

**MUNI  
MED**

# **Informatika a statistika ve zdravotnictví - cvičení**

Tomáš Pavlík ([pavlik@iba.muni.cz](mailto:pavlik@iba.muni.cz))

jaro 2024

# Přínos kurzu

- Orientace v principech statistiky, plánování a hodnocení experimentů.
- Orientace v základních metodách, jejich výpočetní podstatě a jejich předpokladech.
- Schopnost aplikace metod při řešení základních problémů z oblasti biologie a medicíny a interpretace výsledků.

# Požadavky k zápočtu

1. 100% docházka
2. Zápočet bude vycházet z přednášek
3. Zápočtová písemka = 50 bodů (zápočet za 25 a více bodů)

# Literatura

1. Přednášky

2. Skripta (<https://www.matematickabiologie.cz/studenti>)

# Přednáška I. Základní aspekty zpracování a analýzy dat

*Tato prezentace je autorským dílem vytvořeným zaměstnanci Masarykovy univerzity. Studenti předmětu mají právo pořídit si kopii prezentace pro potřeby vlastního studia. Jakékoliv další šíření prezentace nebo její části bez svolení Masarykovy univerzity je v rozporu se zákonem.*

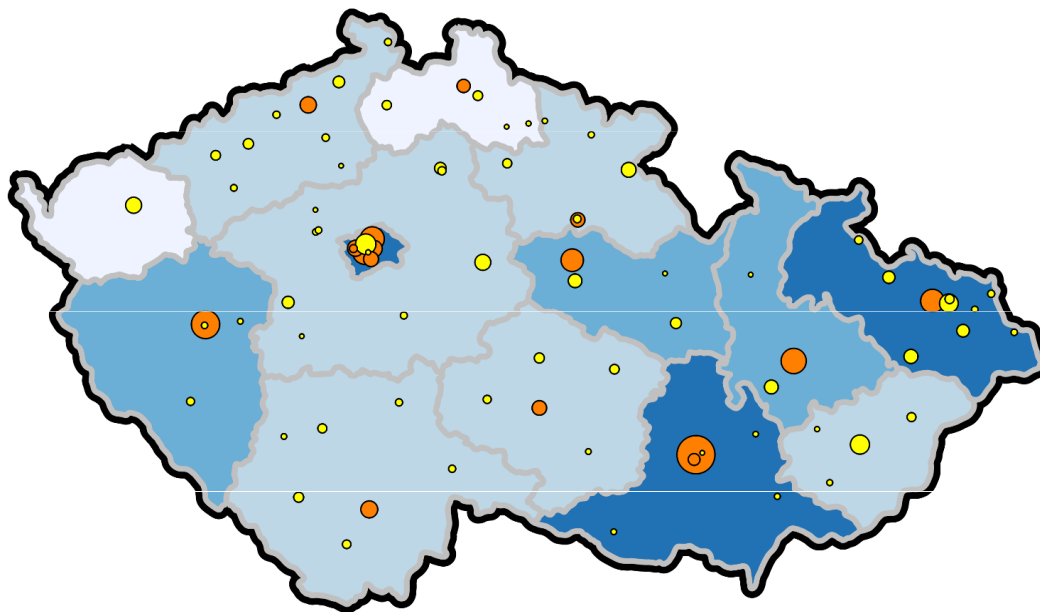
# Příklad: Hodnocení distribuce a dostupnosti zdravotní péče

*Resekce prsu pro ZN včetně odstranění mízních uzlin bez rekonstrukce prsu*

*– data ZP za rok 2017*

N = 5 312

Odstranění prsu i s uzlinami bez následné rekonstrukce je častý výkon, prováděný napříč PALP včetně velkých a specializovaných center.



- TYP ZZ 1:** Fakultní a velké krajské nemocnice, ZZ vysoce specializované péče
- TYP ZZ 2:** Krajské a oblastní nemocnice, Ostatní ZZ s úzkým zaměřením

