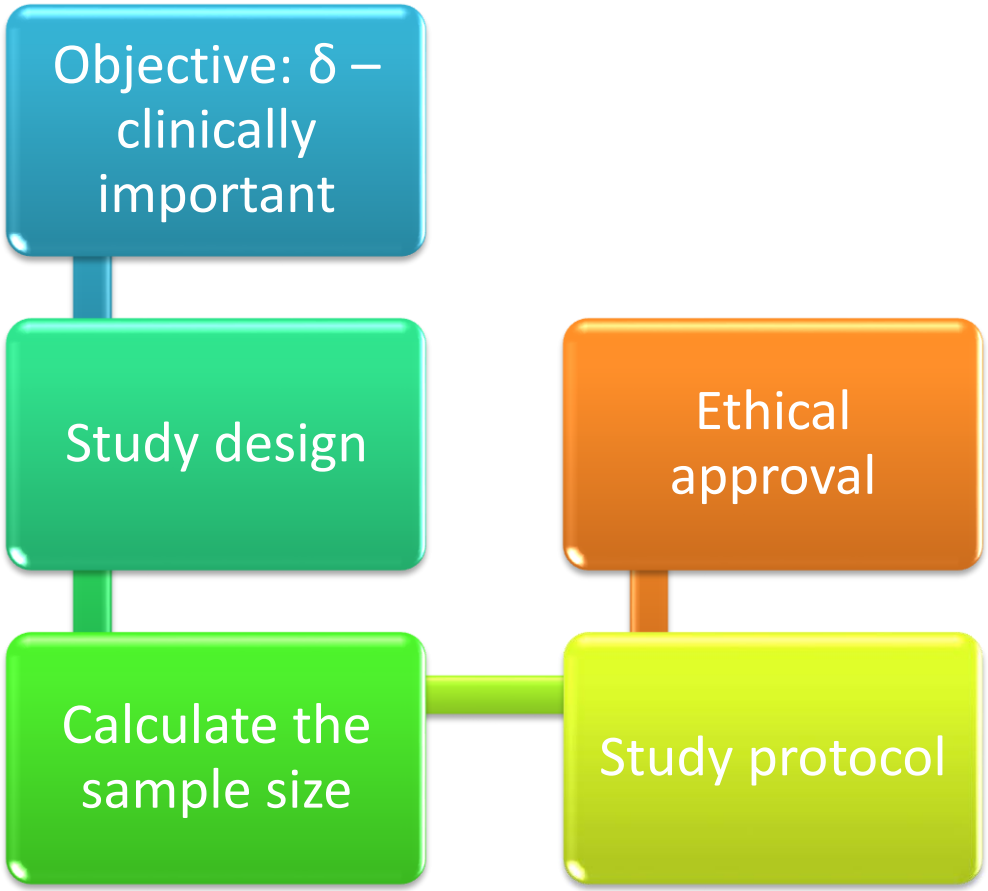
- there are online calculators
- example for two proportion:
- https://statpages.info/proppowr.html

Step 4: making the protocol of the study



- there are templates for each kind of study,
- a protocol of the study have to answer to the questions:
  - where?
  - when?
  - how?
  - who?
  - etc.

Step 5: apply for ethical approval

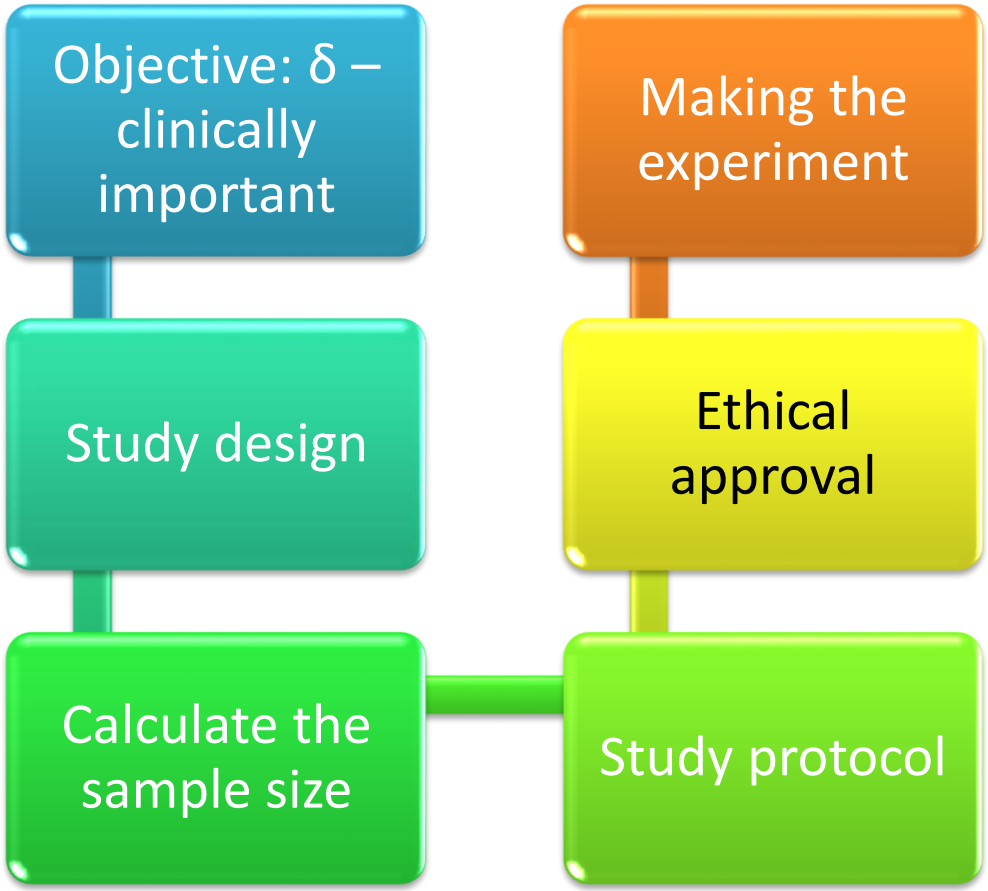


- each health institution should have a research ethics committees
- submit the protocol to the ethic committee
- wait for approval

Step 5: only for clinical trial

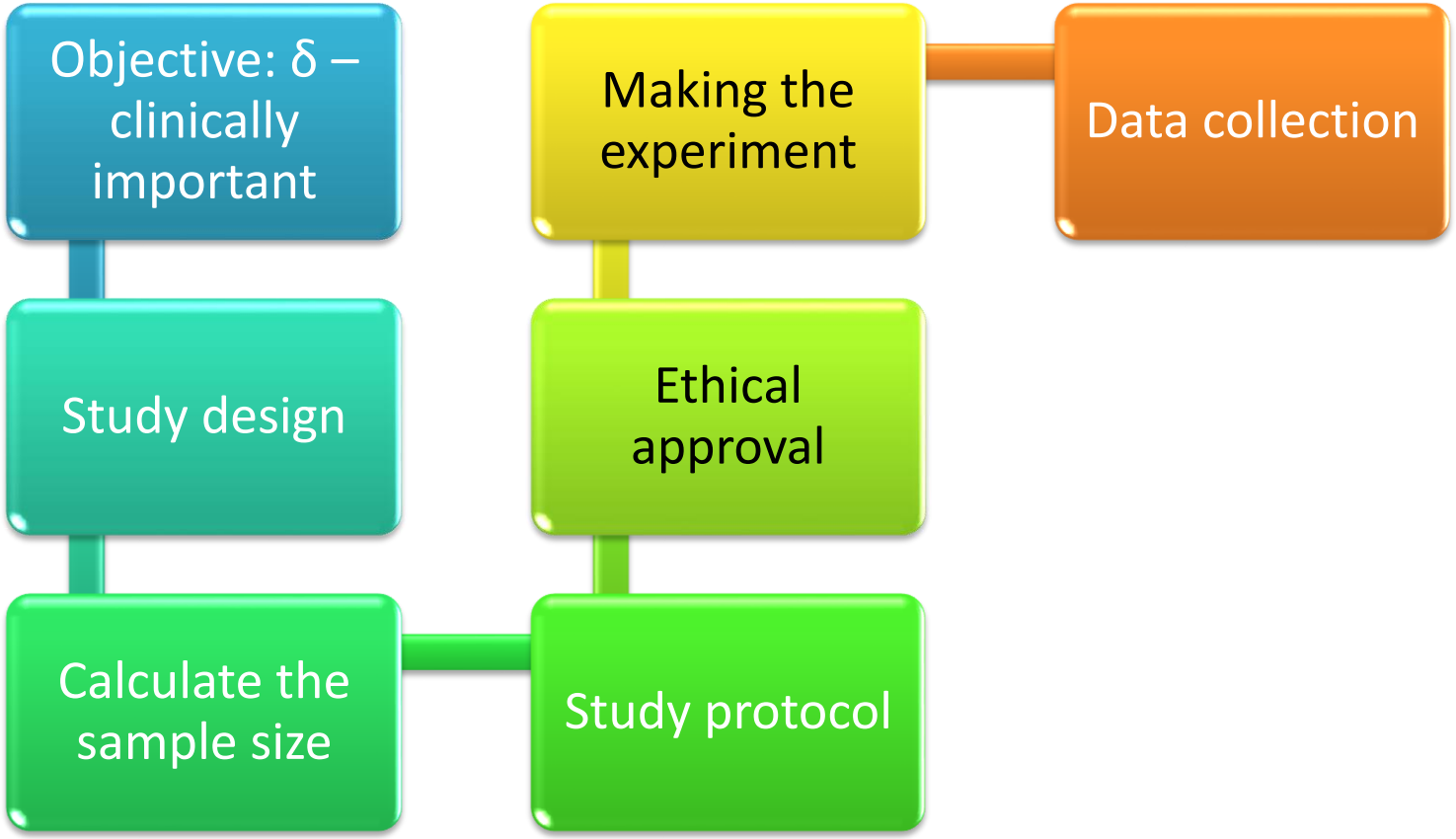
- registration of the trial at the WHO International Clinical Trials Registry Platform (ICTRP) before starting the trial

Step 6: carrying out the study



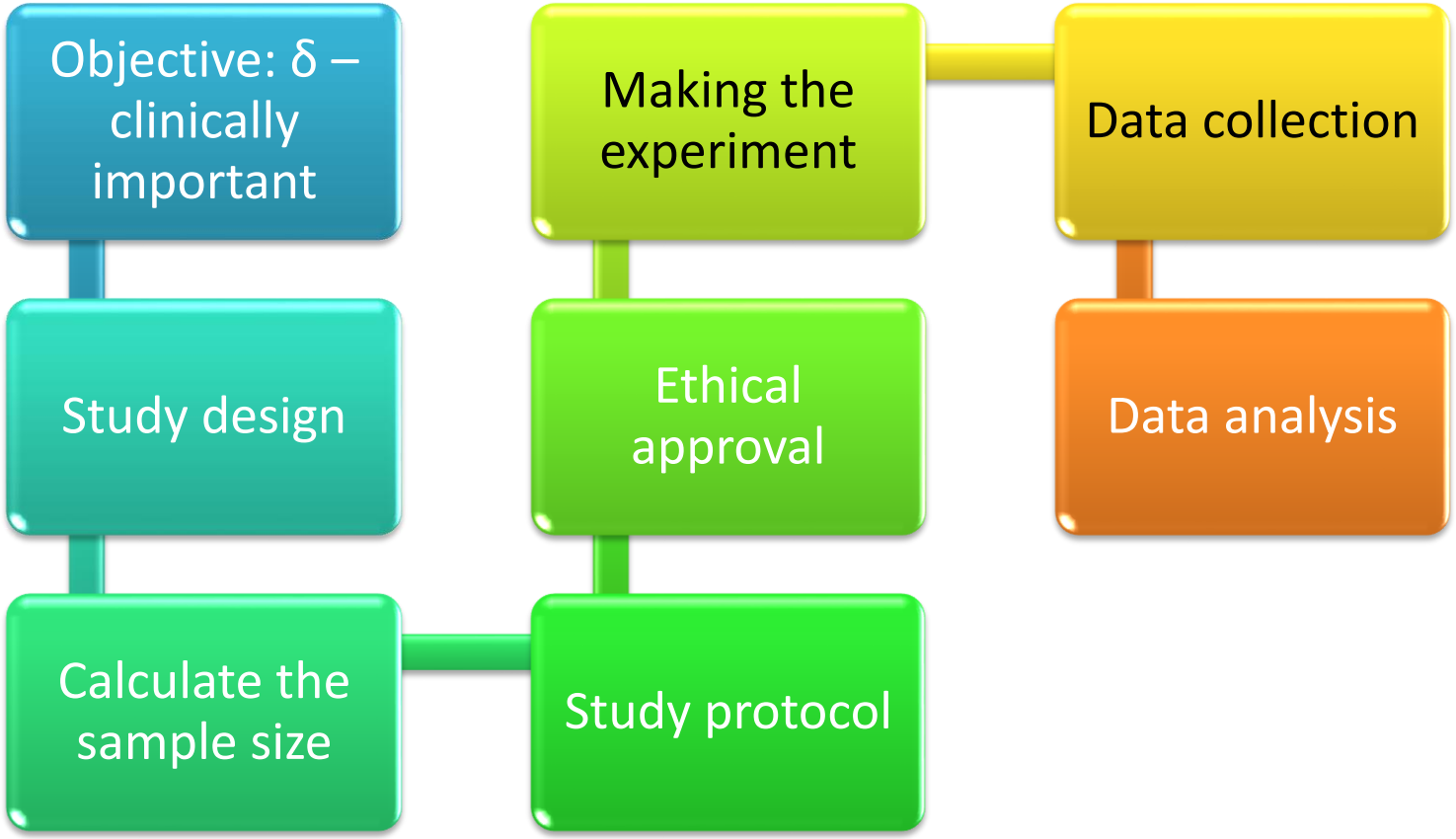
- after you establish the protocol of the study in all details you can start the experiment

