

Principles of Biostatistics and Informatics

Lecture 14th:
Evidence based medicine. Clinical studies.
2018. December 12.
Veres Dániel

Evidence Based Medicine (EBM)

A series of conscientious, unambiguous and logical decisions based on the evidence currently available that serves the patient's personalized treatment.

Known results of „statistical trials”
Personal knowledge, clinical practice
„Common sense”

There is no exact definition for evidence based medicine (abbreviated as EBM), but there are several attempt to clarify it. I prefer the next one: A series of conscientious, unambiguous and logical decisions based on the evidence currently available that serves the patient's personalized treatment. I'd like to highlight two phrase:
1. Evidence currently available – where it is come from?
a) Based on known statistical trials
b) Based on personal knowledge, practice
2. Logical – it means here: don't forget to use your mind („common sense”)

In this lecture I tried to write everything into the slides, therefore you can see less notes.

Biostatistics – why to learn?

- „To decide whether we should believe in something we are reading or to see where the mistake is, that is to say, do not fall so easily into statistical „juggling”, artifacts and mistakes. (see excel – panacea,...)
- „To judge better whether we were lucky or not – or none of them....”
- „To judge better what is worth , whether it is worth for risking it...”
- „So that we can do our best to design and evaluate our own statistics in our work (diploma...).”
- „I got an interested, unexpected result? I just discovered something or just the game of chance I see?”
- „To make our results more understandable and effective, we can highlight the essence. ”
- „To have a clear understanding of the literature. ”
(J. Reiczigel)

Why to learn Biostatistics?
In general and in your medical studies.

Known results of „statistical trials” – how they will be?

- We collecting data and analysing them.

How the results of „statistical trials” created.
Collecting data then analysing?

Known results of „statistical trials” – how they will be?

- ▲ ~~We collecting data and analysing them:~~
- HOW TO GET DATA? – first PLAN it!

Considerational considerations :

- What is the aim, the question?
- What „mistakes” should be considered?
- How much should the samplesize be?
- What methods can be applied?
- Which sampling techniques are available?
- ... So THEN collect data...

NO!!! First plan – then analysing!!

Known results of „statistical trials” – how they will be?

- ▲ ~~We deal with existing, existing data that we have collected:~~
- HOW TO GET DATA?

Considerational considerations :

- What is the aim, the question?
- What „mistakes” should be considered?
- What methods can be applied?
- Which sampling techniques are available?
- How much should the sample size be?
- ... So THEN collect data...

- **The most sophisticated, most accurate data analysis does not compensate for a poorly planned, designed or executed data collection or survey !!!!!!!**

PLANNING!

What is the aim?

Is there any difference?
Is there any relation (correlation)?
Is there any effect?
...

First look what can be a typical question.

Aim + Relevant?

Is there any difference?
Is there any relation (correlation)?
Is there any effect?
...

There may be a „difference” – but is it relevant (clinically important)?
Based on clinical practice – that is NOT A STATISTICAL
QUESTION – but it is really IMPORTANT

– How much is the difference...? EFFECT SIZE
...

If there will be a difference – is it important? How much
have to be the effect that can be transformed to clinical
benefit? This quantity is the effect size in general. This
„size” is high enough? –that is NOT A STATISTICAL
QUESTION – but it is really IMPORTANT, it is based on
clinical practice.

Effect size

There could be a difference, but is it relevant?

– How much is the „difference“? : How to express it?

Difference between means, medians; ratio of means, medians

How much (how many times) does it change in another group?

Correlation, determination coefficient

To what extent does the change in y affected by the change in x?
slope

If the x (independent) variable increased by 1 unit what will be the average change in the y (dependent) variable.

Odds ratio, Risk ratio

How many times does the odds or risk increased if the risk factor present?

...

What could be an effect size – how to express it?

Error 1 – why? Significance

Problem: we couldn't examine everybody (the population)!

Solution: sampling – but...

...sampling error (chance)... Hypothesis testing!

The observed effect could be by chance?

Is the „difference“ significant?

(*see excel, radiation)

We have learned before the main problematic in statistics: we couldn't measure the population therefore we sampling it. But in the observed „difference“ may be because of chance.

To calculate „this chance“ we perform hypothesis tests.

See example in excel (radiation tab)

Effect size and „significance“ together

Confidence interval!

(NOTE: learn this carefully – we like to ask it in the exam!)

Is there a way to express observed effect size and significance together? YES. With using confidence intervals.
See example in excel: radiation tab solution, confidence interval 2.