**Počet hospitalizačních případů v jednotlivých nemocnicích**  
(čím větší bod, tím více hospitalizačních případů)

**Celkový počet hospitalizačních případů v daném kraji:**

1-150   150-300   300-450   450 a více

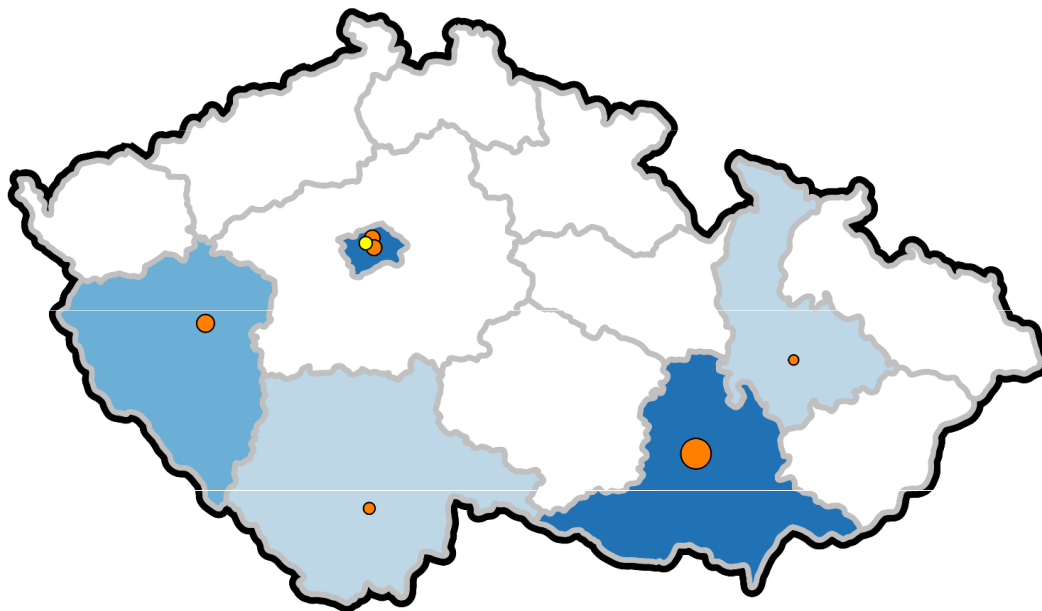
# Příklad: Hodnocení distribuce a dostupnosti zdravotní péče

*Resekce prsu pro ZN včetně odstranění mízních uzlin s rekonstrukcí prsu*

– data ZP za rok 2017

N = 54 = 1 procento!

Odstranění prsu i s uzlinami s následnou rekonstrukcí je prováděno **naprosto minimálně**, a to včetně velkých a specializovaných center.



- TYP ZZ 1: Fakultní a velké krajské nemocnice, ZZ vysoce specializované péče
- TYP ZZ 2: Krajské a oblastní nemocnice, Ostatní ZZ s úzkým zaměřením
- Počet hospitalizačních případů v jednotlivých nemocnicích** (čím větší bod, tím více hospitalizačních případů)

Celkový počet hospitalizačních případů v daném kraji:

0    1-3    3-10    10 a více

# Příklad: Monitoring epidemie COVID-19



## COVID-19: Přehled aktuální situace v ČR

Základní přehled | Očkování | PES: Protiepidemický systém ČR | Kapacity lůžkové péče | Krajské zpravodajství | Kumulativní přehledy | Přehledy dle KHS

<b>Provedené PCR testy</b> 5 214 322 (+ 39 052 za 23.02.) k datu: 24. 2. 2021 v 9.01 h	<b>Provedené antigenní testy</b> 2 461 866 (+ 47 538 za 23.02.) k datu: 24. 2. 2021 v 9.01 h	<b>Vykázaná očkování</b> 581 542 (+ 16 663 za 23.02.) k datu: 24. 2. 2021 v 9.01 h	<b>Aktuálně hospitalizováni</b> 6 817 k datu: 24. 2. 2021 v 9.01 h
---	---	---	--

<b>Potvrzené případy</b> 1 184 352 (+ 15 672 za 23.02.) k datu: 24. 2. 2021 v 9.01 h	<b>Aktivní případy</b> 127 240 k datu: 24. 2. 2021 v 9.01 h	<b>Vyléčení</b> 1 037 430 k datu: 24. 2. 2021 v 9.01 h	<b>Úmrtí</b> 19 682 k datu: 24. 2. 2021 v 9.01 h
---	---	--	--

[Vysvětlení základních pojmů spojených s onemocněním COVID-19 na tomto webu](#) | [Přidej se k nám a zapoj se do vývoje webových aplikací proti COVID-19](#)

[Datové sady COVID-19](#) | [Popis zpracování dat](#) | [Zprávy a metodiky](#) | [COVID-19 Widget](#) | [Často kladené otázky](#) | [Aktuálně o COVID-19](#)

## + OTEVŘENÉ DATOVÉ SADY

### Kapacity lůžkové péče

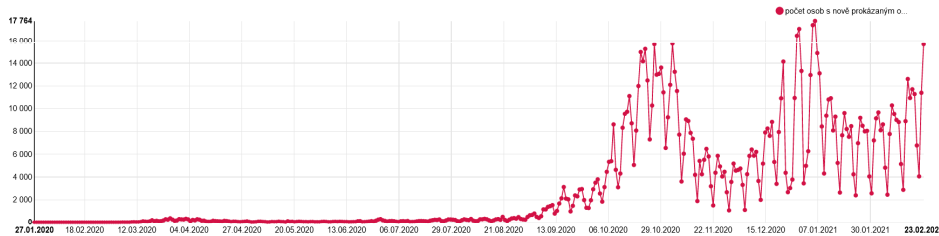
**Přehled ČR**  
Přehled aktualizován 24. 2. 2021 v 15.36 h

