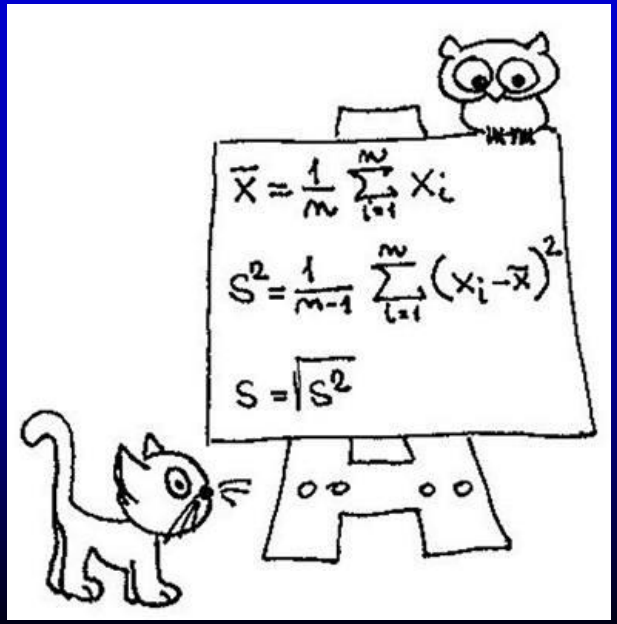
Využití a zneužití statistických metod v medicíně

Martin Hynek

Gennet, Centre for Fetal Medicine, Prague

EuroMISE Centre, First Faculty of Medicine of Charles University in Prague



GENNET

EuroMISE

Statistika

Existují tři druhy lži: lež prostá, lež odsouzeníhodná a statistika.

Benjamin Disraeli (1804 - 1881)

Statistika

Existují tři druhy lži: lež prostá, lež odsouzeníhodná a statistika.

Benjamin Disraeli (1804 - 1881)

Statistika se těší pochybnému vyznamenání tím, že je nejvíce nepochopeným vědním oborem.

H. Levinson

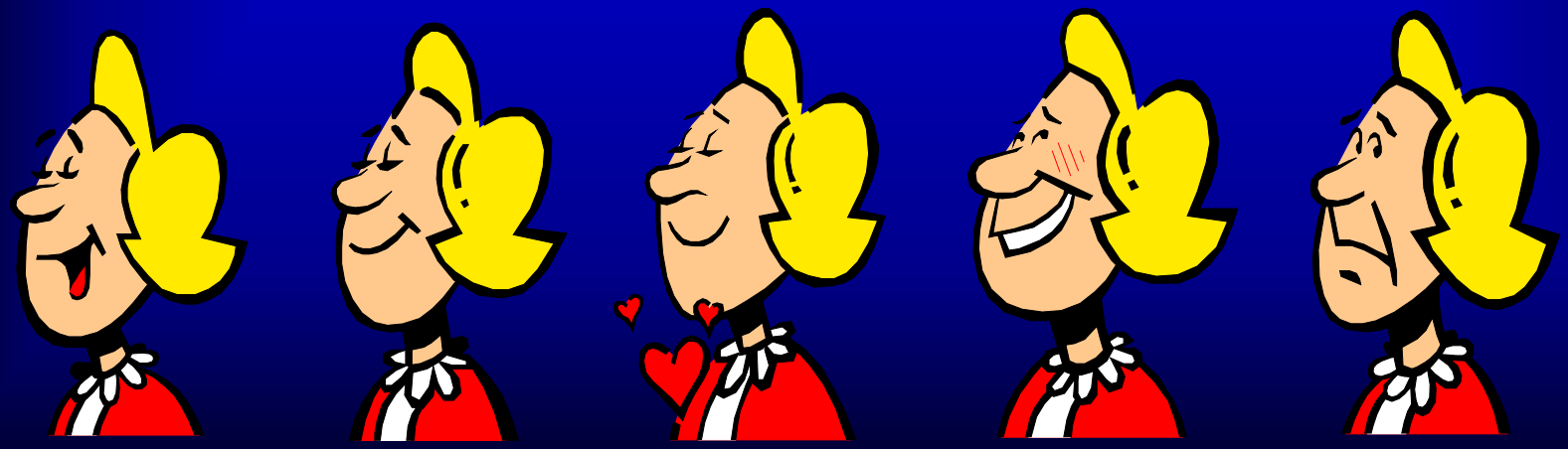
Statistika vs. lékař

Several studies have shown that the statistical understanding by doctors of basic statistical methods and ideas is inadequate. *(Altman and Bland, 1991)*

One of the most important skills a physician should have is the ability to critically analyse original contribution to the medical literature. *(Albert, 1981)*

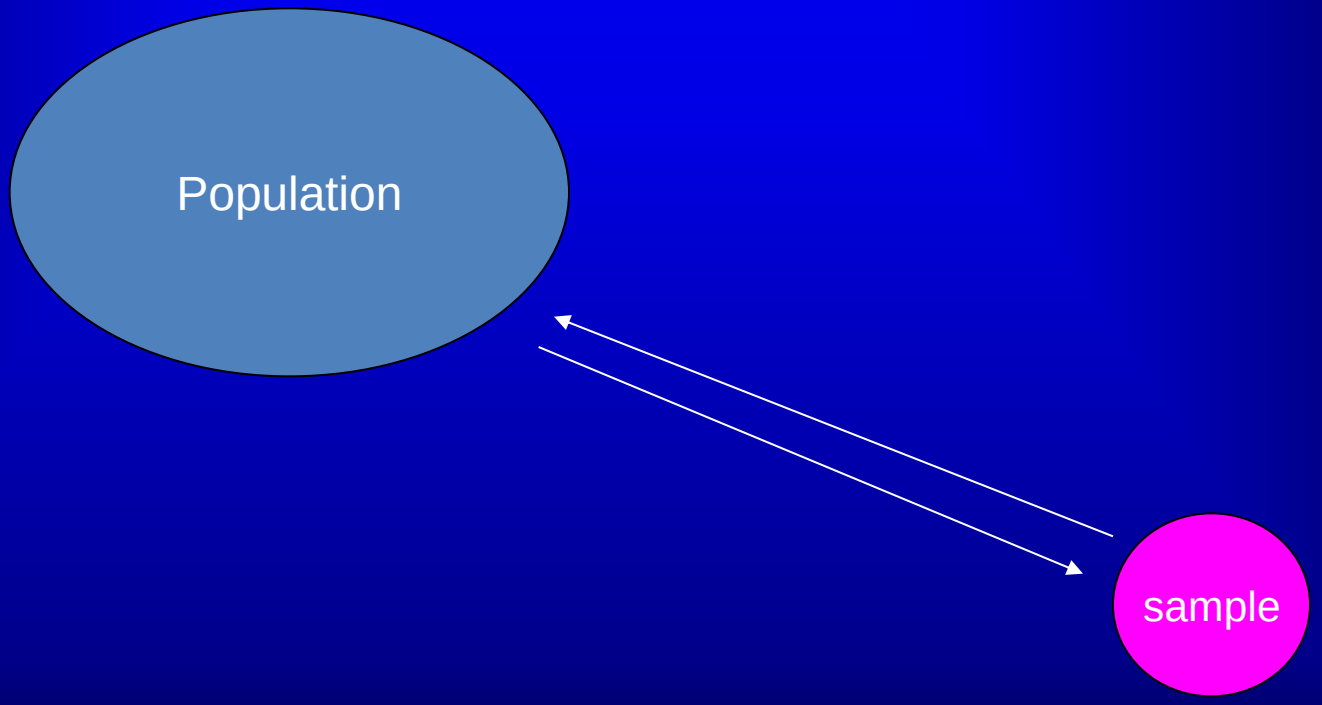
Proč ji v medicíně potřebujeme ?

- Biologická variabilita
  - Jednotlivé atributy se liší jak mezi jedinci, tak u jednoho jedince v průběhu času



Proč ji v medicíně potřebujeme ?