Step 7: collect the data



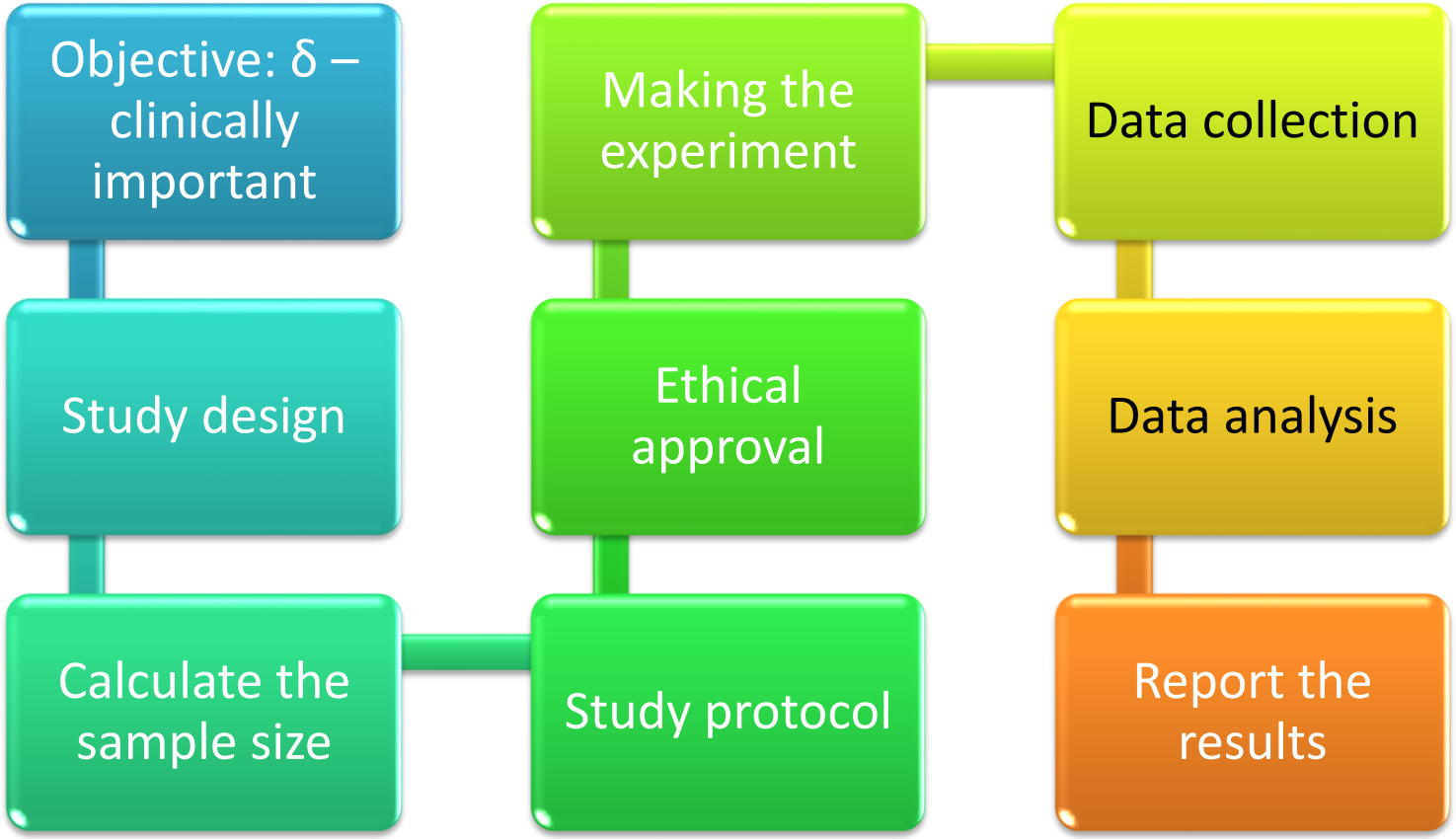
- collect the data about the variables of interest in Excel table or other applications
- data should be deidentified (no personal data should be collected)

Step 8: analyze the data



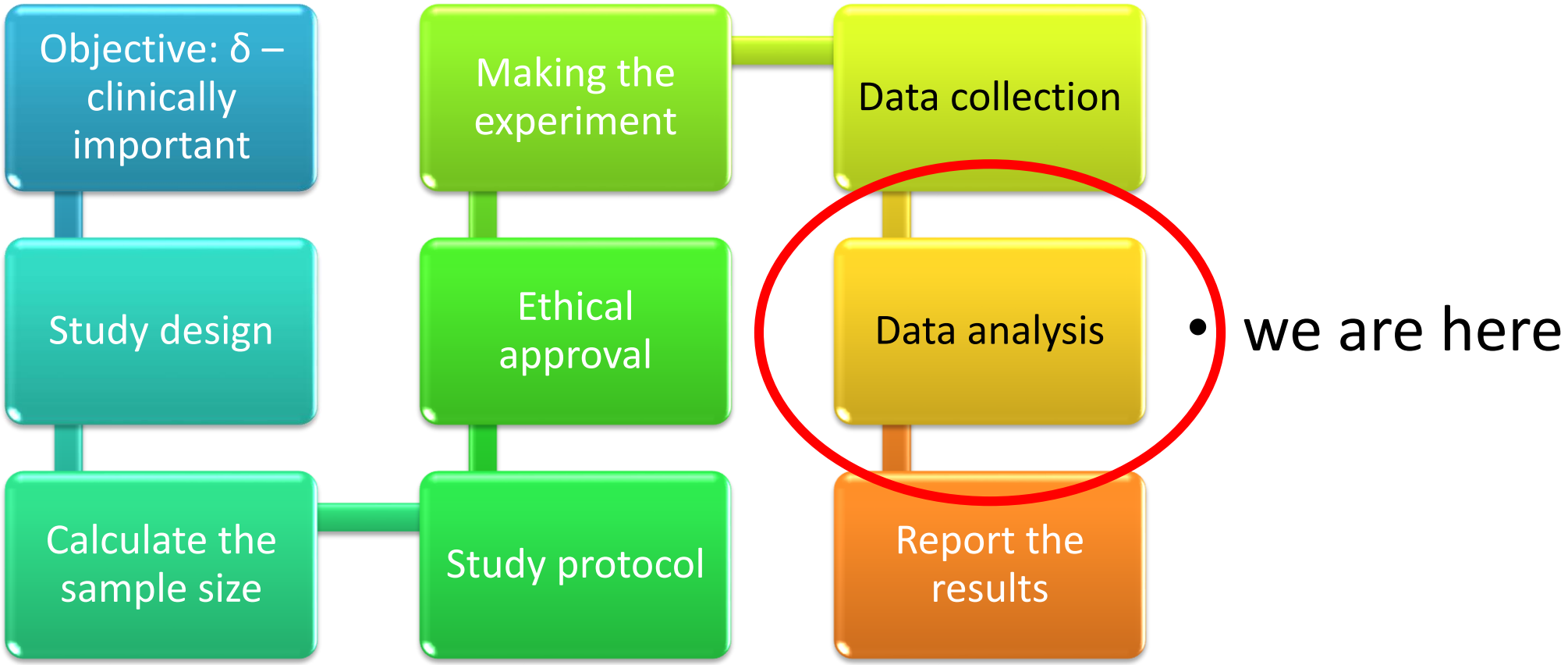
- use an application to analyze the data with statistical tests
  - SPSS
  - SAS
  - STATA
  - STATISTICA
  - R language (Jamovi)
  - etc.
- choose the conclusion for the medical question

Step 9: report the results



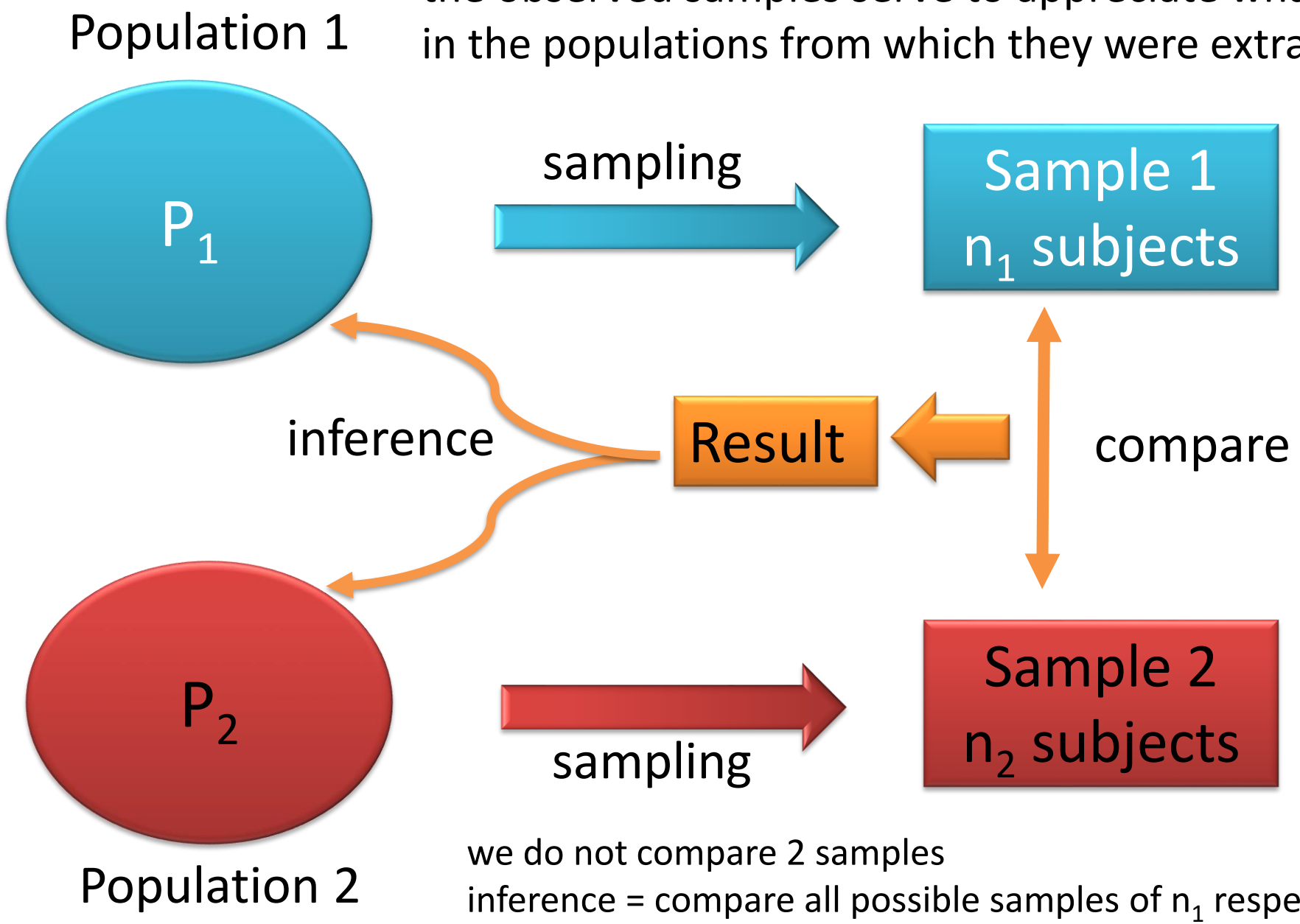
- report the results of the study in:
  - medical conferences
  - mass media
  - journals as original articles
  - for clinical trial in WHO International Clinical Trials Registry Platform