<b>Přístroj: ECMO</b> kapacita: volná / celkem 38 / 85 <b>44 %</b>	<b>Přístroj: UPV</b> kapacita: volná / celkem 701 / 2 300 <b>30 %</b>	<b>Přístroj: CRRT (kontinuální dialýza)</b> kapacita: volná / celkem 210 / 323 <b>65 %</b>
<b>Přístroj: IHD (intermitentní dialýza)</b> kapacita: volná / celkem 131 / 225 <b>58 %</b>	<b>Lůžka (ARO+JIP dospělí)</b> kapacita: volná / celkem 571 / 3 961 COVID+ 170 COVID- 401 <b>14 %</b>	<b>Lůžka (standardní s kyslíkem)</b> kapacita: volná / celkem 5 117 / 21 444 COVID+ 708 COVID- 4 409 <b>23 %</b>
<b>Ventilátory (přenosné)</b> kapacita: volná / celkem 521 / 548 <b>95 %</b>	<b>Ventilátory (operační sál)</b> kapacita: volná / celkem 774 / 883 <b>87 %</b>	<b>Reprofilizovaná kapacita: Lůžka IP (UPV +/-)</b> kapacita: volná / skutečná 144 / 768 kapacita: skutečná / plánovaná 768 / 1 627 <b>18 %</b> <b>47 %</b>

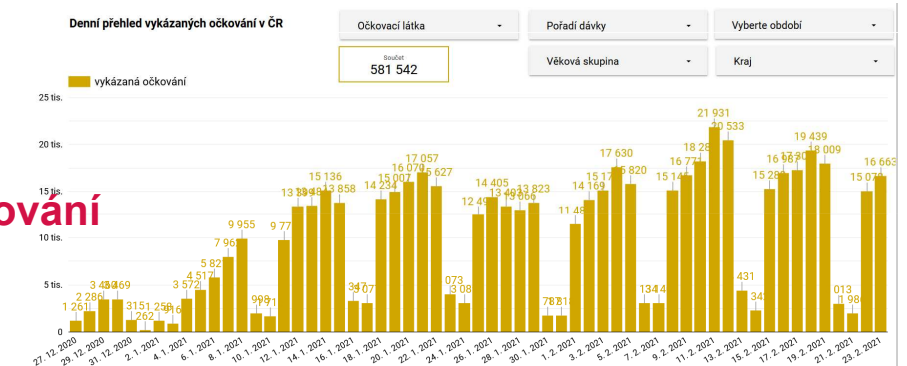
## Testování a epidemiologie

Denní přehled počtu osob s nově prokázaným onemocněním COVID-19 dle hlášení krajských hygienických stanic a laboratoří

Přehled za posledních 14 dní | [Kompletní přehled za celé období](#) | [Tabulkový přehled](#)



## Očkování



<https://onemocneni-aktualne.mzcr.cz/covid-19>



# Další příklady použití statistiky v biologii a medicíně?

# **O čem ta statistika vlastně je?**

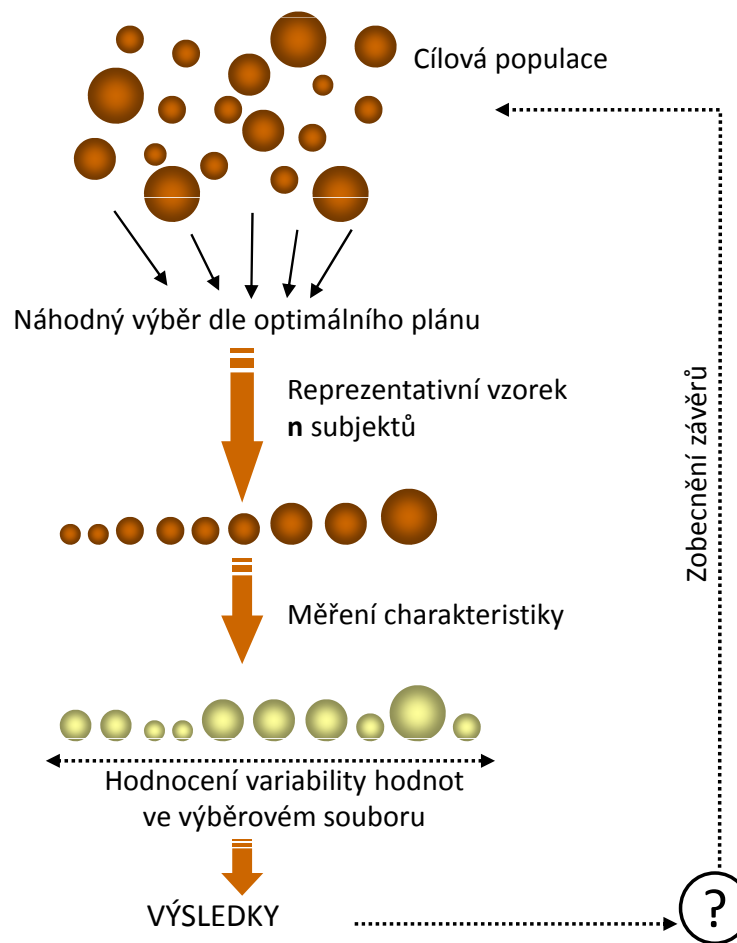
# Statistika a biostatistika

- Biostatistika je **aplikace statistických metod** v řešení biologických a klinických problémů.
- Snahou je **získat z pozorovaných dat užitečnou informaci**. V popředí zájmu je pozorovaná variabilita mezi studovanými subjekty, kterou chceme vysvětlit.
- Biostatistika je **orientována na konkrétní problém**, ne na teoretické aspekty. To však neznamená, že lze statistické metody používat bezhlavě.