- Sampling



Nemusíte snist celého vola na to, aby jste poznali, že maso je tuhé.
S. Johnson

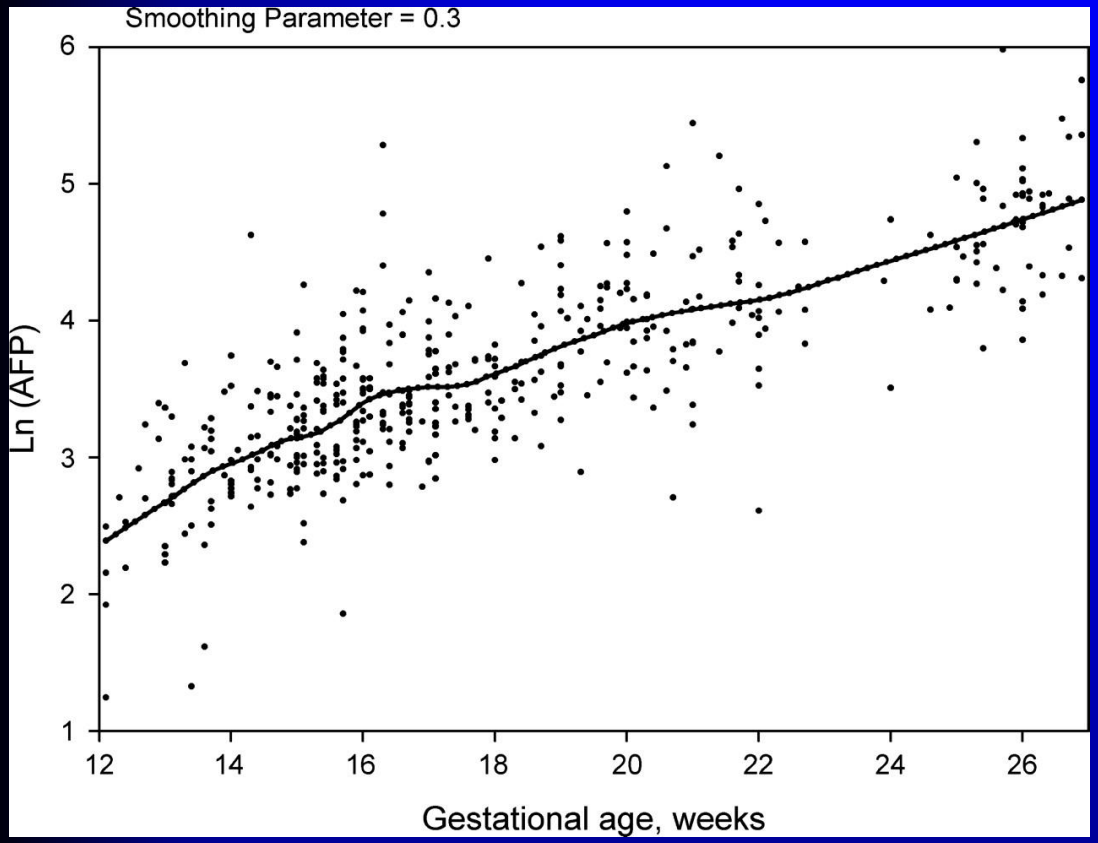
Proč je dobré ji alespoň trochu rozumět ?

- Výzkum
- Odborné publikace
- Běžná praxe

Proč je dobré ji alespoň trochu rozumět ?

- Běžná praxe
  - screening a screenigové algoritmy
  - Z-score a růstové grafy
  - MoM - násobek mediánu