Compare samples



- we are here

the observed samples serve to appreciate what is happening in the populations from which they were extracted



we do not compare 2 samples
inference = compare all possible samples of n_1 respectively n_2 subjects that can be taken from the populations

Compare two groups with confidence intervals method

Research question example

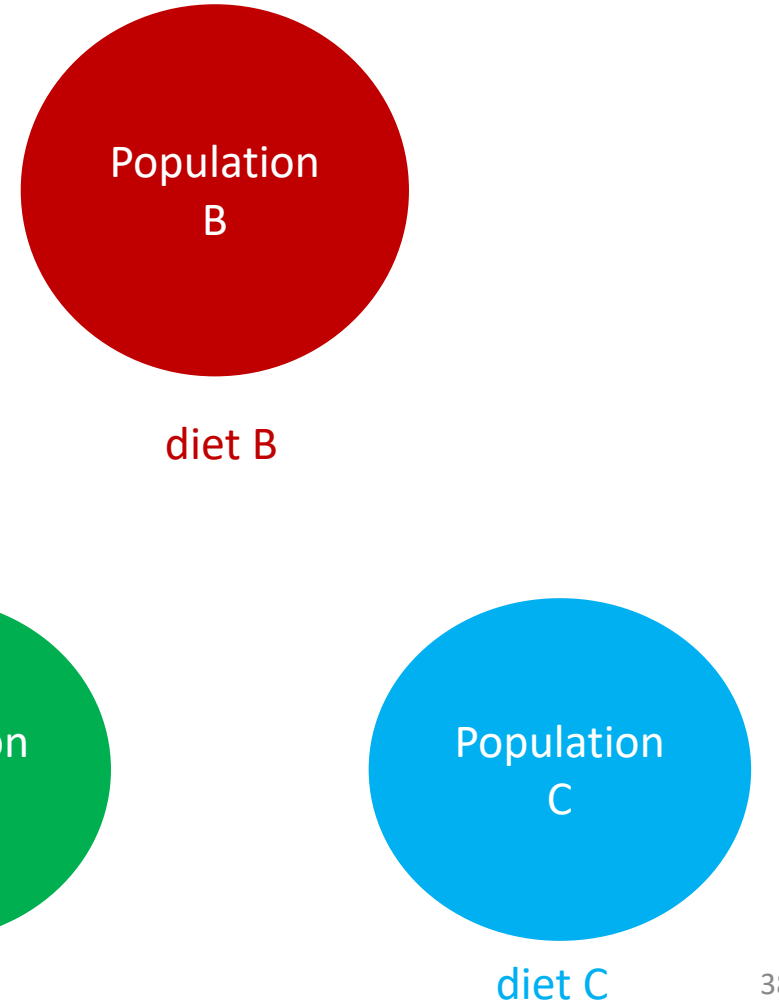
- Which diet is the best diet in lowering the cholesterol in people with hipercholesterolemia?
  - mediteranean diet
  - paleo diet
  - high-protein diet

Compare two groups with confidence intervals method

- subjects with hipercholesterolemia

Objectiv: To compare cholesterol

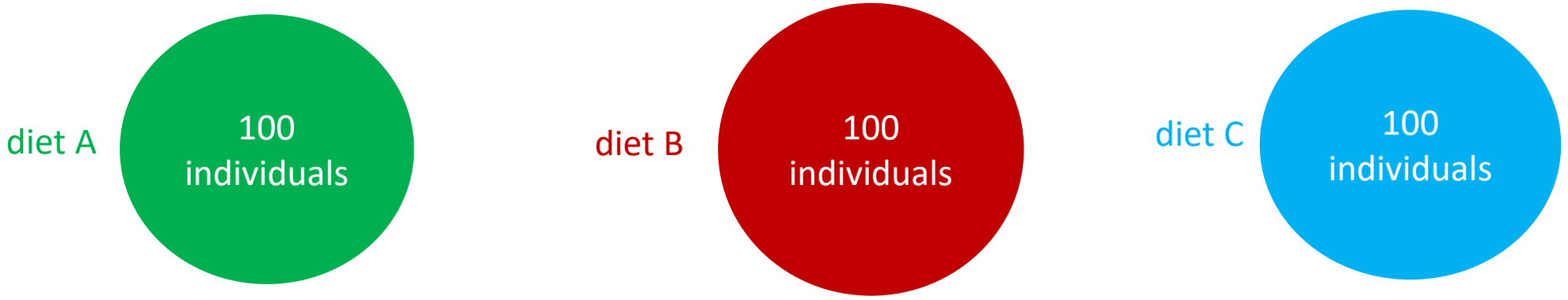
- diet A (mediteranean diet)
- diet B (paleo diet)
- diet C (high-protein diet).



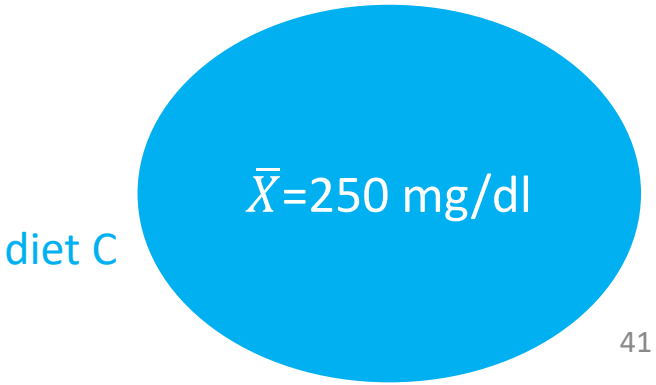
Compare two groups with confidence intervals method

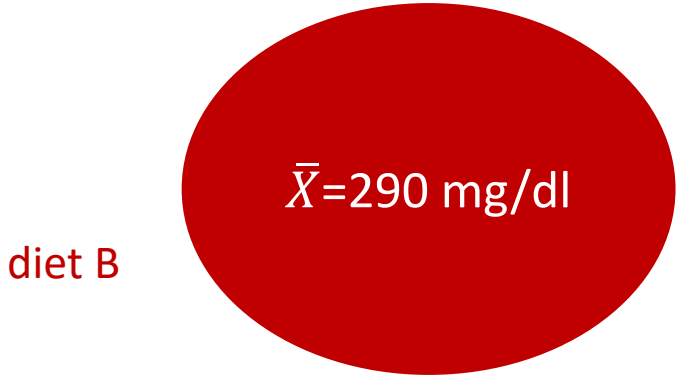
- Objective: To compare **cholesterol** in subjects with hypercholesterolemia who followed **diet A** (Mediterranean diet) versus **diet B** (Paleo diet) versus **diet C** (High-protein diet).

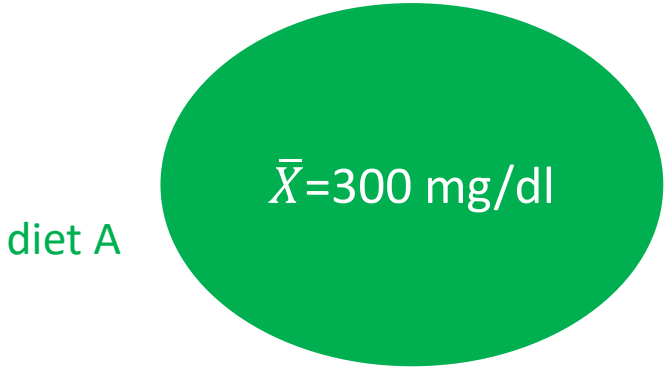
- Method:
- a difference by 10mg/dl we consider it clinically important.
- We design the study to detect a difference of the cholesterol by 10mg/dl or more between the diets in the population with a 95% probability
  - we calculate the needed sample size take into the study 300 subjects randomly selected

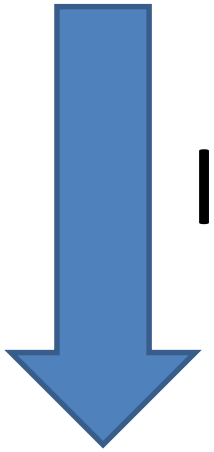


Findings on the samples





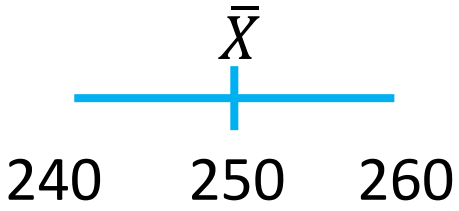


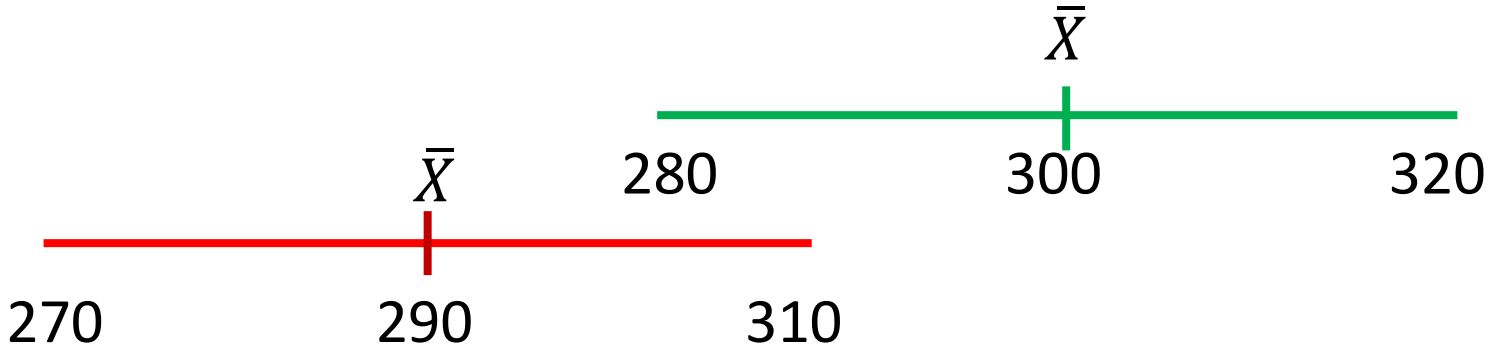


Inference to the population

we estimate 95% confidence intervals

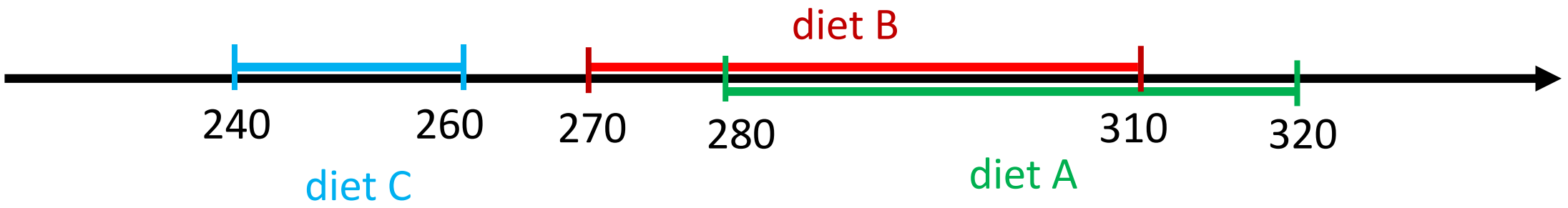
diet A: $\bar{X}$= 300mg/dl, 95%CI 280-320
diet B: $\bar{X}$= 290mg/dl, 95%CI 270-310
diet C: $\bar{X}$= 250mg/dl, 95%CI 240-260





Inference to the population

we estimate 95% confidence intervals



diet A: $\bar{X}$= 300mg/dl, 95%CI 280-320
diet B: $\bar{X}$= 290mg/dl, 95%CI 270-310
diet C: $\bar{X}$= 250mg/dl, 95%CI 240-260

- 95% CI 240-260

Interpretation: if we measure the whole population of subjects following diet C (high protein diet) than average cholesterol of whole population will be somewhere between 240mg/dl and 260mg/dl with 5% error,

(we do not know where exactly in this interval will be, only that it is in this interval with 95% probability.