# Jaké úlohy můžeme s pomocí statistiky řešit?

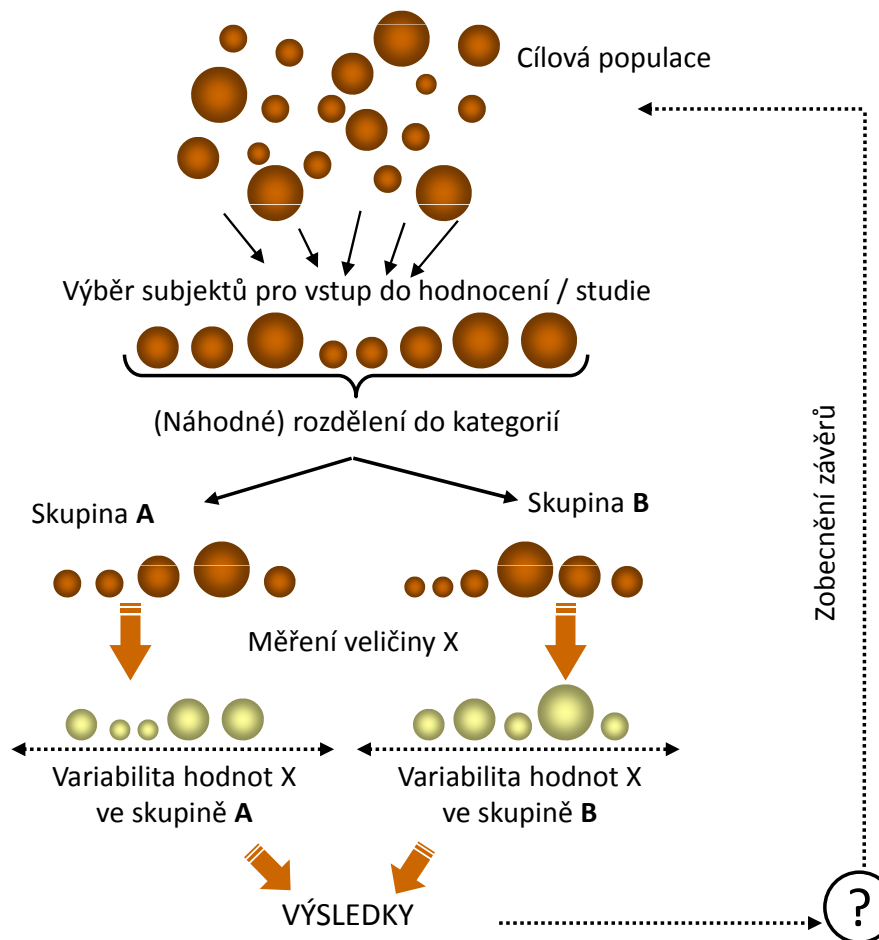
- **Popis cílové populace** – odhady charakteristik cílové populace
- **Srovnání skupin** – testování hypotéz
- **Regresní analýza** – stochastické modelování pro vysvětlení variability