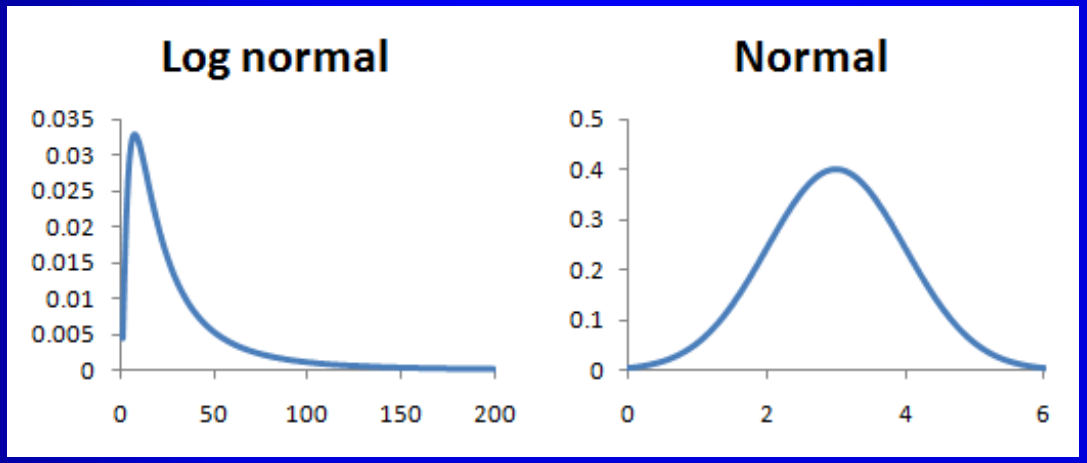
MoM – násobek mediánu

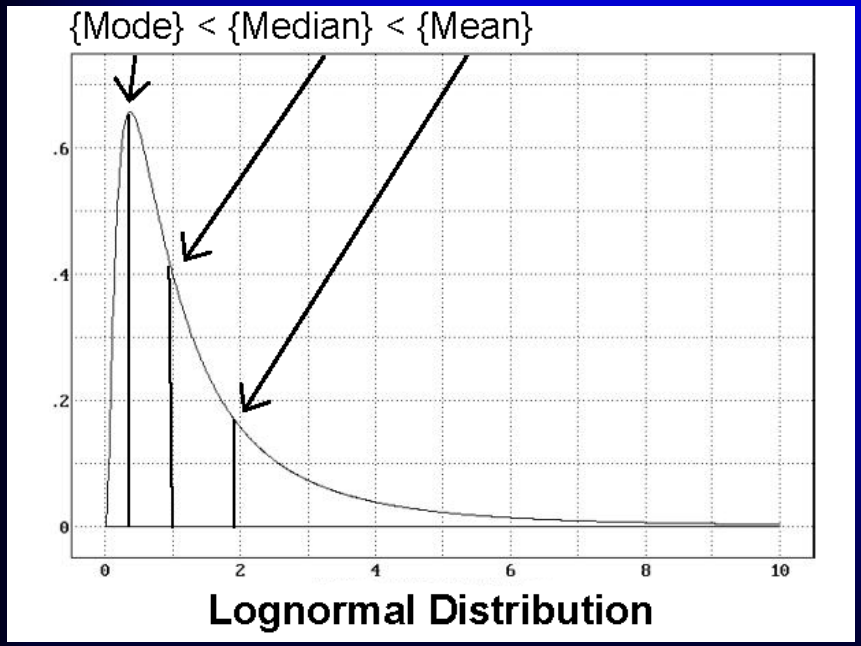


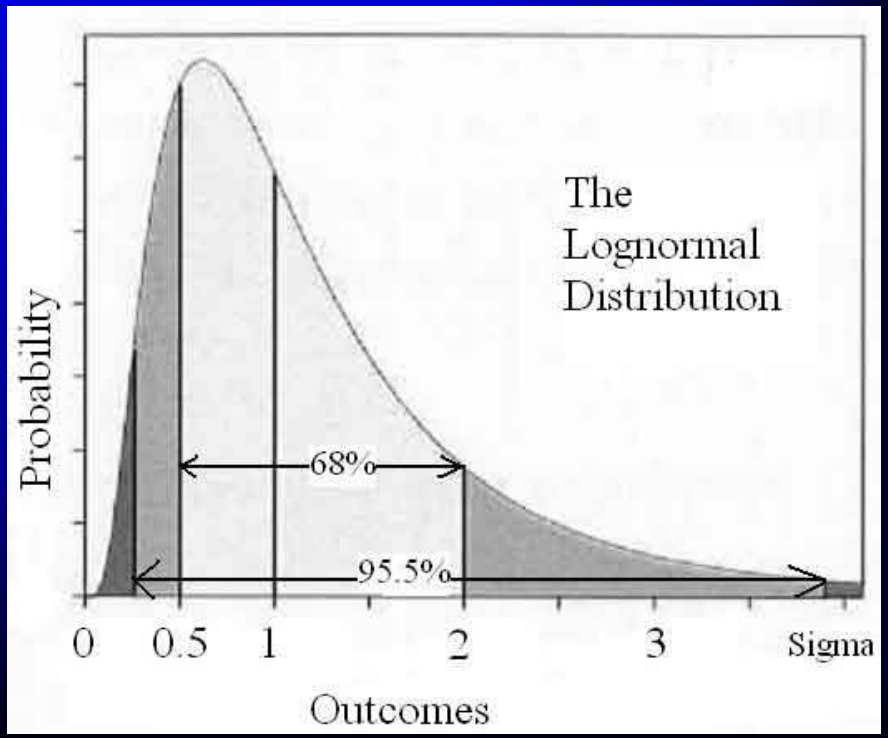
AFP

- Standardizace
- Hodnoty porovnatelné nezávisle na gestačním stáří, mezi laboratořemi, jednoznačný cut-off

MoM – násobek mediánu





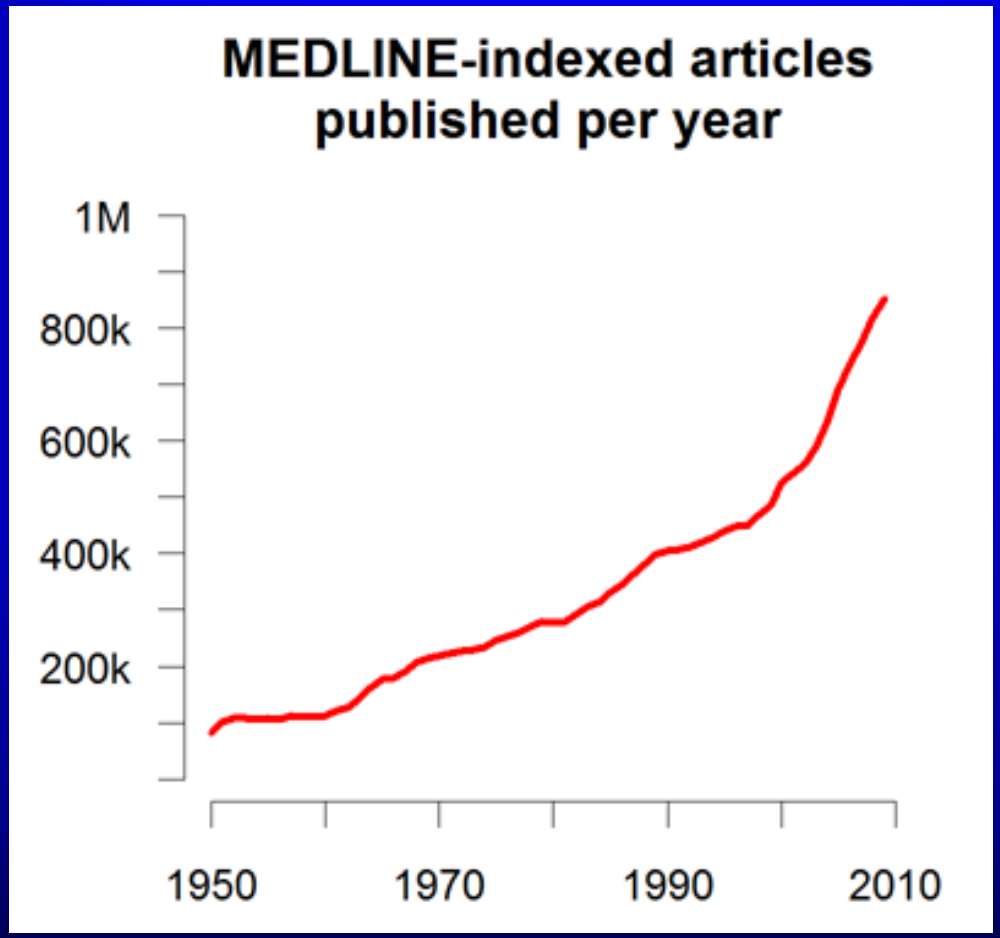


Proč je dobré ji alespoň trochu rozumět ?

- Běžná praxe

Patient: Will I survive this risky operation?
Surgeon: Yes, I'm absolutely sure that you will survive the operation.
Patient: How can you be so sure?
Surgeon: Well, 9 out of 10 patients die in this operation, and yesterday my ninth patient died.

Statistika v lékařských časopisech



Statistika v lékařských časopisech

Table 16.1 Use of statistical procedures in *Pediatrics* (Hayden, 1983)

	Year			
	1952	1962	1972	1982
Number of papers	67	98	115	151
No statistical procedures	66%	59%	45%	30%
Statistical procedures other than t, χ^2 and r	3%	5%	12%	35%

Altman (1991) *Practical Statistics for Medical Research*, Chapman and Hall.

Statistika v lékařských časopisech

?? Kvalita ??

Over several decades there has been considerable evidence of the wide misuse of statistics. (Altman,2000)

- 75% z 59 článků v Transfusion s nevhodným statistickým testem či chybou v interpretaci a 22% článků se závěrem, který nevyplýval z prezentovaných dat (Kanter and Taylor, 1994)
- 40% z 164 článků v British Journal of Psychiatry (McGuigan, 1995)
- Jen 30% článků bylo v pořádku v Amer Journal of Obstet and Gynecol (Welch, 1996)
- Statistické chyby v 38% článků v Nature a v 22% článku v BMJ (Garcia-Berthou and Alcaraz, 2004)

Statistika v lékařských časopisech

Table 16.4 Summary of review of 86 therapeutic trials in perinatal medicine (Tyson *et al.*, 1983)

	% of studies fulfilling criteria		
	Yes	Unclear	No
Statement of purpose	94	6	0
Clearly defined outcome variables	74	1	25
Planned prospective data collection	48	30	22
Predetermined sample size (or a sequential trial)	3	16	71
Sample size specified	93	6	1
Disease/health status of subjects specified ($n = 85$)	51	20	29
Exclusion criteria specified ($n = 81$)	46	9	45
Randomization (if feasible) appropriately performed and documented ($n = 69$)	9	12	79
Blinding used, or lack of blinding unlikely to have biased results ($n = 83$)	49	47	4
Adequate sample size	15	44	41
Statistical methods identified, appropriately used and interpreted	26	0	74
Recommendations/conclusions justified	10	71	19