If there is a difference can we recognize it?

| | | In population (in reality) the null hypothesis is: | |
|------------------------------|---------------------------|---|---|
| | | True | False |
| Decision on null hypothesis: | Accepting (Not rejecting) | Good decision | Error (type II) (β) (false negative result) |
| | Rejecting | Error (type I) (α) (false positive result) | Good decision power(1-β) |

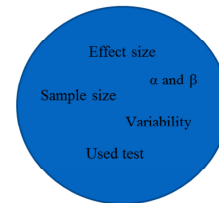
(NOTE: learn this strictly – we like to ask it in the exam!)

Power.

Power

Is there a difference can we recognize it?
it can be recognized easier, if:
higher sample size
high effect size
(may be others: small SD...)
high power hypothesis test:
higher scale!!!
normal distribution
„effectively“ paired

Power ($1-\beta$)



They depend on each other.

Relevant, but not significant...

Reasons:

small power:
small sample size (limitation: money, ethical issues)*
large variability
less powerfull statistical test

we could not measure it accurately
violated assumptions for the test

we were unlucky (sampling error)

other errors

Plan ahead!!

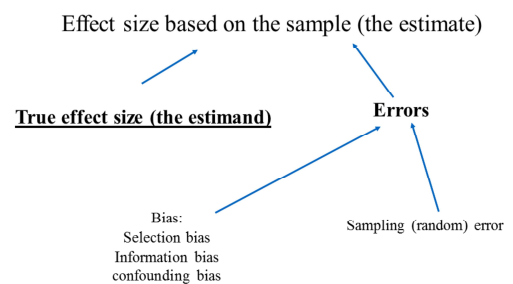
-*Ask yor statisticians...

(© eg: <https://www.youtube.com/watch?v=PbODigCZqL8>)

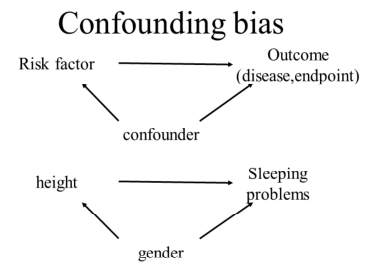
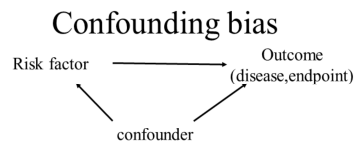
Sample size video:

<https://www.youtube.com/watch?v=Hz1fyhVOjr4>

Other errors



(NOTE: learn this strictly – we like to ask it in the exam!)



Most common confounders: gender, age – always think about them!

(NOTE: learn this strictly – we like to ask it in the exam!)

Selection bias, Information bias

Selection bias:

There is a difference between the selected and not selected individuals, or difference between assignment to groups (erroneous selection with respect to an outcome influencing parameter)

typical: age, gender different in the groups
different population
different follow-up time

Information bias:

erroneous data collection about or from subjects (which affects the outcome)

typical: recall bias
more careful monitoring for diseased, young

(NOTE: learn this strictly – we like to ask it in the exam!)

Selection bias:

AGE: Knowing that gender has an effect on sleeping problems, the proportion of women in a „new drug group“ is different than in the „placebo group“.

LOSS OF FOLLOW UP: – we know that AIDS is more frequent in i.v. drug users and homosexuals: those who has AIDS will more often "quit" the follow-up, than those who do not get AIDS, as well as IDU users more often „disappear“, than homosexuals

DIFFERENT POPULATION: Fracture in Women and Nutrition Relationships: We choose bone trauma from a trauma class, control of the hospital's internal medicine (But there are other more frequent illnesses in the internal medicine, eg diabetes is more common that has an effect on fracture !!)

Information bias:

RECALL BIAS: Parents of children diagnosed with cancer may be more likely to recall infections earlier in the child's

life than parents of children without cancer.

Good source:

Catalogue of Bias Collaboration, Spencer EA, Brassey J, Mahtani K., 2017. <https://catalogofbias.org/>

Which methods?

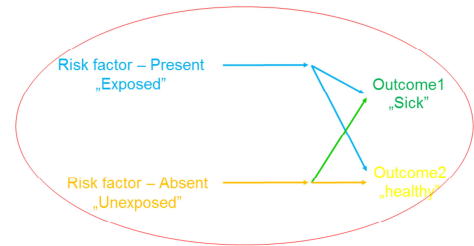
Main type of Clinical Studies (Study Designs):

Observation: no intervention, just observation
 Cross-sectional studies – at a given time
 Case-control studies – pro-/retrospective
 Cohort – pro-/retrospective

Experimental: intervention („treating”)
 main type: randomized controlled and clinical trials

(NOTE: learn this strictly – we like to ask it in the exam!)