- **Predikce a klasifikace** – stochastické modelování a klasifikační algoritmy pro předpovídání neznámých hodnot

# Popis cílové populace – popis pozorované variability



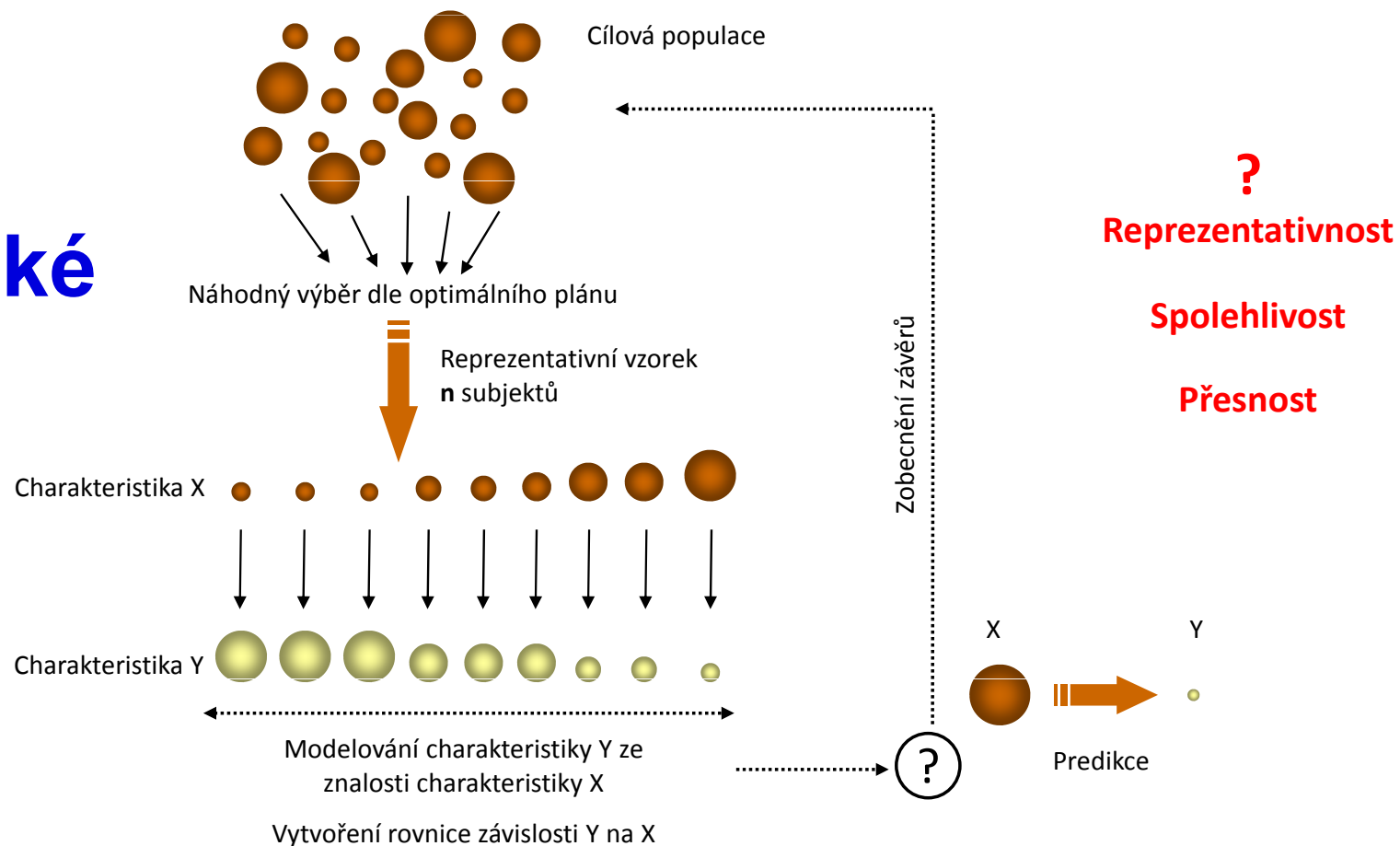
**?**  
**Reprezentativnost**  
**Spolehlivost**  
**Přesnost**