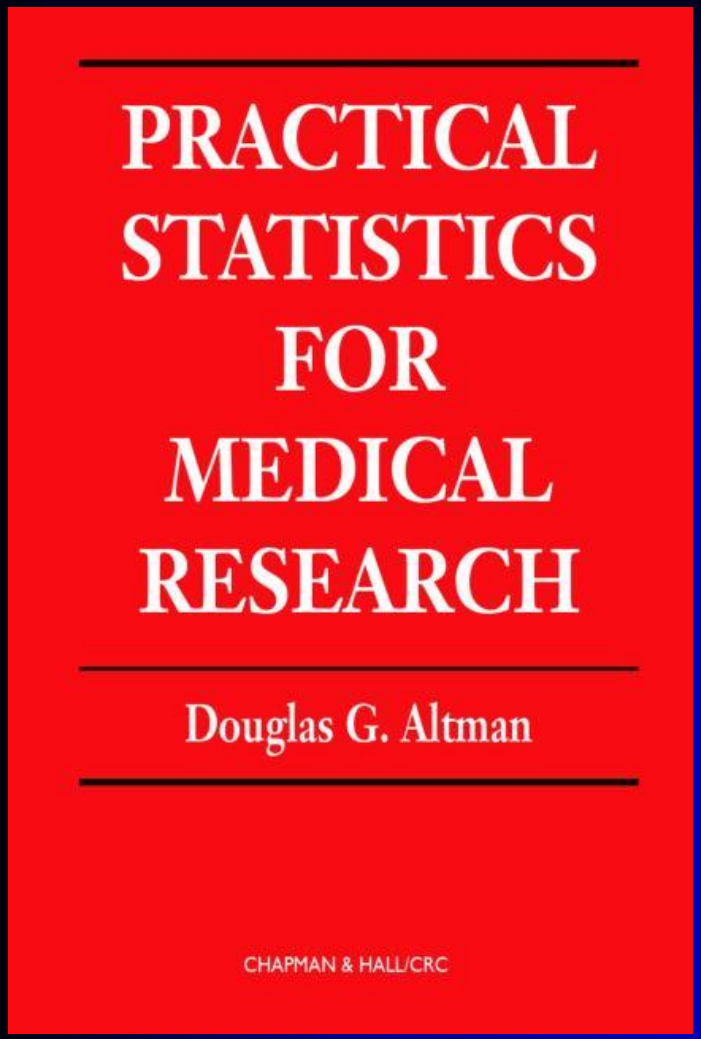
Altman (1991) *Practical Statistics for Medical Research*, Chapman and Hall.

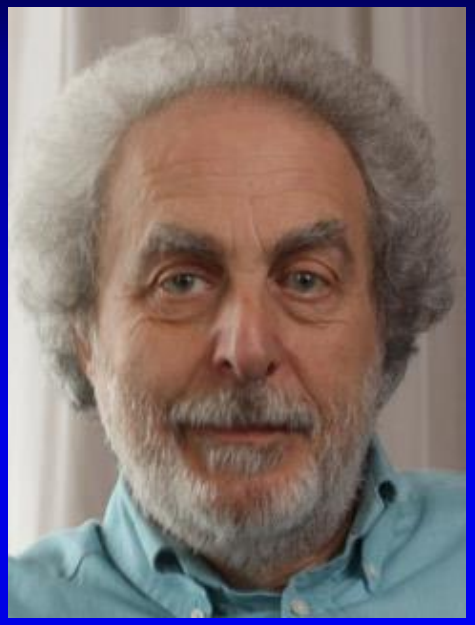
Statistika v lékařských časopisech

Table II. Assessment of the quality of methodology of 364 randomized controlled trials published in 10 leading surgical journals (1988–1994) [88].

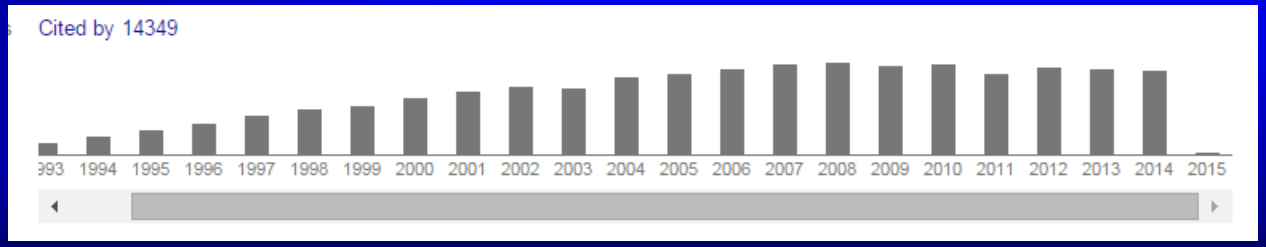
Criterion	Acceptable (%)
Clear description of intervention	94
Adequate control group	93
Inclusion criteria	75
Randomization technique	27
Sample size calculation	19
Definition of endpoint	65
Unbiased outcome assessment	48
Adverse events documented	77

Altman (2000) Stat in Med





Douglas G. Altman

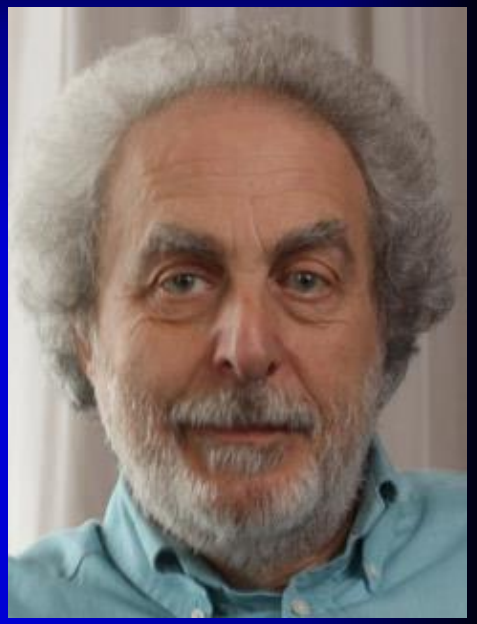


GENNET
www.gennet.cz

Statistics in medical journals: some recent trends

Douglas G. Altman*,†

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, Old Road, Oxford OX3 7LF, U.K.



STATISTICAL REVIEWING FOR MEDICAL JOURNALS

DOUGLAS G. ALTMAN*

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, Old Road, Headington, Oxford OX3 7LF, U.K.

STATISTICS IN MEDICAL JOURNALS: DEVELOPMENTS IN THE 1980s

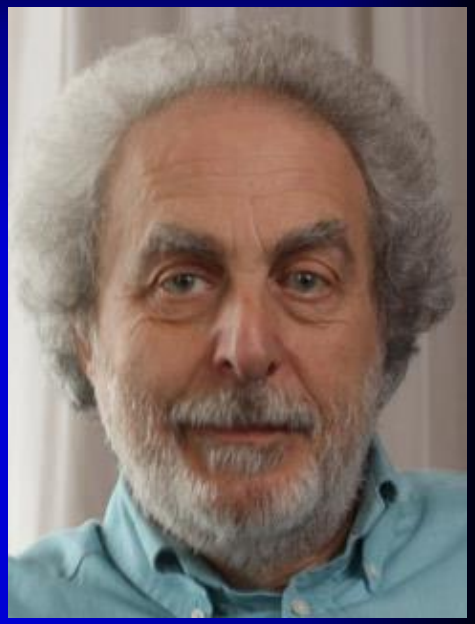
DOUGLAS G. ALTMAN

Medical Statistics Laboratory, Imperial Cancer Research Fund, P.O. Box No 123, Lincoln's Inn Fields, London WC2A 3PX, U.K.

STATISTICS IN MEDICAL JOURNALS*

DOUGLAS G. ALTMAN

MRC Clinical Research Centre, Watford Road, Harrow, Middlesex, U.K.



BMJ

LONDON, SATURDAY 29 JANUARY 1994

The scandal of poor medical research

We need less research, better research, and research done for the right reasons

What should we think about a doctor who uses the wrong treatment, either wilfully or through ignorance, or who uses the right treatment wrongly (such as by giving the wrong dose of a drug)? Most people would agree that such behaviour was unprofessional, arguably unethical, and certainly unacceptable.

What, then, should we think about researchers who use the wrong techniques (either wilfully or in ignorance), use the right techniques wrongly, misinterpret their results, report their results selectively, cite the literature selectively, and draw unjustified conclusions? We should be appalled. Yet numerous studies of the medical literature, in both general and specialist journals, have shown that all of the above phenomena are common.[1-7] This is surely a scandal.

When I tell friends outside medicine that many papers published in medical journals are misleading because of methodological weaknesses they are rightly shocked. Huge sums of money are spent annually on research that is seriously flawed through the use of inappropriate designs, unrepresentative samples, small samples, incorrect methods of analysis, and faulty interpretation. Errors are so varied that a whole book on the topic,[7] valuable as it is, is not comprehensive; in any case, many of those who make the errors are unlikely to read it.