Cross-sectional



Risk factor and outcomes observed at the same time

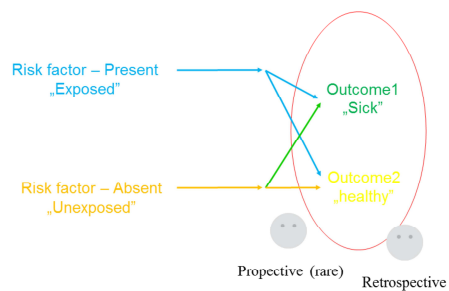
Risk factor and outcomes observed at the same time

Eg.

Family head reads a newspaper - knows that HIV can spread from mother to child.
 Could NOT examine causality!

Case-control

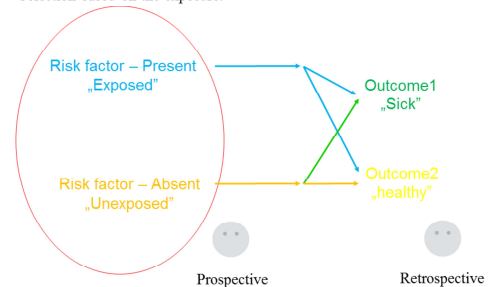
Selection based on the outcomes.



Selection based on the outcome (presence or absence of disease).
 Effective, if the disease is rare!

Cohort

Selection based on the exposure.



Selection based on the risk factor (exposition).
 Effective if the risk factor is rare.

Note

Prevalence - Measures the existence of a disease (or exposure, treatment) at a given point in time (point prevalence)

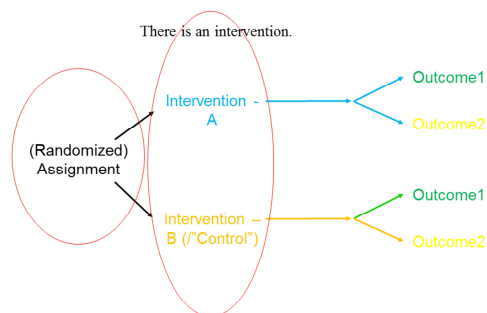
Incidence - Measures occurrence of a disease in a population over a specified period in time (**new cases!**)

Acute disease: prevalence ~ incidence

chronic: prevalence ? incidence

(NOTE: learn this strictly – we like to ask it in the exam!)

Experimental (RCT)



There is an intervention, not only observation.
It creates the possibility for randomization!
With randomization we could reduce measured and UNmeasured(!) biases.

Summary table of study design

| | Cross-sectional | Case-control | Cohort | RCT |
|----------------|---|--|---|---------------------------------------|
| Property | Selection at a given time point | Selection based on: outcome (case/control) | Selection based on: risk factor (exposure) | There is an intervention |
| Advantages | Logistically easier and faster Cheap | Good for rare disease, Logistically easier and faster Less expensive | Good for rare risk factor | Reduced bias |
| Dis-advantages | No causality | Hard to select controls – selection, information bias (eg. recall bias) | Long follow-up time Information bias (eg. recall bias) | Expensive, logistically hard and slow |

(NOTE: learn this strictly – we like to ask it in the exam!)

...

What's the difference between a physicist, a mathematician, and a statistician?

The physicist calculates until he gets a correct result and concludes that he has proven a fact.

The mathematician calculates until he gets a wrong result and concludes that he has proven the contrary of a fact.

The statistician calculates until he gets a correct result about an obviously wrong proposition and concludes NOTHING, because the explanation is the task of the scientist who consulted the statistician.

Source of stat jokes: <http://www.ilstu.edu/~gcramsey/Gallery.html>

Jokes: <http://my.ilstu.edu/~gcramsey/Gallery.html>

+Notes

- Correlation:
 - *GRAPH!!!!*;
 - Correlation NOT equal with causality
 - eg: <http://www.fastcodesign.com/3030529/infographic-of-the-day/hilarious-graphs-prove-that-correlation-isnt-causation>
- Multiplicity
 - eg: *Chocolate Helps Weight Loss*
 - <https://io9.gizmodo.com/i-fooled-millions-into-thinking-chocolate-helps-weight-1707251800>

Correlation:

<https://www.fastcompany.com/3030529/hilarious-graphs-prove-that-correlation-isnt-causation>

Multiplicity: <https://io9.gizmodo.com/i-fooled-millions-into-thinking-chocolate-helps-weight-1707251800>

Non-transitivity



*see excel file

Links:

https://en.wikipedia.org/wiki/Nontransitive_dice

<http://singingbanana.com/dice/article.htm>

<https://plus.maths.org/content/taxonomy/term/789>

Red>blue>olive>red