# Srovnání skupin – srovnání pozorované variability

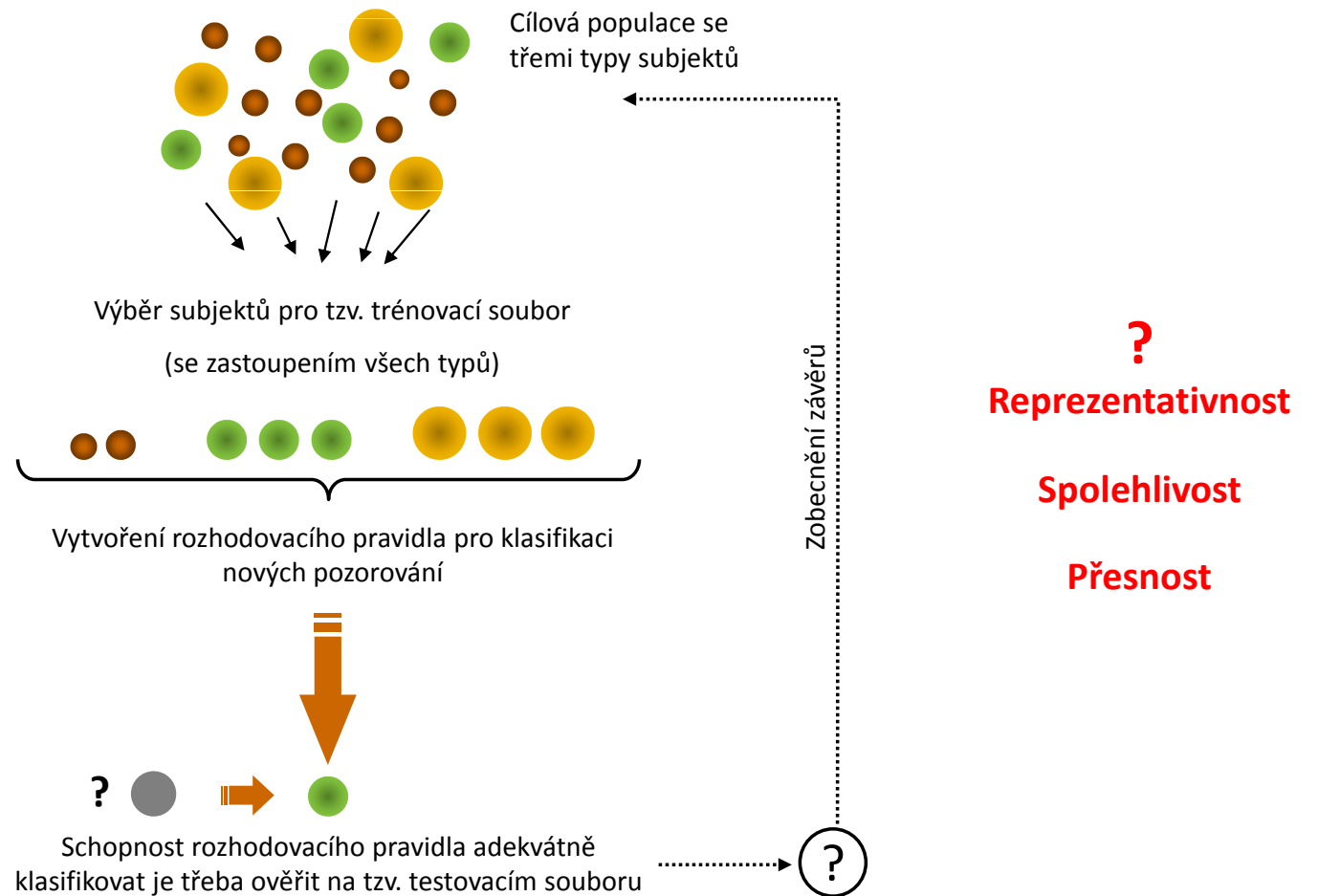


**?**  
**Srovnatelnost**  
**Spolehlivost**  
**Přesnost**

# Predikce neznámých hodnot + stochastické modelování



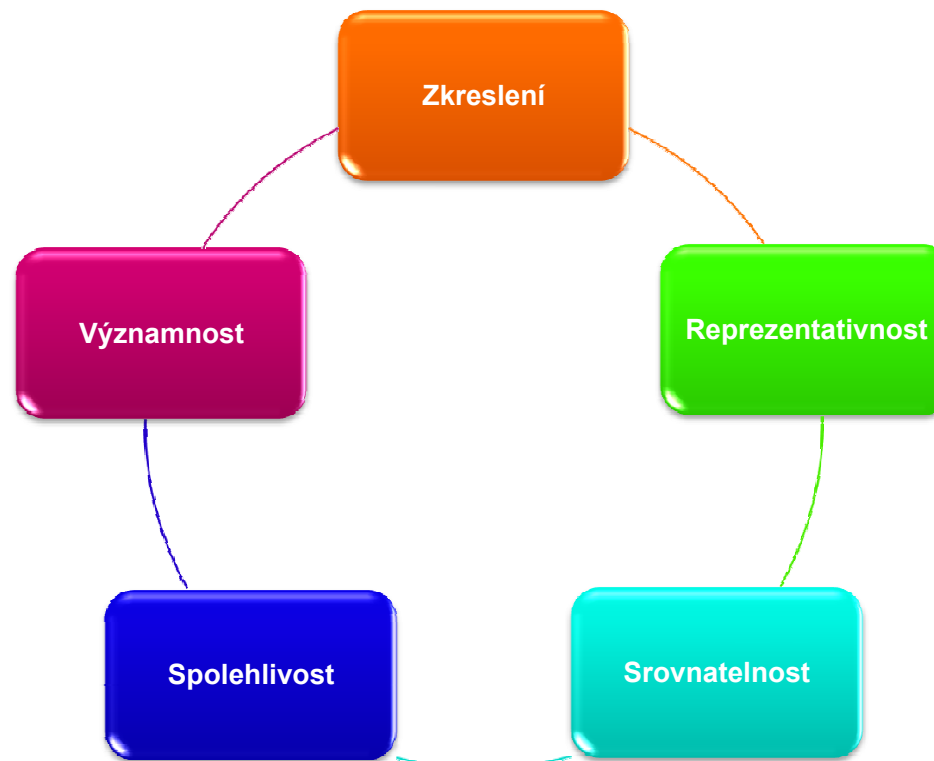
# Klasifikace nových pozorování – klasifikační algoritmy