Why are errors so common? Put simply, much poor research arises because researchers feel compelled for career reasons to carry out research that they are ill equipped to perform, and nobody stops them. Regardless of whether a doctor intends to pursue a career in research, he or she is usually expected to carry out some research with the aim of publishing several papers. The length of a list of publications is a dubious indicator of ability to do good research; its relevance to the ability to be a good doctor is even more obscure. A common argument in favour of every doctor doing some research is that it provides useful experience and may help doctors to interpret the published research of others. Carrying out a sensible study, even on a small scale, is indeed useful, but carrying out an ill designed study in ignorance of scientific principles and getting it published surely teaches several undesirable lessons.

In many countries a research ethics committee has to approve all research involving patients. Although the Royal College of Physicians has recommended that scientific criteria are an important part of the evaluation of research proposals,[8] few ethics committees in Britain include a statistician. Indeed, many ethics committees explicitly take a view of ethics that excludes scientific issues. Consequently, poor or useless studies pass such review even though they can reasonably be considered to be unethical.[9]

The effects of the pressure to publish may be seen most clearly in the increase in scientific fraud,[10] much of which is relatively minor and is likely to escape detection. There is nothing new about the "massage" of data or of "data torture," as it has recently been called[11]—Charles Babbage described its different forms as long ago as 1830.[12] The temptation to behave dishonestly is surely far greater now, when all too often the main reason for a piece of research seems to be to lengthen a researcher's curriculum vitae. Bailar suggested that "there may be greater danger to the public welfare from statistical dishonesty than from almost any other form of dishonesty."[13]

Evaluation of the scientific quality of research papers often falls to statisticians. Responsible medical journals invest considerable effort in getting papers refereed by statisticians; however, few papers are rejected solely on statistical grounds.[14] Unfortunately, many journals use little or no statistical refereeing—bad papers are easy to publish.

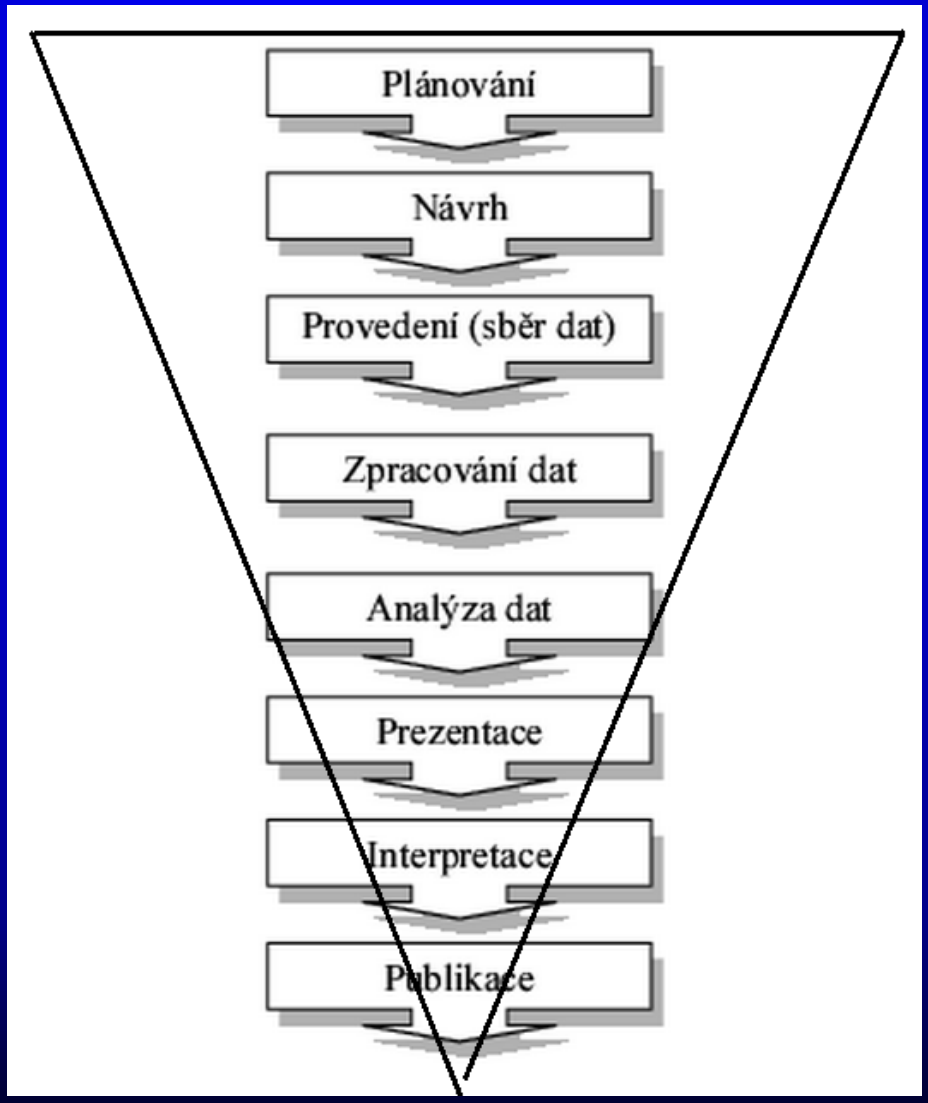
Statistical refereeing is a form of fire fighting. The time spent refereeing medical papers (often for little or no reward) would be much better spent in education and in direct participation in research as a member of the research team. There is, though, a serious shortage of statisticians to teach and, especially, to participate in research.[15] Many people think that all you need to "do" statistics is a computer and appropriate software. This view is wrong even for analysis, but it certainly ignores the essential consideration of study design, the foundations on which research is built. Doctors need not be experts in statistics, but they should understand the principles of sound methods of research. If they can also analyse their own data, so much the better. Amazingly, it is widely considered acceptable for medical researchers to be ignorant of statistics. Many are not ashamed (and some seem proud) to admit that they "don't know anything about statistics."

The poor quality of much medical research is widely acknowledged, yet disturbingly the leaders of the medical profession seem only minimally concerned about the problem and make no apparent efforts to find a solution. Manufacturing industry has come to recognise, albeit gradually, that quality control needs to be built in from the start rather than the failures being discarded, and the same principles should inform medical research. The issue here is not one of statistics

Obecné schéma výzkumného projektu

Zvárová (1998). *Základy statistiky pro biomedicínské obory.*

Obecné schéma výzkumného projektu



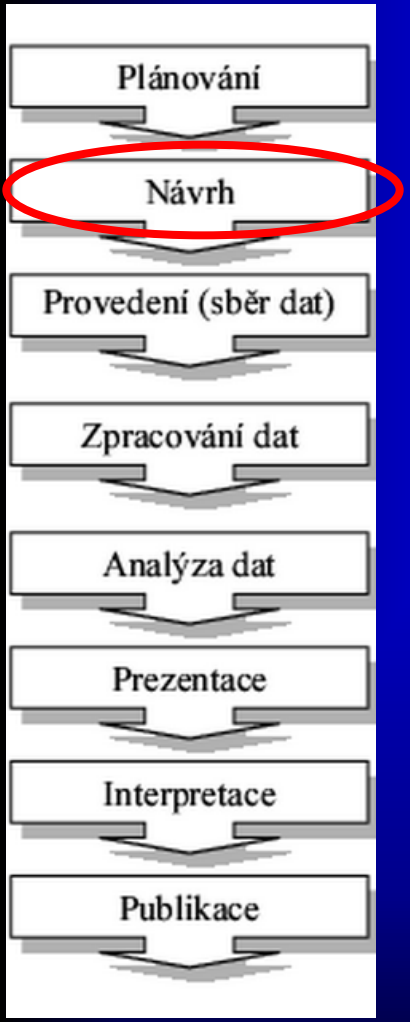
Zvárová (1998). *Základy statistiky pro biomedicínské obory.*

Chyby - příklady

Why are errors so common? Put simply, much poor research arises because researchers feel compelled for career reasons to carry out research that they are ill equipped to perform, and nobody stops them. Regardless of whether a doctor intends to pursue a career in research, he or she is usually expected to carry out some research with the aim of publishing several papers. The length of a list of publications