Observation: if we want more precisse results we should take more subjects into the study).

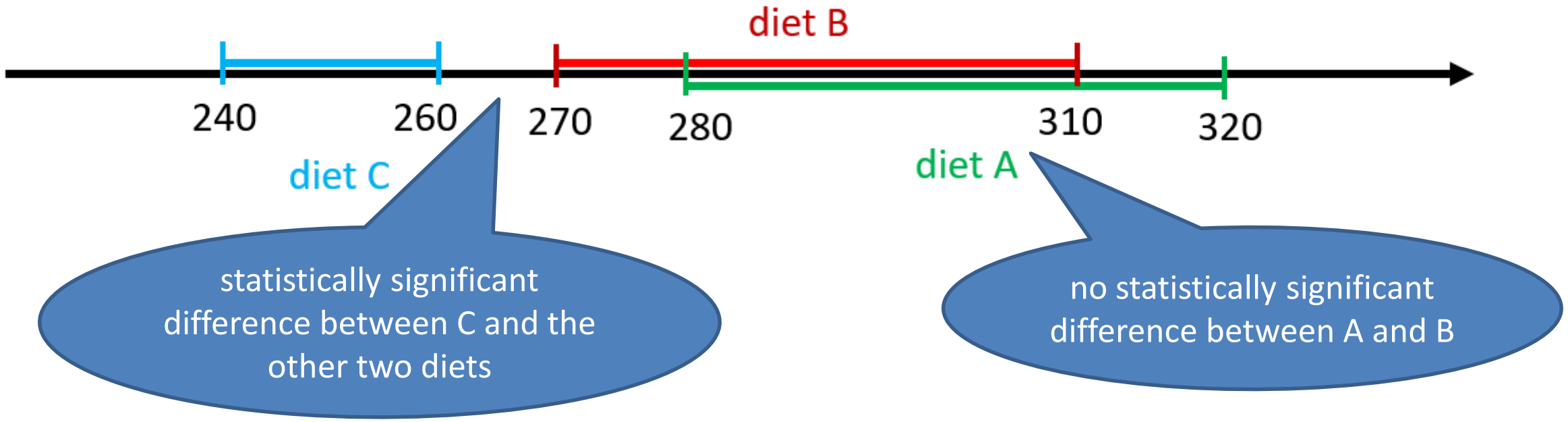
- In the worst case average cholesterol can be 260mg/dl, in the best case 240mg/dl

- 95% CI 270-310

Interpretation: if we measure the whole population of subjects following mediteranean diet than average cholesterol will be somewhere between 270mg/dl and 310mg/dl with 5% error,

we do not know where exactly in this interval is.

- In the worst case it can be 310mg/dl, in the best case can be 270mg/dl



- <u>Conclusion 1:</u> this study demonstrate that subjects that follow diet C have low cholesterol than those who follow diet B (with 10mg/dl) and diet A (with 20mg/dl) (95% CI of diet C did not overlap the 95% CI of diet B and A. Any value in the 95% CI of diet C < any value in the 95% CI of diet B and A)

- <u>Conclusion 2:</u> we fail to demonstrate that there is a significant statistically difference between cholesterol in diet B compare with diet A (95% CIs of diet A and B overlap. Observation: The difference between diet A and B is less than 10mg/dl, if we want to know the difference we should repeat the study with more subjects to have a better precision)

Compare two groups arithmetic means with statistical tests

steps of a statistical test exemplified with t test for independent samples in case of unequal variances

Comparing two means

- General:

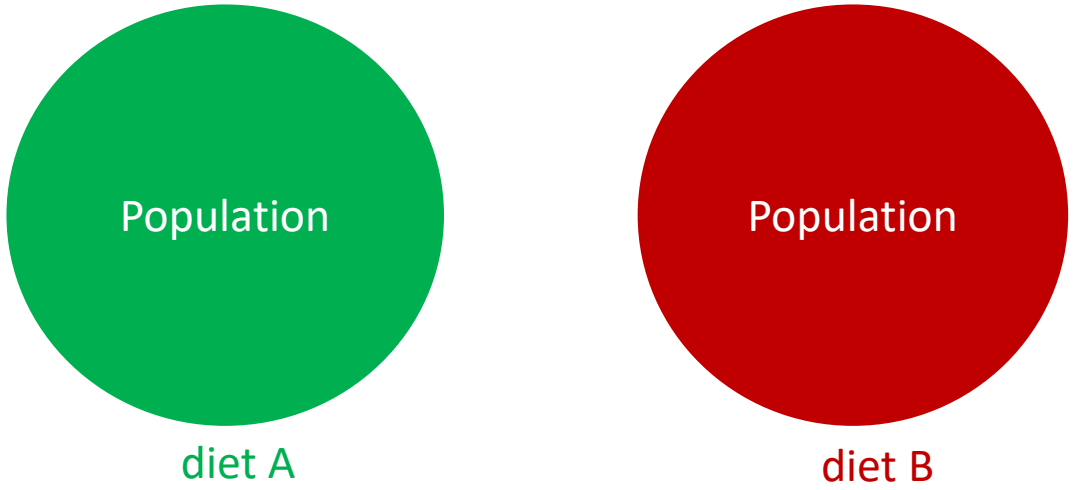
population P_1

population P_2

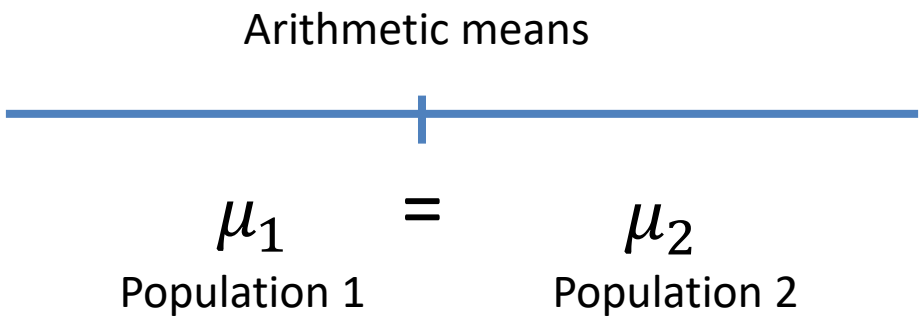
- Example:

patients with

- diet A
- diet B;

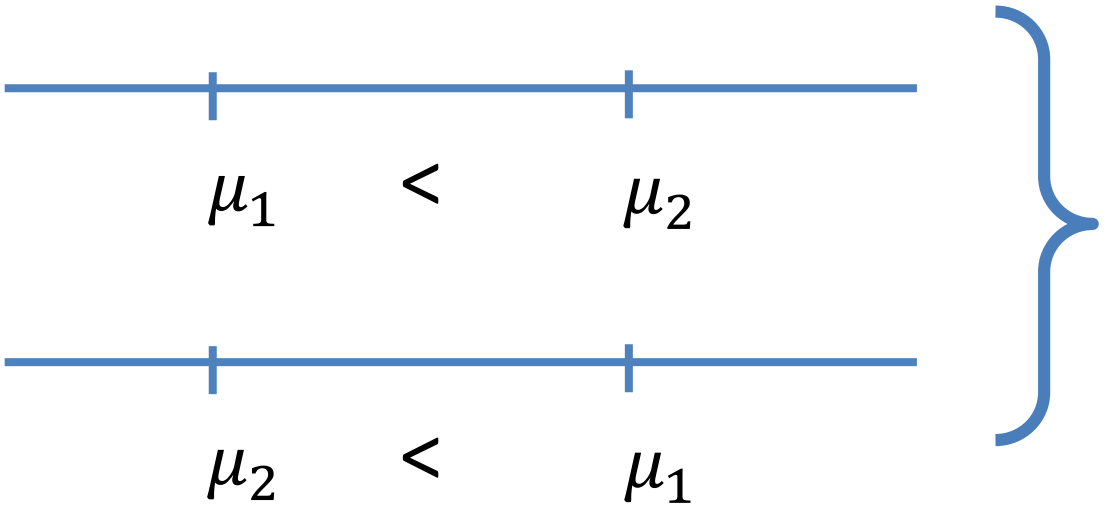


Possible results



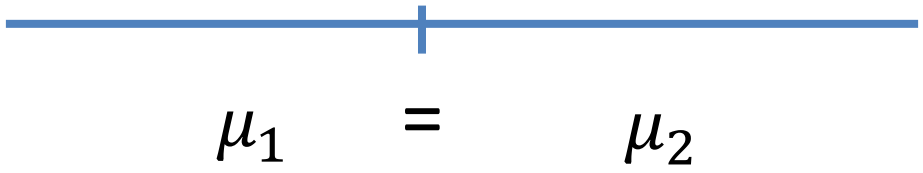
$\mu_1 = \mu_2$(or near equal)

OR



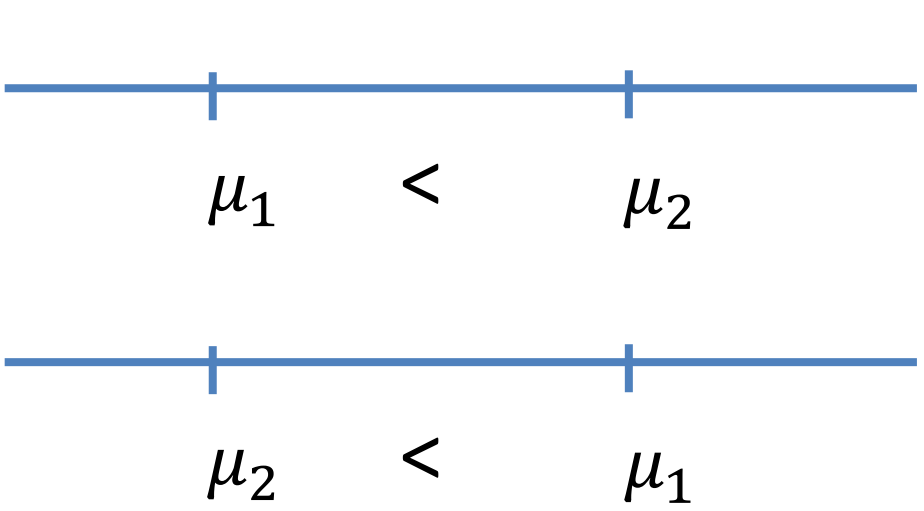
$\mu_1 \neq \mu_2$

Possible results



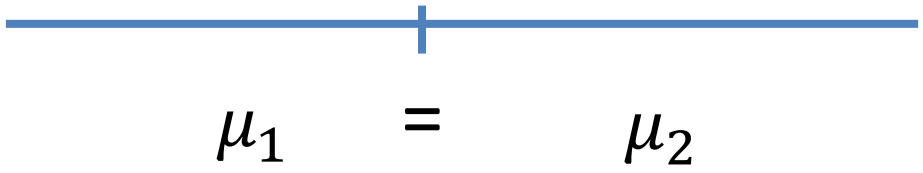
$\mu_1 = \mu_2$ H0: there is no difference