# **Klíčové aspekty statistiky**

# Klíčové aspekty statistických odhadů



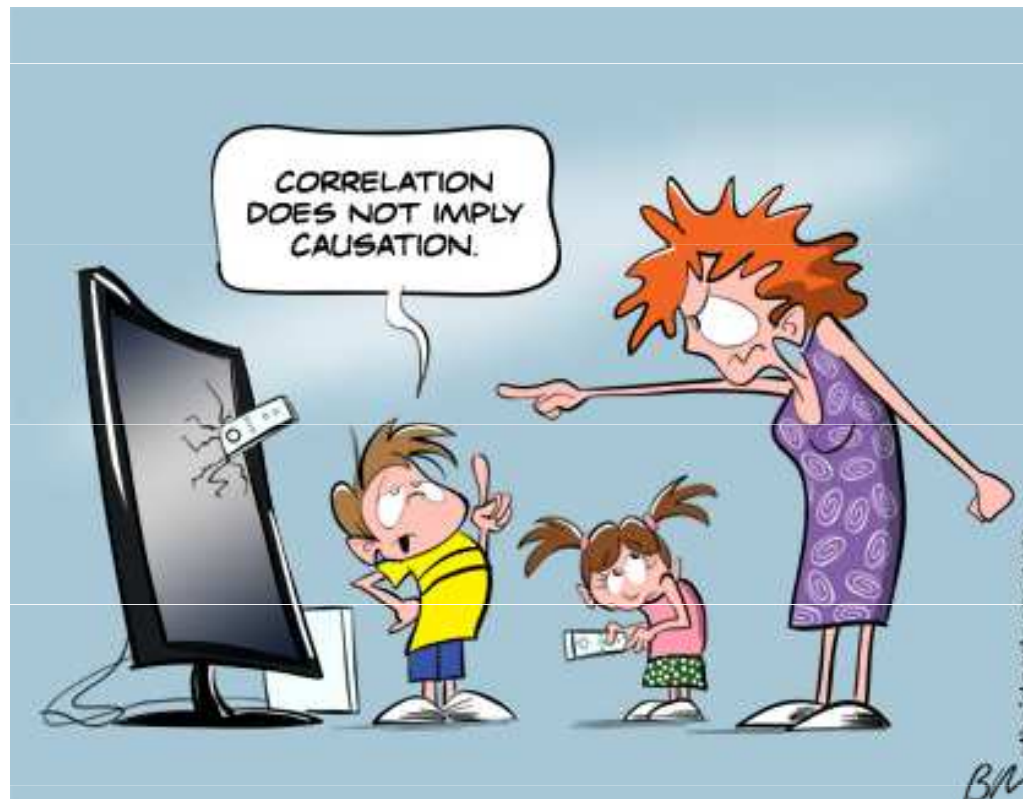
# Klíčové aspekty – zkreslení

- V jakémkoliv hodnocení se **snažíme vyhnout zkreslení výsledků** („*biased results*“), tedy zkreslení výsledků jinými faktory než těmi, které jsou cíli studie.

***„Bias is any process at any stage of inference tending to produce results that differ systematically from the true values.”***

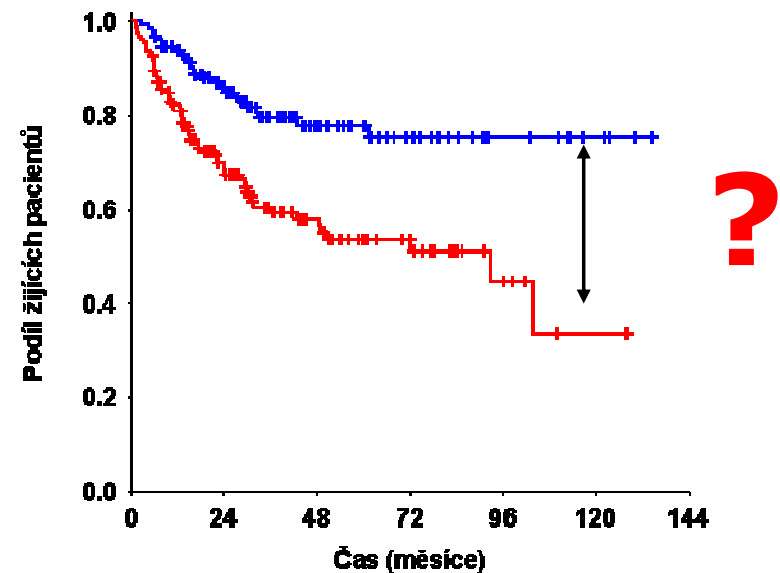
- Statistické srovnání není nikdy 100% spolehlivé, existuje náhoda a tedy i pravděpodobnost chybného úsudku. Chceme použít adekvátní metody pro odstranění vlivů, které by zkreslily výsledky a nebyly přitom náhodné (např. zastoupení pohlaví, věk, apod.).

# Klíčové aspekty – zkreslení



# Klíčové aspekty – zkreslení

- Čím by mohl být způsoben pozorovaný rozdíl v 10letém přežití pacientů s nádorem trávicího traktu?



# Klíčové aspekty – zkreslení

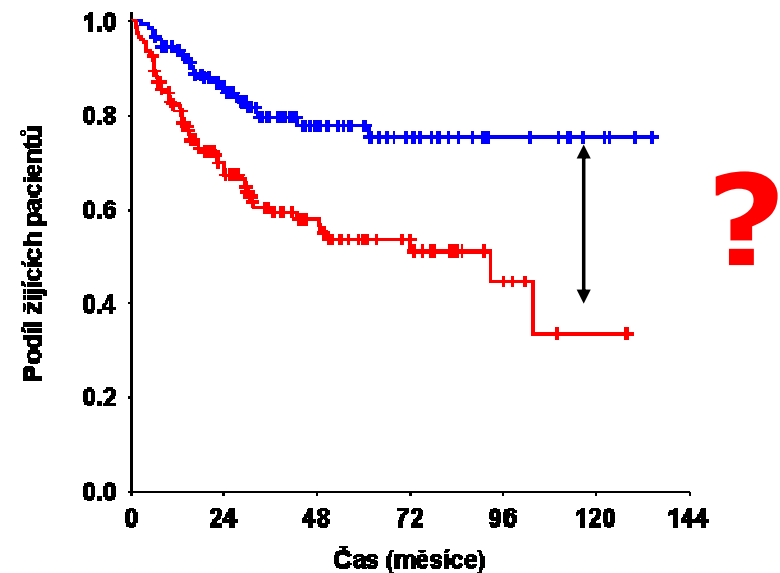
- Čím by mohl být způsoben pozorovaný rozdíl v 10letém přežití pacientů s nádorem trávicího traktu?