Altman (1994) *Scandal of poor medical research.*

Chyby - příklady



> One reason for some of the problems is that many studies are not actually designed but rather 'happen'. They are based on an analysis of pre-existing data that were collected for some other purpose. While many

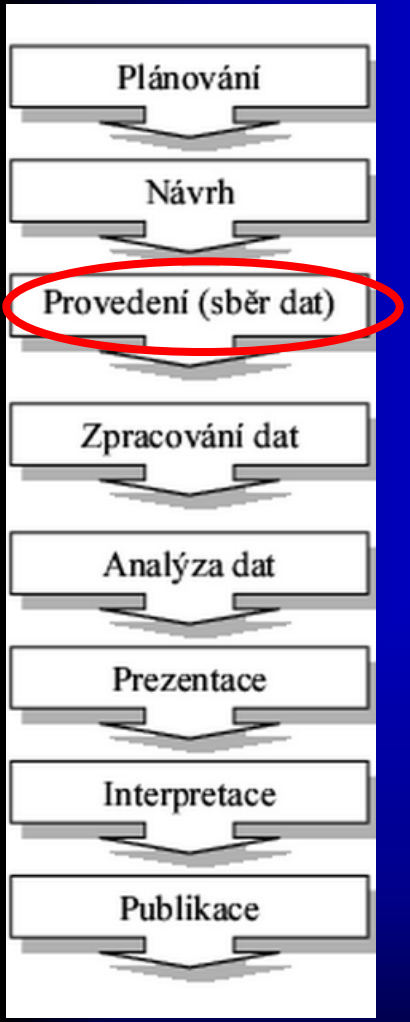
Altman (1991) *Practical Statistics for Medical Research*.

- neadekvátní velikost souboru
  - neprokážeme rozdíl i když je přítomen
  - příliš široké konfidenční intervaly
  - 28 % ze studentů…

    (2 ze 7)

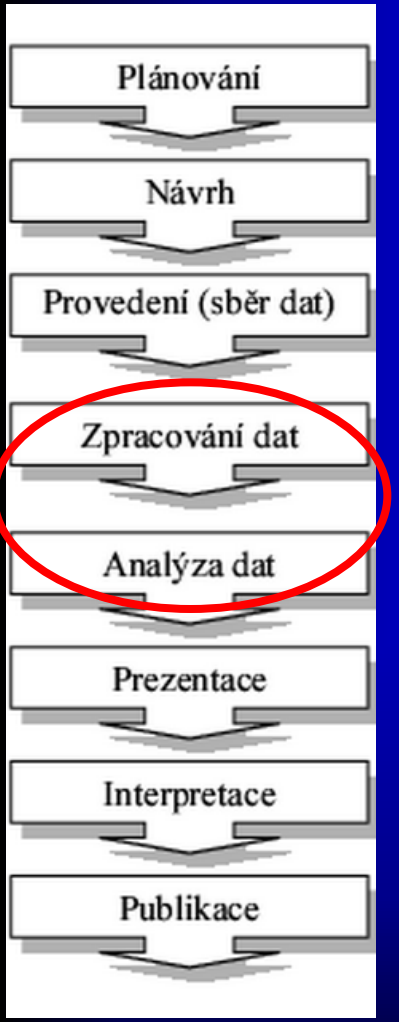
- koncept odhadu velikosti souboru nutného pro studii je v medicíně velmi ojedinělý

Chyby - příklady



- problém mnohočetného porovnávání (opakování statistického testu při stejné 5% hladině významnosti)