OR



$\mu_1 \neq \mu_2$ H1: there is difference

Possible results

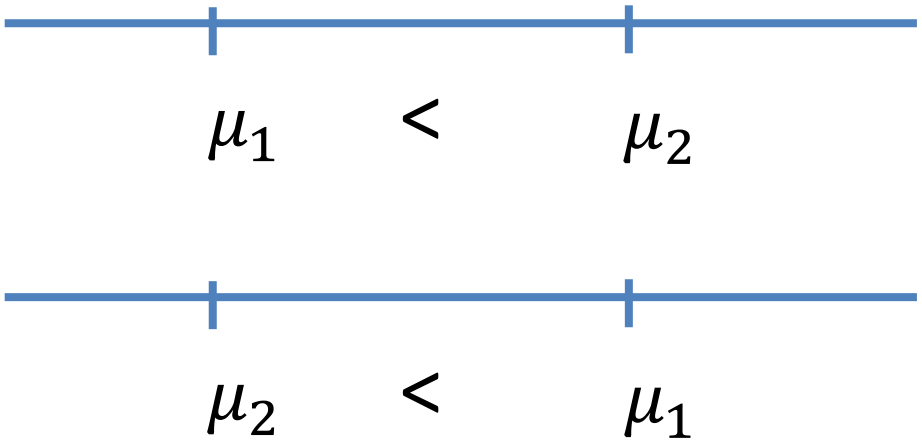


$\mu_1 = \mu_2$ H0: there is no difference

OR

Which one is true?

we will choose H0 or H1 based on a probability



$\mu_1 \neq \mu_2$ H1: there is difference

Select samples, make the study, obtain the results on the samples

- General:

population P_1

population P_2

$\bar{X}_1$ = average for first sample,

$\bar{X}_2$ = average for second sample,

s_1^2 = variances for first sample

s_2^2 = variances for second sample

- Example:

patients with

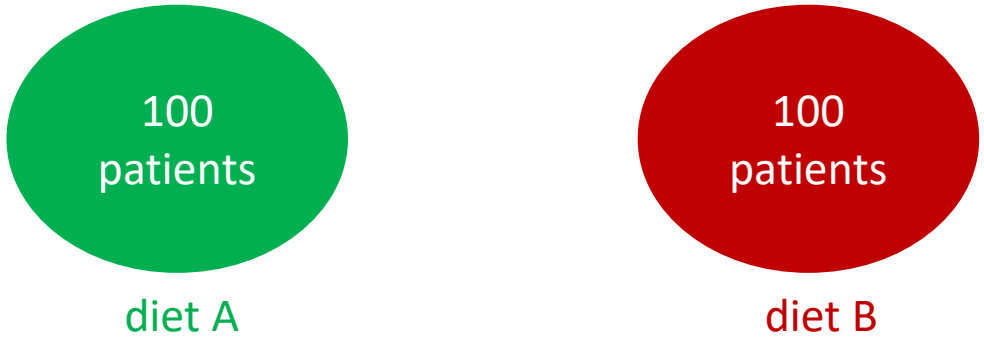
- diet A
- diet B;

= cholesterol $\bar{X}_1$ for first sample 300mg/dl,

= cholesterol $\bar{X}_2$ for second sample 290mg/dl,

= cholesterol s_1^2 for first sample 100

= cholesterol s_2^2 for second sample 100



Select samples, make the study, obtain the results on the samples

- General:

population P_1

population P_2

$\bar{X}_1$ = average for first sample,

$\bar{X}_2$ = average for second sample,

s_1^2 = variances for first sample

s_2^2 = variances for second sample

- Example:

patients with

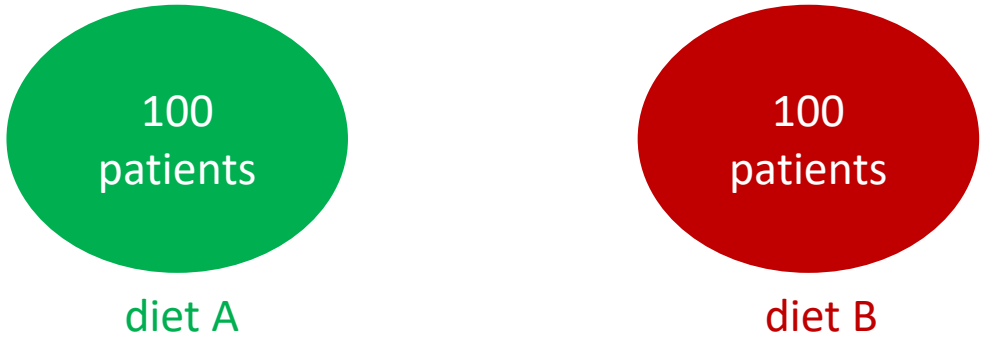
- diet A
- diet B;

= cholesterol $\bar{X}_1$ for first sample 300mg/dl,

= cholesterol $\bar{X}_2$ for second sample 290mg/dl,

= cholesterol s_1^2 for first sample 100

= cholesterol s_2^2 for second sample 100



Step 1 of a statistical test: Formulating Hypothesis

General:

- H0 - null hypothesis: at the population level, there is no statistical significant difference between group 1 average and group 2 average
- H1 - alternative hypothesis: at the population level, there is statistical significant difference between group 1 average and group 2 average

Example:

- H0 - null hypothesis: at the population level there is no statistically significant difference between the averages of cholesterol for patients with diet A versus diet B
- H1 - alternative hypothesis: at the population level there is statistically significant difference between the averages of cholesterol for patients with diet A versus diet B

Step 1 of a statistical test: Formulating Hypothesis

General:

- H0 - null hypothesis: at the population level, there is **no** statistical significant difference between group 1 average and group 2 average

- H1 - alternative hypothesis: at the population level, there is statistical significant difference between group 1 average and group 2 average

Example:

- H0 - null hypothesis: at the population level there is **no** statistically significant difference between the averages of cholesterol for patients with diet A versus diet B

- H1 - alternative hypothesis: at the population level there is statistically significant difference between the averages of cholesterol for patients with diet A versus diet B

Step 1 of a statistical test: Formulating Hypothesis

- H0 - null hypothesis: to demonstrate there is no statistically significant difference between the averages of cholesterol for patients with diet A versus diet B

"statistically significant" = suggest that is at the population level, so the term "population level" is redundant

- H1 - alternative hypothesis: there is no statistically significant difference between the averages of cholesterol for patients with diet A versus diet B

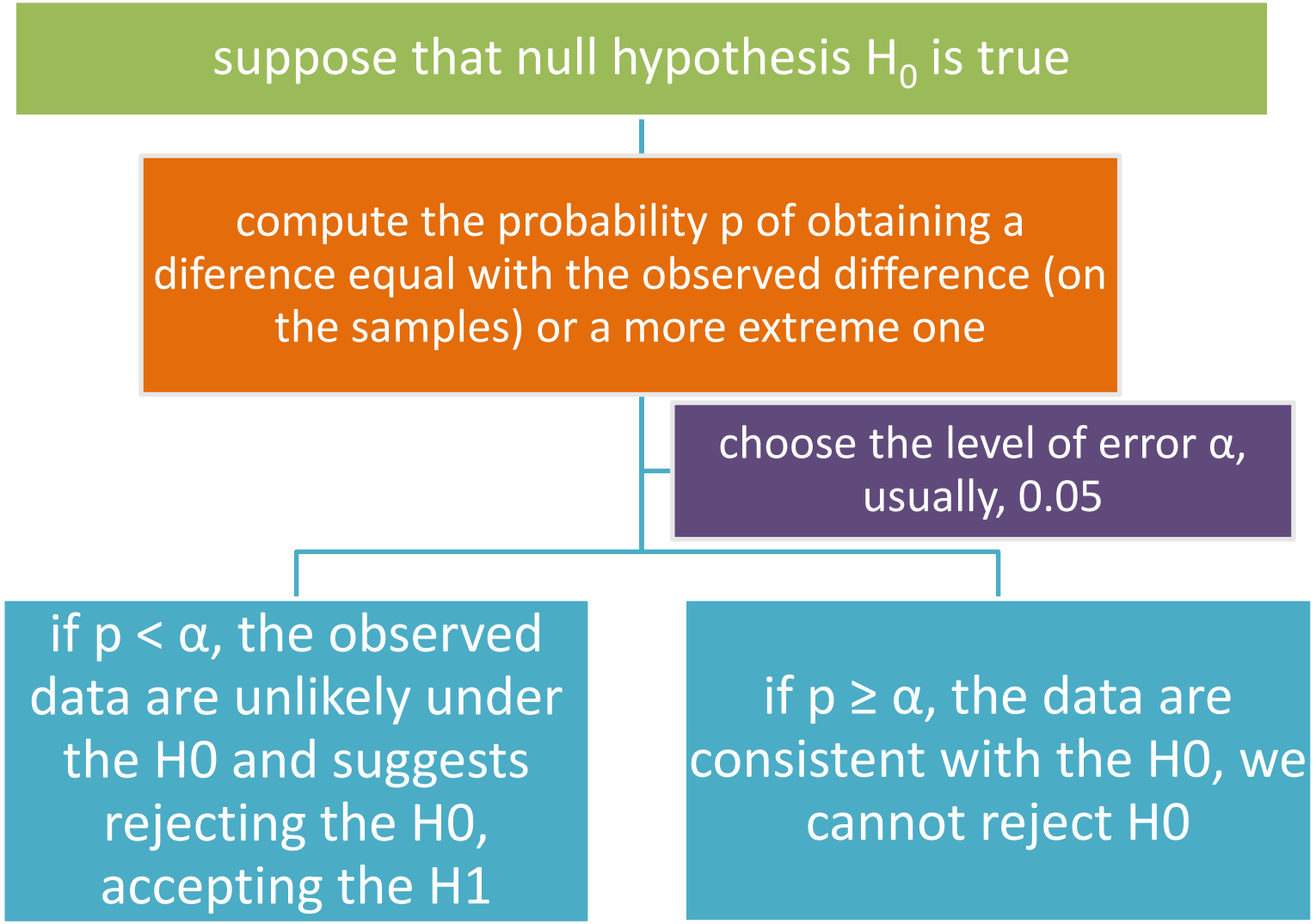
"statistically" – we estimate statistically the cholesterol average in the population of subjects with diet A with the cholesterol average from the sample with diet B

"significant" – we will choose a level of significance, i.e. a probability level for the confidence

Alternative=the other possible situation, when null hypothesis is not true

Method:

absurd assumption



p-value interpretation

$p < \alpha$ (small p-value)

- the observed data are unlikely under the null hypothesis and suggests rejecting the H0, accepting the H1
- suggests that the observed result is statistically significant and unlikely to have occurred by chance. We reject H0, we are in favor of H1
- the result is considered statistically significant
- the obtain difference does not automatically mean is a medical relevant result

- $p \geq \alpha$
- the data are consistent with the null hypothesis, we cannot reject H0
- suggests that the observed result is not statistically significant and is likely to have occurred by chance. We cannot reject H0
- the result is considered statistically not significant
- the obtain difference does not automatically mean that is no medical relevant result

Step 2 of a statistical test – Finding the proper statistical test to apply by answering to the questions:

General:

- What we compare?
  - means,
  - variances,
  - percentages
  - ….
- How is the distribution?
  - normal
  - non-normal,
  - binomial
  - …
- How are the samples?
  - independent
  - dependent
- How is the sample size?
  - small (<30)
  - large
- How are the variances?
  - equal
  - unequal
- How many samples we want to compare?
  - one,
  - two,
  - three
  - many

- Example:
  - normal distributions
    - we should check the distribution,
      - Result: stress is normally distributed in both groups,
  - unequal variances
    - we should check (not the subject of this lecture, see next lecture),
      - Result: variances are not equal,
  - two arithmetic means are compared,
  - groups are independent
    - between groups - within the sample there are no subjects that are extremely similar to the others (repeated observations on the same subjects, measurements on the left hand and on the right hand of the same subject, …)
    - not the case - between the subjects in the same group
- we will choose:
  - **Student t-test for independent samples in the case of unequal variances**

Step 3. Select the level of significance

General:

- α - level of significance
  - probability of incorrectly rejecting the null hypothesis when it is actually true
- **Usually, α is choose to be 0.05**
  - **(corresponding to Z_α=1.96)**
- p-value or p
  - probability to obtain a result more extreme than the result obtained from the observed data.
  - If p < α, then we will **reject the null hypothesis**: we demonstrate that there is difference between the arithmetic means and say that the test is statistically significant: we have difference between the means

Example:

- we choose α = 0.05

Step 4. Determine the value the test parameter must attain to be declared significant

General:

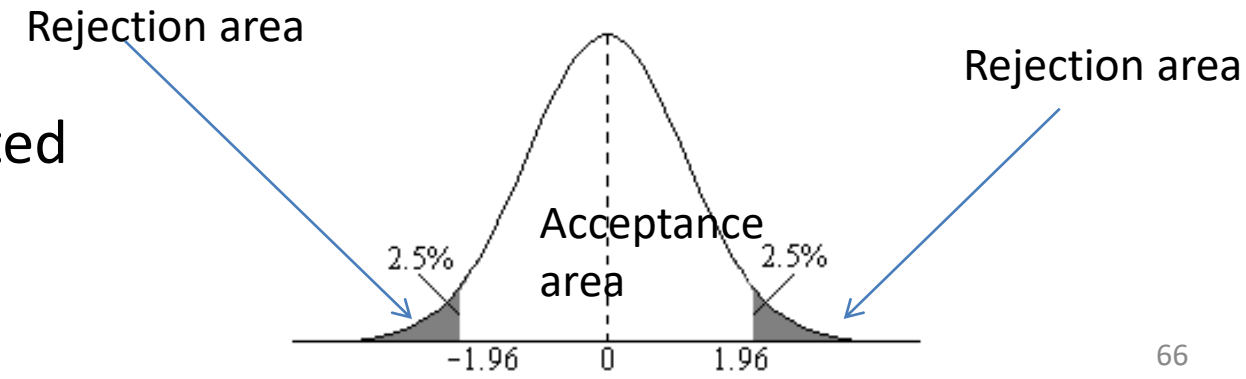
- **critical value**
  - ex. for t-test t_α, for Z test Z_α
- Rejection area – critical interval
  - interval where H_0 should be rejected
  - the interval depend on the distribution of the test and sometimes degrees of freedom
  - ex. $(-\infty, -Z_\alpha) \cup (Z_\alpha, \infty)$
- Acceptance interval:
  - interval where H_0 should be accepted
  - ex.$[-Z_\alpha, Z_\alpha]$
  - for α=0.05, Z_α=1.96

Example:

t test – Student distribution – **critical value** is different for each degree of freedom

- t_α = 1.97 = **critical value for 100+100-2 degree of freedom**
- rejection interval $(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ = $(-\infty, -1.97) \cup (1.97, \infty)$
- acceptance interval: $[-t_\alpha, t_\alpha]$=[-1.97,1.97]

= an interval around 0, if we found a difference in this interval it means is a small difference and H_0 will be accepted (there is no difference)



Step 4. Determine the value the test parameter must attain to be declared significant

General:

- **critical value**
  - ex. for t-test t_α, for Z test Z_α)
- Rejection area – critical interval
  - interval where H_0 should be rejected
  - different depending on the distribution of the test and sometimes degrees of freedom
  - ex. $(-\infty, -Z_\alpha] \cup [Z_\alpha, \infty)$
- Acceptance interval:
  - interval where H_0 should be accepted
  - ex.$(-Z_\alpha, Z_\alpha)$
  - for α=0.05, Z_α=1.96

Example:

t test – Student distribution – **critical value** is different for each degree of freedom

- t_α = 1.97 = **critical value for 100+100-2 degree of freedom**
- rejection interval $(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ = $(-\infty, -1.97) \cup (1.97, \infty)$
- acceptance interval: $[-t_\alpha, t_\alpha]$=[-1.97,1.97]

  = an interval around 0, if we found a difference in this interval it means is a small difference and H_0 will be accepted (there is no difference)

Step 5: computing the formula for test parameter

General:

- any statistical test have a test parameter
- test two means, then $\bar{X}_1 - \bar{X}_2$ should be in the formula, divided by standard errors

if parameter of the test is small, then the difference between $\bar{X}_1$ and $\bar{X}_2$ is small, near 0.

Example:

- the parameter of the test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

difference between averages.

material A: $\bar{X}_1$ = 300MPa, s = 100MPa
material B: $\bar{X}_2$ = 290MPa, s = 100MPa

$$t = \frac{300-290}{\sqrt{\frac{100^2}{100}+\frac{100^2}{100}}} = \frac{10}{\sqrt{200}} = \frac{10}{14.14} = 0.71,$$

p=0.48

p can be computed only by computers – have a very complicated formula

Step 5: computing the formula for test parameter

General:

any statistical test have a test parameter based on the difference between what we test

test two means than $\bar{X}_1 - \bar{X}_2$ should be in the formula, divided by standard errors

if parameter of the test is small, then the difference between $\bar{X}_1$ *and* $\bar{X}_2$ is small, near 0.

Example:

- the parameter of the test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

difference between averages.

diet A: $\bar{X}_1$ = 300mg/dl, s = 100mg/dl
diet B: $\bar{X}_2$ = 290mg/dl, s = 100mg/dl

$$t = \frac{300-290}{\sqrt{\frac{100^2}{100}+\frac{100^2}{100}}} = \frac{10}{\sqrt{200}} = \frac{10}{14.14} = 0.71,$$

p=0.48

p can be computed only by computers – have a very complicated formula

Step 6. Draw and state the conclusion

General:

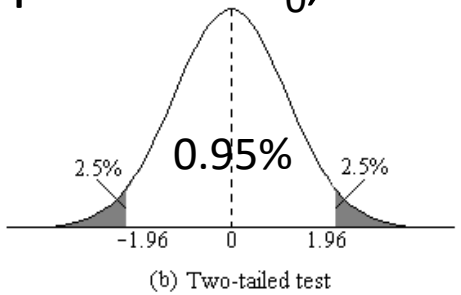
If parameter of the test ∈ acceptance interval --> we fail to reject null hypothesis H_0

If parameter of the test ∈ rejection interval --> we reject null hypothesis H_0, accept H_1

the same as

If p≥0.05 we fail to reject null hypothesis H_0

If p<0.05 we reject null hypothesis H_0, accept H_1



(b) Two-tailed test

Example:

if t ∈$[-t_\alpha, t_\alpha]$ – acceptance interval =small differences.

if t ∈$(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ – rejection interval =high differences.

the same as

If p≥0.05, then H_0 is in favor – fail to reject H_0

If p<0.05 than H_0 is false, we reject H_0 – accept H_1

Step 6. Draw and state the conclusion

General:

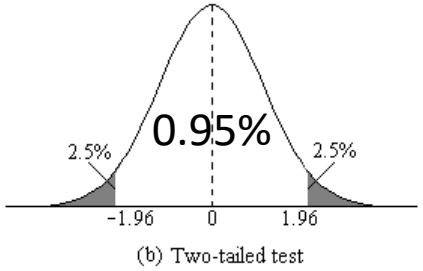
If parameter of the test ∈ acceptance interval --> we fail to reject null hypothesis H_0

If parameter of the test ∈ rejection interval --> we reject null hypothesis H_0, accept H_1

the same as

If p≥0.05 we fail to reject null hypothesis H_0

If p<0.05 we reject null hypothesis H_0, accept H_1



Example:

if t ∈$[-t_\alpha, t_\alpha]$ – acceptance interval =small differences.

if t ∈$(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ – rejection interval =high differences.

the same as

If p≥0.05, then H_0 is in favor – fail to reject H_0

If p<0.05 than H_0 is false, we reject H_0 – accept H_1

Example:

$-1.97 \leq t = 0.71 \leq 1.97$

we fail to reject null hypothesis,

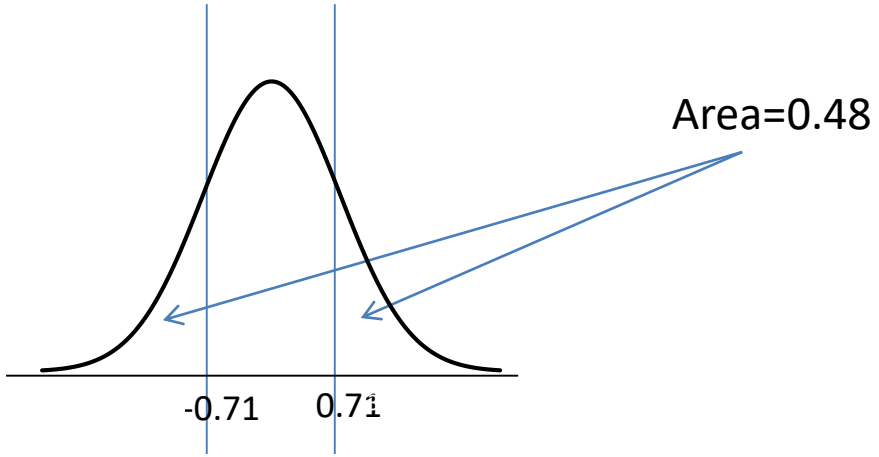
the average cholesterol in people who have diet A did not differ statistically significant from the average cholesterol in people who have diet B

Area under the curve for $t = 0.71$ is 0.48, $p = 0.48 > 0.05$ we fail to reject null hypothesis,

the average cholesterol in people who have diet A did not differ statistically significant from the average cholesterol in people who have diet B

~~We accept the null hypothesis~~

We fail to reject the null hypothesis



Example 2

diet C compare with B

Comparing two means

- General:

population P_1

population P_2

$\bar{X}_1$ = average for first sample,

$\bar{X}_2$ = average for second sample,

s_1^2 = variances for first sample

s_2^2 = variances for second sample

- Example:

= subjects with diet C

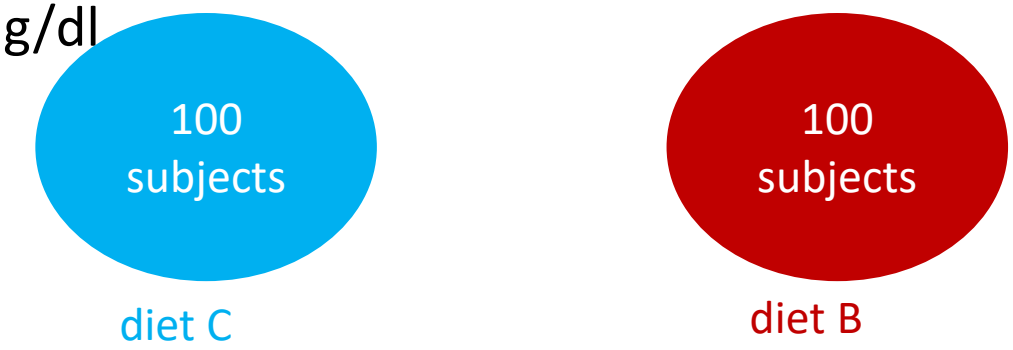
= subjects with diet B;

= average cholesterol for first sample 250mg/dl,

= average cholesterol for second sample 290 mg/dl,

= cholesterol variances for first sample 50mg/dl

= cholesterol variances for second sample 100 mg/dl



Step 1 of a statistical test: Formulating Hypothesis

General:

- H0 - null hypothesis: there is **no** statistically significant difference between group 1 average and group 2 average
- H1 - alternative hypothesis: there is statistically significant difference between group 1 average and group 2 average

Example:

- H0 - null hypothesis: there is **no** statistically significant difference between the averages of cholesterol for subjects with diet C versus diet B
- H1 - alternative hypothesis: there is statistically significant difference between the averages of cholesterol for subjects with diet C versus diet B

Step 2 of a statistical test – Finding the proper statistical test to apply after answering to the questions:

General:

- What we compare?
  - means,
  - variances,
  - percentages
  - ….
- How is the distribution?
  - normal
  - non-normal,
  - binomial
  - …
- How are the samples?
  - independent
  - dependent
- How is the sample size?
  - small (<30)
  - large
- How are the variances?
  - equal
  - unequal
- How many samples we want to compare?
  - one,
  - two,
  - three
  - many

- Example:
  - normal distribution
    - usually we should check the distribution,
      - Result: cholesterol is normally distributed in both groups,
  - unequal variances
    - usually we should check,
      - Result: unequal variances,
  - two arithmetic means are compared,
  - groups are independent
  - subjects are independent in the groups
- we will choose:
  - t-test for independent samples in the case of unequal variances

Step 3. Select the level of significance of the statistic test

General:

- α - level of significance
  - probability of incorrectly rejecting the null hypothesis when it is actually true
- **Usually, α is choose to be 0.05**
  - **(corresponding to Z_{α}=1.96)**
- p-value or p
  - the probability to obtain a result more extreme than the result obtained from the observed data.
  - If $p < \alpha$, then we will reject the null hypothesis: we demonstrate that there is difference between the arithmetic means and say that the test is significant: we have difference between the means

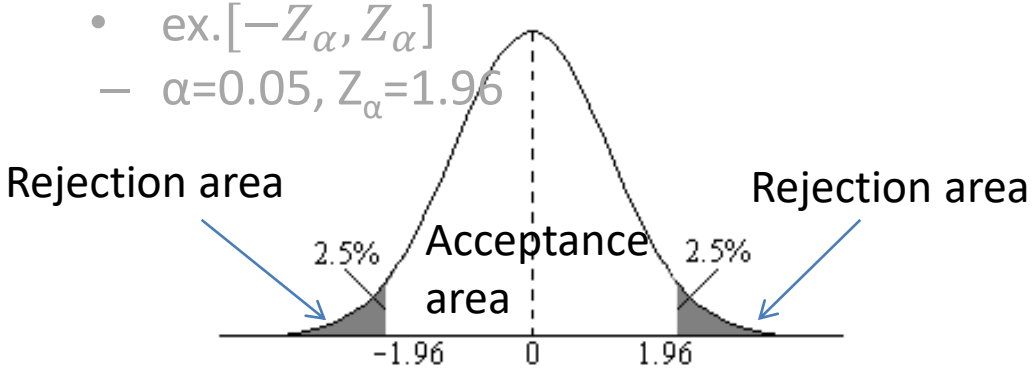
Example:

- we choose $\alpha = 0.05$

Step 4. Determine the value the test parameter must attain to be declared significant

General:

- **critical value**
  - for Z test Z_α
- Rejection area – critical interval
  - different depending on the distribution of the test and sometimes degrees of freedom
  - interval where H_0 should be rejected
  - ex. $(-\infty, -Z_\alpha) \cup (Z_\alpha, \infty)$
- Acceptance interval:
  - interval where H_0 should be rejected
  - ex.$[-Z_\alpha, Z_\alpha]$
  - α=0.05, Z_α=1.96



Example:

t test – Student distribution – different for each degree of freedom

- t_α = 1.97 = **critical value for 100+100-2 degree of freedom**

- rejection interval $(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ = $(-\infty, -1.97) \cup (1.97, \infty)$

- acceptance interval: $[-t_\alpha, t_\alpha]$= [-1.97,1.97]

Step 5: computing the formula for test parameter

General:

any statistical test have a test parameter based on the difference between what we test

test two means, then $\bar{X}_1 - \bar{X}_2$ should be in the formula, divided by standard errors

Example:

- the parameter of the test:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

diet C: $\bar{X}_1$ = 250mg/dl, s = 50mg/dl
diet B: $\bar{X}_2$ = 290mg/dl, s = 100mg/dl

$$t = \frac{250-290}{\sqrt{\frac{50^2}{100}+\frac{100^2}{100}}} = -\frac{40}{\sqrt{125}} = -\frac{40}{11.18} = 3.58,$$

p=0.0004

Step 6. Draw and state the conclusion

General:

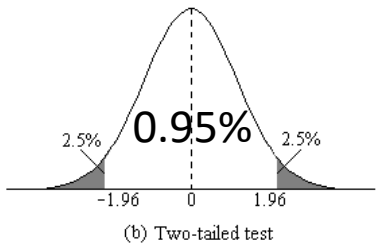
If parameter of the test ∈ acceptance interval --> we fail to reject null hypothesis H_0

If parameter of the test ∈ rejection interval --> we reject null hypothesis H_0, accept H_1

the same as

If p≥0.05 we fail to reject null hypothesis H_0

If p<0.05 we reject null hypothesis H_0, accept H_1



Example:

if t ∈$[-t_\alpha, t_\alpha]$ – acceptance interval =small differences.

if t ∈$(-\infty, -t_\alpha) \cup (t_\alpha, \infty)$ – rejection interval =high differencies.

the same as

If p≥0.05, than H_0 is in favor – fail to reject H_0

If p<0.05 than H_0 is false, we reject H_0 – accept H_1

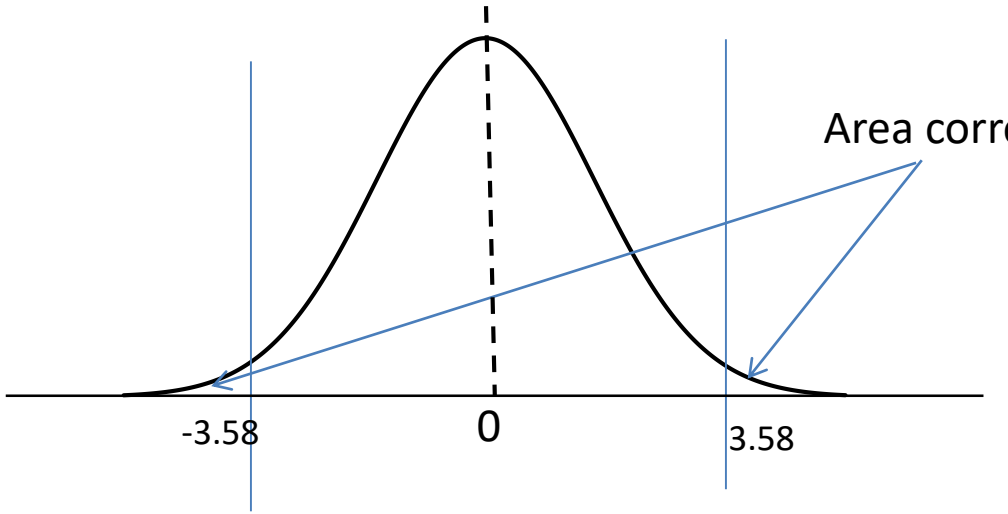
Step 6. Draw and state the conclusion

Example:

$3.58 \in (-\infty, -1.97) \cup (1.97, \infty)$ Rejection area

We reject the null hypothesis H_0, we accept the alternative hypothesis H_1: The average cholesterol of subjects with diet C differ statistically significant from the average cholesterol of subjects with diet B

Area under the curve for t = 3.58 is 0.0004, p = 0.0004<0.05 we reject the null hypothesis H_0, we accept the alternative hypothesis H_1: The average cholesterol of subjects with diet C differ statistically significant from the average cholesterol of subjects with diet B



Area corresponding to $(-\infty, -3.58] \cup [3.58, \infty)$ =0.0004

Randomized Controlled Trial › Circulation. 2018 Mar 13;137(11):1103-1113.
doi: 10.1161/CIRCULATIONAHA.117.030088. Epub 2018 Feb 26.

Low-Calorie Vegetarian Versus Mediterranean Diets for Reducing Body Weight and Improving Cardiovascular Risk Profile: CARDIVEG Study (Cardiovascular Prevention With Vegetarian Diet)

Francesco Sofi [1 2 3], Monica Dinu [4], Giuditta Pagliai [4], Francesca Cesari [5], Anna Maria Gori [4 6], Alice Sereni [4 6], Matteo Becatti [7], Claudia Fiorillo [7], Rossella Marcucci [4 6], Alessandro Casini [4 2]

Affiliations + expand

PMID: 29483085 DOI: 10.1161/CIRCULATIONAHA.117.030088

Abstract

Background: Only a few randomized dietary intervention studies that investigated the effects of lacto-ovo vegetarian diet (Vd) in clinically healthy omnivorous subjects are available.

Methods: We randomly assigned to overweight omnivores with a low-to-moderate cardiovascular risk profile a low-calorie Vd compared with a low-calorie Mediterranean diet (MD), each lasting 3 months, with a crossover design. The primary outcome was the difference in body weight, body mass index, and fat mass changes between the 2 groups. Secondary outcomes were differences in circulating cardiovascular disease risk parameters changes between the 2 groups.

Results: One hundred eighteen subjects (mean age: 51.1 years, females: 78%) were enrolled. The total participation rate at the end of the study was 84.7%. No differences between the 2 diets in body weight were observed, as reported by similar and significant reductions obtained by both Vd (-1.88 kg) and MD (-1.77 kg). Similar results were observed for body mass index and fat mass. In contrast, significant differences between the 2 interventions were obtained for low-density lipoprotein cholesterol, triglycerides, and vitamin B_{12} levels. The difference between the Vd and MD groups, in terms of end-of-diet values, was recorded at 9.10 mg/dL for low-density lipoprotein cholesterol (P=0.01), 12.70 mg/dL for triglycerides (P<0.01), and 32.32 pg/mL for vitamin B_{12} (P<0.01). Finally, no significant difference was found between Vd and MD interventions in oxidative stress markers and inflammatory cytokines, except for interleukin-17, which improved only in the MD group. Forty-six participants during the Vd period and 35 during the MD period reached the target values for ≥1 cardiovascular risk factor.

Conclusions: Both Vd and MD were effective in reducing body weight, body mass index, and fat mass, with no significant differences between them. However, Vd was more effective in reducing low-density lipoprotein cholesterol levels, whereas MD led to a greater reduction in triglyceride levels.

Clinical trial registration: URL: https://www.clinicaltrials.gov. Unique identifier: NCT02641834.

- Italy
- vegetarian versus Mediterranean diet in obese people
- 118 subjects were randomly assigning in groups
- 3 months
- after 3 months change the diets (cross-over design)
- trial was registrated

Randomized Controlled Trial › Circulation. 2018 Mar 13;137(11):1103-1113.
doi: 10.1161/CIRCULATIONAHA.117.030088. Epub 2018 Feb 26.

Low-Calorie Vegetarian Versus Mediterranean Diets for Reducing Body Weight and Improving Cardiovascular Risk Profile: CARDIVEG Study

Table 1. Baseline Characteristics of the Study Population According to the First Randomization

Characteristic	All (n=118)	Vegetarian Diet (n=60)	Mediterranean Diet (n=58)	*P* Value
Total cholesterol, mg/d (mean±SD)	202.7±109.2	198.8±94.5	206.8±124.9	0.96

index, and fat mass changes between the 2 groups. Secondary outcomes were differences in circulating cardiovascular disease risk parameters changes between the 2 groups.

Results: One hundred eighteen subjects (mean age: 51.1 years, females: 78%) were enrolled. The total participation rate at the end of the study was 84.7%. No differences between the 2 diets in body weight were observed, as reported by similar and significant reductions obtained by both Vd (-1.88 kg) and MD (-1.77 kg). Similar results were observed for body mass index and fat mass. In contrast, significant differences between the 2 interventions were obtained for low-density lipoprotein cholesterol, triglycerides, and vitamin B_{12} levels. The difference between the Vd and MD groups, in terms of end-of-diet values, was recorded at 9.10 mg/dL for low-density lipoprotein cholesterol (P=0.01), 12.70 mg/dL for triglycerides (P<0.01), and 32.32 pg/mL for vitamin B_{12} (P<0.01). Finally, no significant difference was found between Vd and MD interventions in oxidative stress markers and inflammatory cytokines, except for interleukin-17, which improved only in the MD group. Forty-six participants during the Vd period and 35 during the MD period reached the target values for ≥1 cardiovascular risk factor.

Conclusions: Both Vd and MD were effective in reducing body weight, body mass index, and fat mass, with no significant differences between them. However, Vd was more effective in reducing low-density lipoprotein cholesterol levels, whereas MD led to a greater reduction in triglyceride levels.

Clinical trial registration: URL: https://www.clinicaltrials.gov. Unique identifier: NCT02641834.