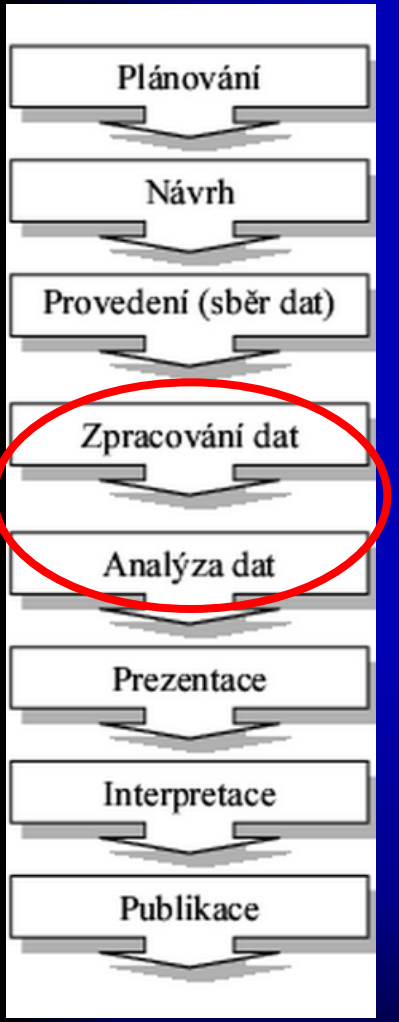
Chyby – příklady
Počítače

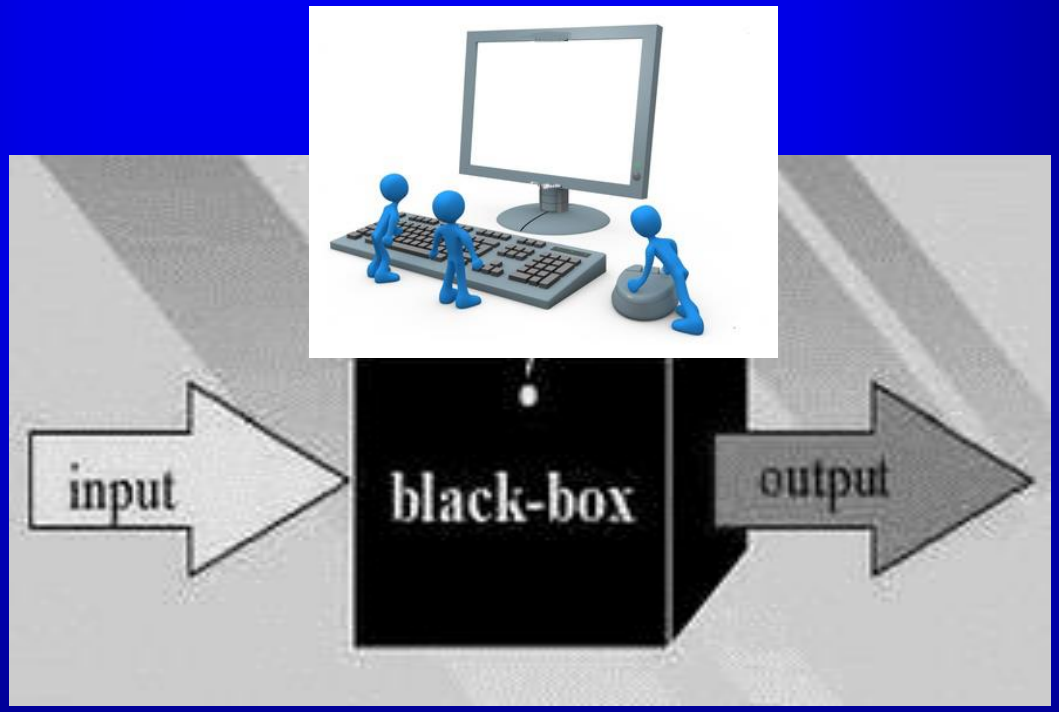


STATISTICA

GraphPad Prism

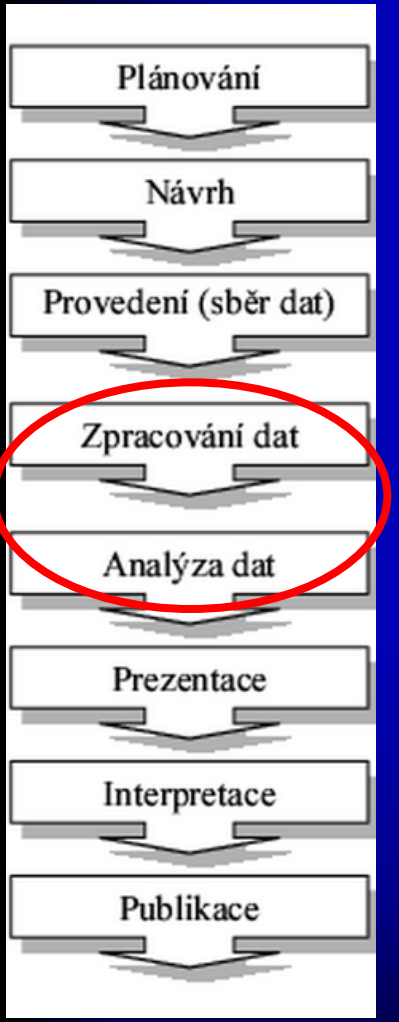
Chyby – příklady
Počítače





Black-box approach

Chyby – příklady
Počítače



- Použití nevhodné metody
- Absence interpretace výsledků

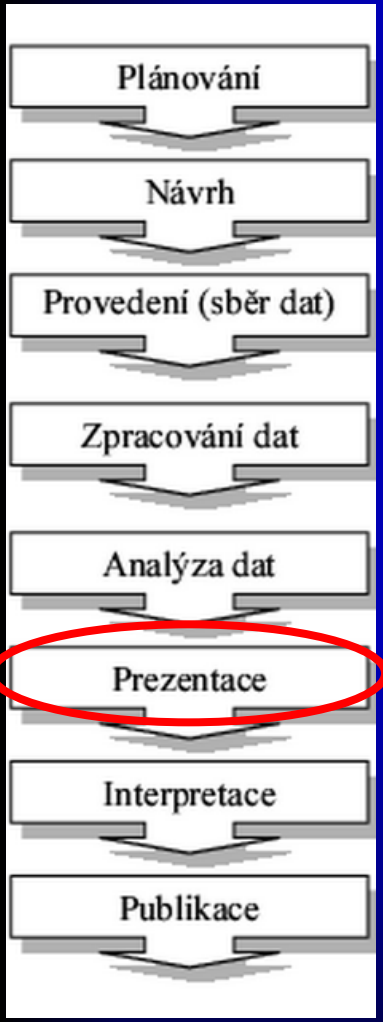
> Most software does nothing to assist in the understanding of statistical principles and cannot help researchers to design sensible studies.

Altman (2000)

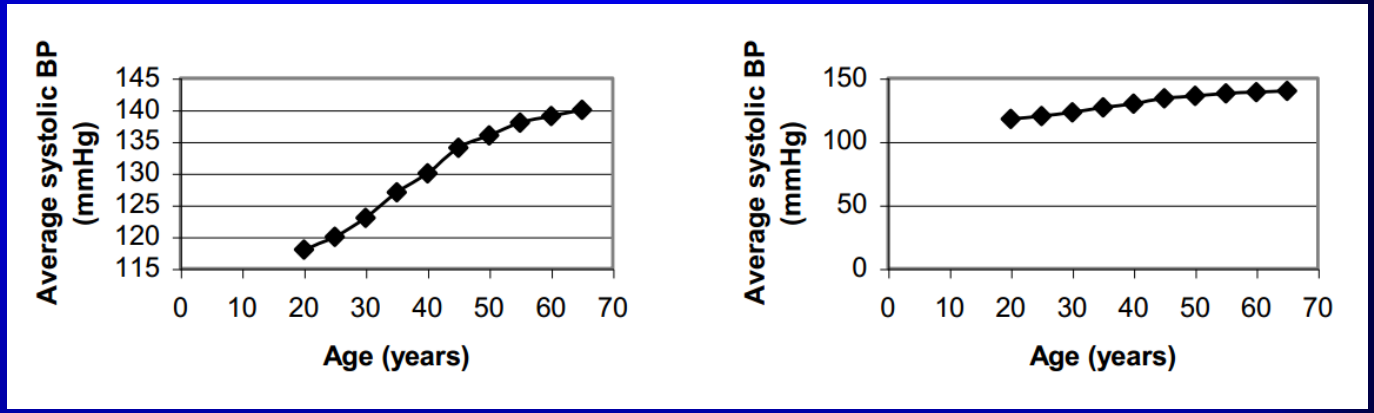
- Black box
  - Exponential regression for survival analysis
    - proportional hazard model $\lambda (t/z) = \lambda e^{\beta z}$
    - log-linear model $\log (T) = -\alpha + \beta^* z$

$\beta^* = - \beta$

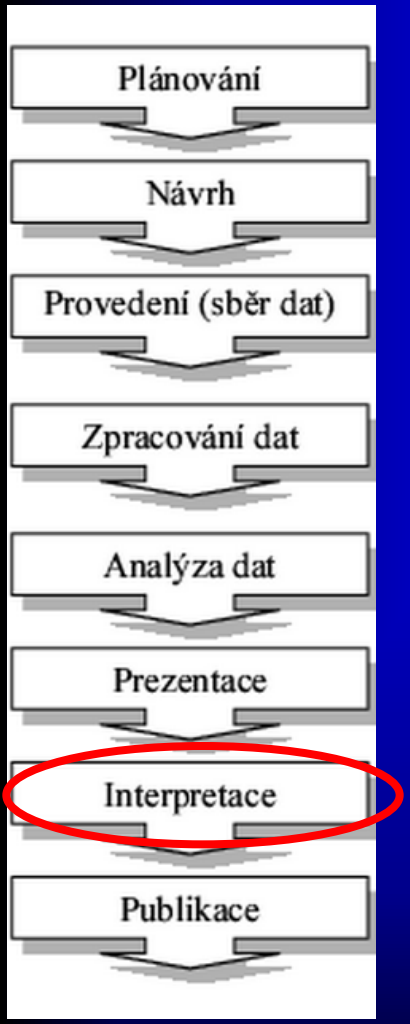
Chyby – příklady



- zavádějící grafická prezentace

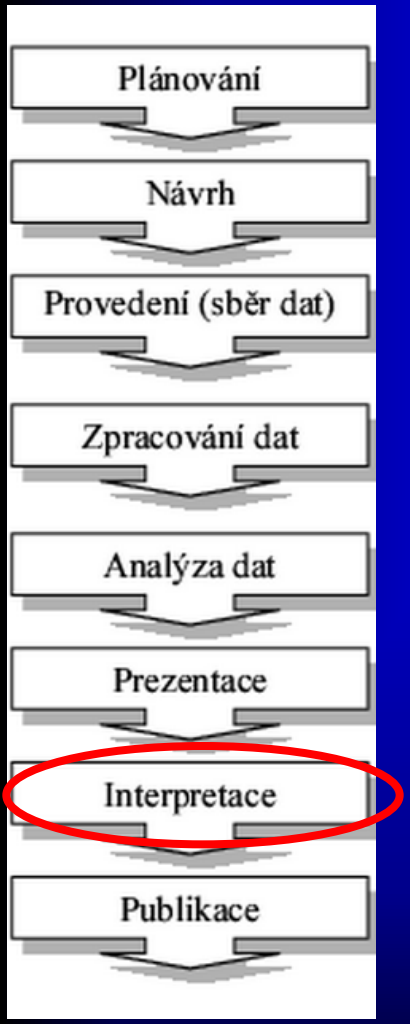


Chyby – příklady



- p-value (p-hodnota) – nepochopení významu
- pravděpodobnost, že pozorujeme naše data nebo data stejně či více extrémní, když je nulová hypotéza pravdivá
- p = 0,001 p = 0,01

Chyby – příklady



- interpretace výsledků studie s nesignifikantním výsledkem jako negativní

- u studie s nízkou sílou statistického testu (malý soubor) není výsledek negativní, ale neprůkazný

The absence of proof is not the proof of absence.