- 198.8 compare with 206.8
- p=0.96 >0.05 we cannot reject H0: there are no significant difference between baseline cholesterol between groups (i.e. in the population from where they select this samples)

	Vegetarian Diet: Before (n=104)	Vegetarian Diet: After (n=104)	Mediterranean Diet: Before (n=103)	Mediterranean Diet: After (n=103)	P (Δ_{VD} VersusΔ_{MD})[†]
TC, mg/dL	207.89 (200.74–215.29)	202.55 (195.98–209.56)[*]	205.41 (197.95–212.94)	205.30 (198.34–212.72)	0.15

- they compare the difference (delta) between baseline and final cholesterol (TC=total cholesterol) between groups
- p=0.15 >0.05 we cannot reject H0: there are no significant difference between the decrease of the cholesterol between vegetarian and Mediterranean diet (i.e. there is no difference in cholesterol decrease between the population which adopt 3 months period a vegetarian and a population who adopt a Mediterranean diet)

A Mediterranean Diet and Low-Fat Vegan Diet to Improve Body Weight and Cardiometabolic Risk Factors: A Randomized, Cross-over Trial

Neal D. Barnard, Jihad Alwarith, Emilie Rembert, Liz Brandon, Minh Nguyen, Andrea Goergen, ...show all

Pages 127-139 | Received 23 Nov 2020, Accepted 22 Dec 2020, Published online: 05 Feb 2021

Cite this article https://doi.org/10.1080/07315724.2020.1869625 Check for updates

Full Article Figures & data References Supplemental Citations Metrics Licensing Reprints & Permission

View PDF View EPUB

Abstract

Objective: Evidence suggests that both Mediterranean and vegan diets improve body weight and cardiometabolic risk factors, but their relative efficacy has not been compared in a randomized trial.

Methods: In a randomized crossover trial, 62 overweight adults were randomly assigned to a Mediterranean or vegan diet for a 16-week period. Body weight, plasma lipids, blood pressure, and body composition (dual X-ray absorptiometry) were measured. Secondary measures included insulin resistance (Homeostasis Model Assessment, HOMA-IR), oral glucose insulin sensitivity (OGIS), and predicted insulin sensitivity (PREDIM) indices. Thereafter, participants were asked to return to their baseline diets for 4 weeks, after which they began the opposite diet for 16 weeks. The same parameters were measured before and after this 2nd 16-week period.

Results: Overall net weight changes were 0.0 (Mediterranean) and −6.0 kg (vegan), (treatment effect −6.0 kg [95% CI −7.5 to −4.5]; $p < 0.001$). HOMA-IR decreased and OGIS increased on the vegan diet with no significant change on the Mediterranean diet (treatment effect −0.7 [95% CI, −1.8 to +0.4]; $p = 0.21$; and +35.8 mL/min/m^2 [95% CI, +13.2 to +58.3]; $p = 0.003$, respectively). PREDIM did not change significantly in either group. Among participants with no medication changes, total and LDL-cholesterol decreased 18.7 mg/dL (0.5 mmol/L) and 15.3 mg/dL (0.4 mmol/L), respectively, on the vegan diet, compared with no significant change on the Mediterranean diet (treatment effect −15.6 [-24.6 to −6.6]; $p = 0.001$ and −14.8 [-23.5 to −6.2]; $p = 0.001$, respectively); systolic and diastolic blood pressure decreased 9.3 and 7.3 mmHg on the Mediterranean diet, compared with 3.4 and 4.1 mmHg on the vegan diet (treatment effect +5.9 [95% CI +1.0 to +10.9]; $p = 0.02$; and +1.8 [95% CI −4.6 to +8.1]; $p = 0.58$, respectively).

Conclusions: A low-fat vegan diet improved body weight, lipid concentrations, and insulin sensitivity, both from baseline and compared with a Mediterranean diet. Blood pressure decreased on both diets, more on the Mediterranean diet.

Clinical trial registration: ClinicalTrials.gov number, NCT03698955

Related rese

People also read

Vegetarian, vega outcomes: A syst observational stu

Monica Dinu et a Critical Reviews in Published online:

Evidence of a veg – an umbrella re observational an

Eliška Selinger e Critical Reviews in Published online:

Mediterranean d mortality in diab analysis of prosp randomized clini

Nerea Becerra-T Critical Reviews in Published online:

- USA
- 62 obese
- randomly assign to Mediteranean or Vegan diet
- trial was registrated

A Mediterranean Diet and Low-Fat Vegan Diet to Improve Body Weigh Cross-

Neal D. Barnar

Cite this artic

Full Article

View PDF

Abstract

Objective: E cardiometab

Methods: In Mediterrane composition (Homeostasi sensitivity (P after which t after this 2nd

Results: Ov [95% CI −7.5 significant ch +35.8 mL/mi either group mg/dL (0.5 n significant ch [-23.5 to −6. the Mediterr +1.0 to +10.9

Conclusions from baselin the Mediterr

Methods: In a randomized crossover trial, 62 overweight adults were randomly assigned to a Mediterranean or vegan diet for a 16-week period. Body weight, plasma lipids, blood pressure, and body composition (dual X-ray absorptiometry) were measured. Secondary measures included insulin resistance (Homeostasis Model Assessment, HOMA-IR), oral glucose insulin sensitivity (OGIS), and predicted insulin sensitivity (PREDIM) indices. Thereafter, participants were asked to return to their baseline diets for 4 weeks, after which they began the opposite diet for 16 weeks. The same parameters were measured before and

- delta cholesterol -15.6mg/dl
- p=0.001<0.05 we reject H0: there was significant difference of mean cholesterol between vegan and mediteranean diet (i.e.there is difference in cholesterol decrease between the population which adopt 3 weeks period a vegetarian and a population who adopt a mediteranean diet)

effect −6.0 kg

n no

.21; and

ificantly in

either group. Among participants with no medication changes, total and LDL-cholesterol decreased 18.7 mg/dL (0.5 mmol/L) and 15.3 mg/dL (0.4 mmol/L), respectively, on the vegan diet, compared with no significant change on the Mediterranean diet (treatment effect −15.6 [-24.6 to −6.6]; $p = 0.001$ and −14.8 [-23.5 to −6.2]; $p = 0.001$, respectively); systolic and diastolic blood pressure decreased 9.3 and 7.3 mmHg on the Mediterranean diet, compared with 3.4 and 4.1 mmHg on the vegan diet (treatment effect +5.9 [95% CI +1.0 to +10.9]; $p = 0.02$; and +1.8 [95% CI −4.6 to +8.1]; $p = 0.58$, respectively).

Clinical trial registration: ClinicalTrials.gov number, NCT03698955

A Mediterranean Diet and Low-Fat Vegan Diet to Improve Body Weigh Cross-

Neal D. Barnar

Cite this artic

Full Article

View PDF

Abstract

Objective: E cardiometab

Methods: In Mediterrane composition (Homeostasi sensitivity (P after which t after this 2nd

Results: Ov [95% CI −7.5 significant ch +35.8 mL/mi either group mg/dL (0.5 n significant ch [-23.5 to −6.2 the Mediterr +1.0 to +10.9

Conclusions from baselin the Mediterr

Methods: In a randomized crossover trial, 62 overweight adults were randomly assigned to a Mediterranean or vegan diet for a 16-week period. Body weight, plasma lipids, blood pressure, and body composition (dual X-ray absorptiometry) were measured. Secondary measures included insulin resistance (Homeostasis Model Assessment, HOMA-IR), oral glucose insulin sensitivity (OGIS), and predicted insulin sensitivity (PREDIM) indices. Thereafter, participants were asked to return to their baseline diets for 4 weeks, after which they began the opposite diet for 16 weeks. The same parameters were measured before and after this 2nd 16-week period.

Results: Overall net weight changes were 0.0 (Mediterranean) and −6.0 kg (vegan), (treatment effect −6.0 kg [95% CI −7.5 to −4.5]; $p < 0.001$). HOMA-IR decreased and OGIS increased on the vegan diet with no significant change on the Mediterranean diet (treatment effect −0.7 [95% CI, −1.8 to +0.4]; $p = 0.21$; and +35.8 mL/min/m^2 [95% CI, +13.2 to +58.3]; $p = 0.003$, respectively). PREDIM did not change significantly in either group. Among participants with no medication changes, total and LDL-cholesterol decreased 18.7 mg/dL (0.5 mmol/L) and 15.3 mg/dL (0.4 mmol/L), respectively, on the vegan diet, compared with no significant change on the Mediterranean diet (treatment effect −15.6 [-24.6 to −6.6]; $p = 0.001$ and −14.8 [-23.5 to −6.2]; $p = 0.001$, respectively); systolic and diastolic blood pressure decreased 9.3 and 7.3 mmHg on the Mediterranean diet, compared with 3.4 and 4.1 mmHg on the vegan diet (treatment effect +5.9 [95% CI +1.0 to +10.9]; $p = 0.02$; and +1.8 [95% CI −4.6 to +8.1]; p

95% confidence interval between -24.6 to -6.6

vegan diet versus mediteranean diet in population best: -24.6mg/dl, worst -6.6mg/dl

Clinical trial registration: ClinicalTrials.gov number, NCT03698955

The OMNIVEG STUDY: Health outcomes of shifting from a traditional to a vegan Mediterranean diet in healthy men. A controlled crossover trial

Miguel López-Moreno [a,b] · Ujué Fresán [c] · Juan Del Coso [d] · ... · Javier Pena-Fernández [e] · Alejandro Muñoz [e] · Jorge Gutiérrez-Hellín [e]
... Show more

Affiliations & Notes · Article Info

Abstract

Background and aim

The Mediterranean diet is a plant-based dietary pattern with well-established health benefits such as the reduced risk of cardiovascular disease. Additionally, incorporating more plant-based foods into a Mediterranean diet may provide further health benefits. The study aimed to assess the effect of shifting from a traditional Mediterranean diet to a vegan Mediterranean diet on cardiorespiratory fitness and lipid profile in physically active and healthy men.

Methods and Results

Participants underwent a baseline period with adhesion to the general patterns of the Mediterranean diet for three weeks and then they changed to an isocaloric vegan version of the Mediterranean diet for four weeks, with a 7-day washout period between diets. The shift from the traditional Mediterranean diet to the vegan Mediterranean diet required substituting animal-based foods with plant-based foods that contain comparable amounts of protein and fat. Fourteen participants with a mean age of 24.6 ± 7.0 years (range: 18–37 years), completed the study protocol. The change from the traditional to the vegan Mediterranean diet reduced blood concentration of total cholesterol (−22.6 mg/dl, $p < 0.01$, Effect size [ES] = 1.07) and low-density lipoprotein cholesterol (−12.8 mg/dl, $p < 0.01$, ES = 0.72). An inverse correlation was observed between the intake of dietary fibre and LDL-C (partial rho = −0.43, $p = 0.040$).

Conclusions

The adoption of a vegan Mediterranean diet with plant-based proteins and fats instead of the traditional Mediterranean diet improved several cardiometabolic health outcomes in physically active and healthy men.

Clinical Trial Registry

NCT06008886.

- Spain
- vegetarian versus Mediterranean diet
- 14 subjects
- 3 weeks
- after 3 weeks change the diets (cross-over design)
- trial was registrated

The OMNIVEG STUDY: Health outcomes of shifting from a traditional to a vegan Mediterranean diet in healthy men. A controlled crossover trial

Miguel López-Moreno [a,b] · Ujué Fresán [c] · Juan Del Coso [d] · ... · Javier Pena-Fernández [e] · Alejandro Muñoz [e] · Jorge Gutiérrez-Hellín [e]
... Show more

Affiliations & Notes · Article Info

Abstract

Methods and Results

Participants underwent a baseline period with adhesion to the general patterns of the Mediterranean diet for three weeks and then they changed to an isocaloric vegan version of the Mediterranean diet for four weeks, with a 7-day washout period between diets. The shift from the traditional Mediterranean diet to the vegan Mediterranean diet required substituting animal-based foods with plant-based foods that contain comparable amounts of protein and fat. Fourteen participants with a mean age of 24.6 ± 7.0 years (range: 18–37 years), completed the study protocol. The change from the traditional to the vegan Mediterranean diet reduced blood concentration of total cholesterol (−22.6 mg/dl, $p < 0.01$, Effect size [ES] = 1.07) and low-density lipoprotein cholesterol (−12.8 mg/dl, $p < 0.01$, ES = 0.72). An inverse correlation was observed bet

$p = 0.040$).

Conclusions

The adoption of a vegan Mediterranean diet with plant-based proteins and fats instead of the traditio improved several cardiometabolic health outcomes in physically active and healthy men.

Clinical Trial Registry

NCT06008886.

- delta cholesterol -22.6mg/dl
- p<0.01 we reject H0: there was significant difference of mean cholesterol between vegan and mediteranean diet (i.e.there is difference in cholesterol decrease between the population which adopt 3 weeks period a vegetarian and a population who adopt a mediteranean diet)

Thank you !!!