Chyby – příklady

- statisticky významný

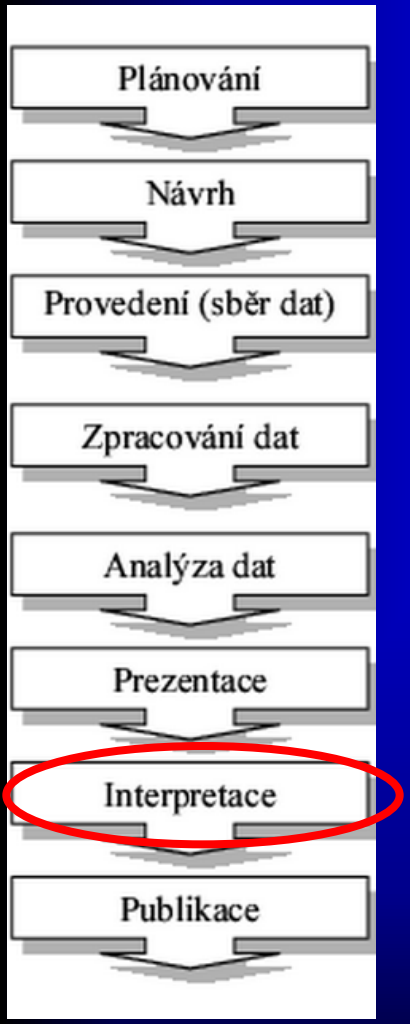
≠

- klinicky významný

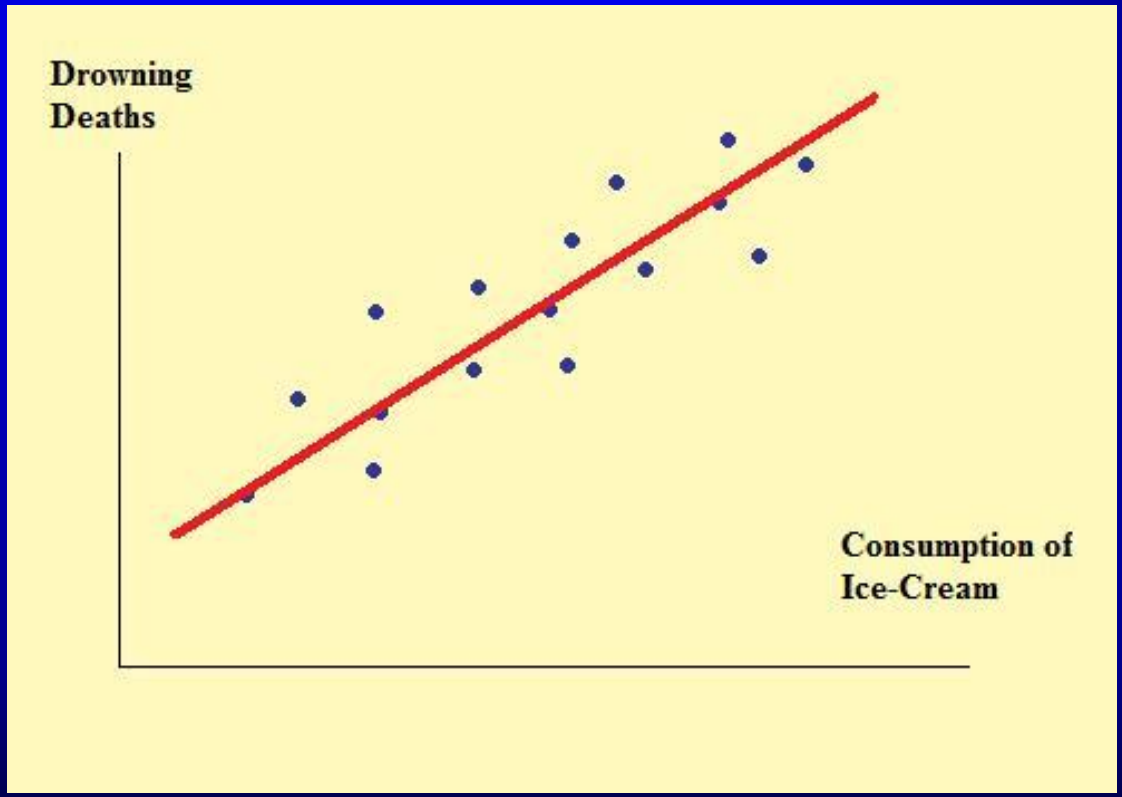
Lék snížil TK v průměru o 18 torr (95% CI = 2-34, p=0,02)

výsledek statisticky významný, klinicky neprůkazný

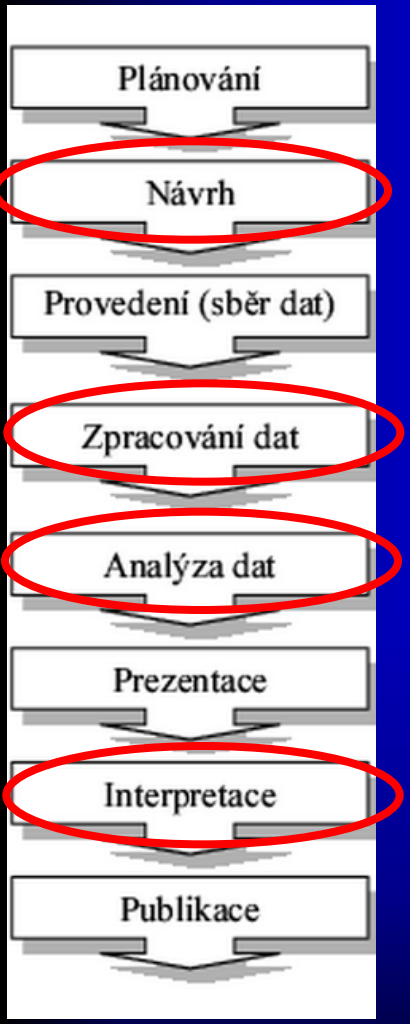
Chyby – příklady



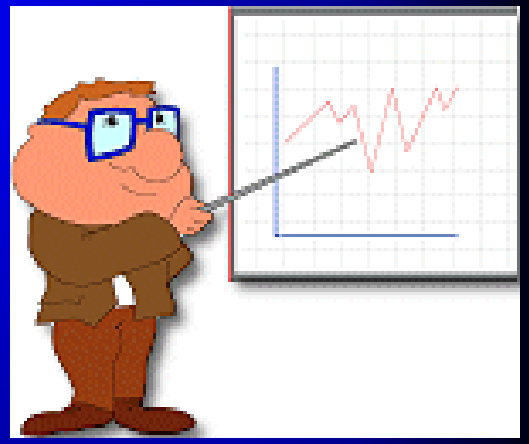
- korelace ≠ kauzalita



Chyby – příklady



- složitá problematika → spolupráce se statistikem



The difference between medical research and agricultural research is that medical research is done by doctors but agricultural research is not done by farmers. *(Michael Healy)*

GENNET
www.gennet.cz

Závěr

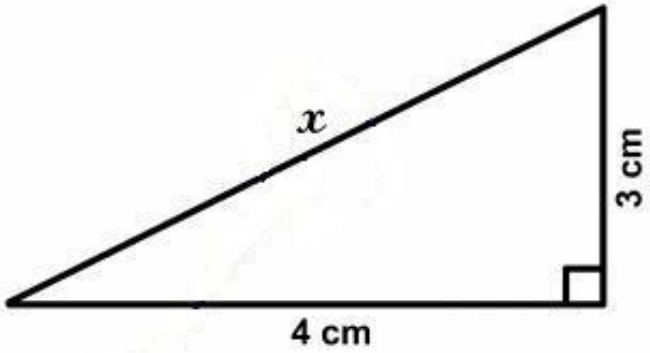
(i) The misuse of statistics is very important.
(ii) A general climate of sloppiness is bad for science.
(iii) Statistics is much more subjective (and difficult) than is usually acknowledged (this is why statisticians have not been replaced by computers).
(iv) Major improvements in the quality of research published in medical journals are unlikely in present research climate.
(v) Too much research is done primarily to benefit the careers of researchers.
(vi) It need not be like this!

(Altman, 2000)

Děkuji za pozornost

Děkuji za pozornost

3. Find x.



Děkuji za pozornost