**Léčba?**

**Nějaký prognostický faktor?**

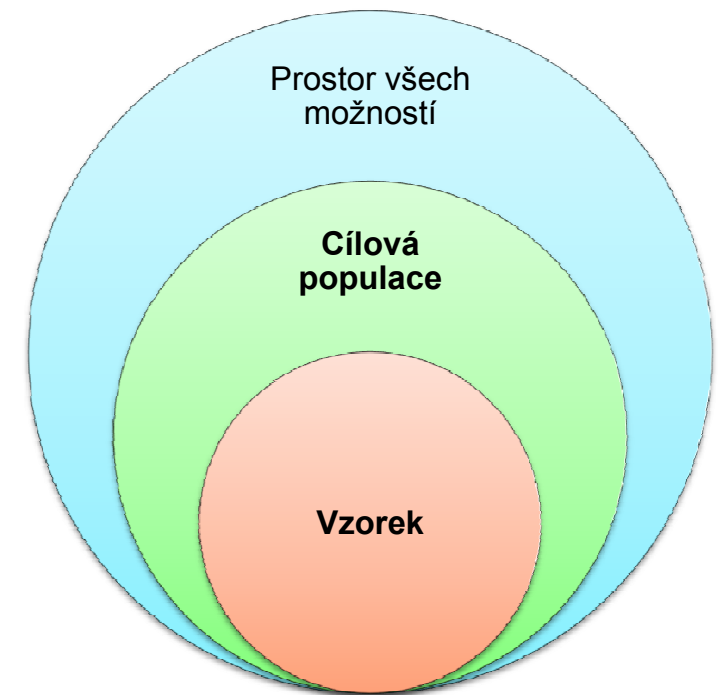
**Stadium nemoci?**

**Věk?**

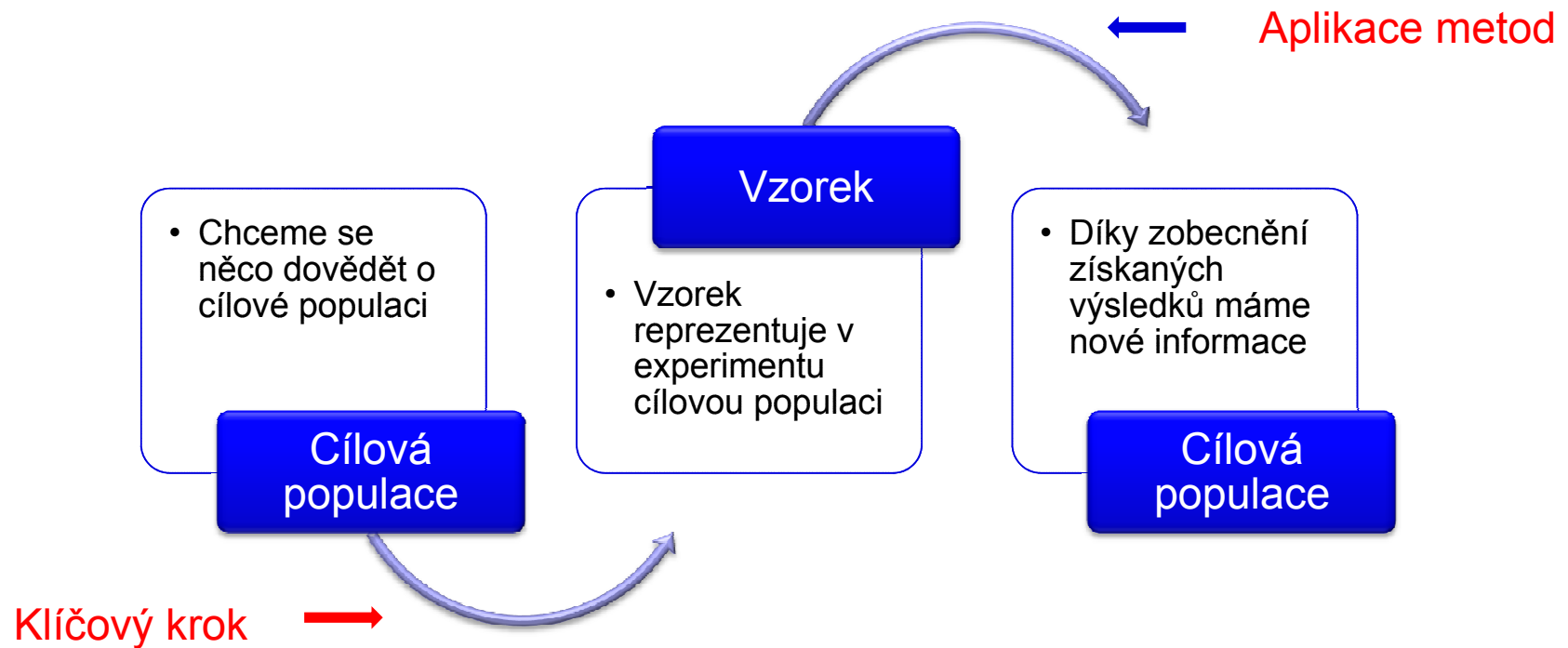


# Klíčové aspekty – reprezentativnost

- **Cílová populace** – skupina subjektů, o které chceme zjistit nějakou informaci.
- **Experimentální vzorek** – podskupina cílové populace, kterou „máme k dispozici“.
- Vzorek musí odpovídat svými charakteristikami cílové populaci.
- Chceme totiž zobecnit výsledky na celou cílovou populaci.



# Klíčové aspekty – reprezentativnost





# Klíčové aspekty – srovnatelnost

- Korektní výsledky při srovnávacích analýzách lze získat **pouze při srovnávání srovnatelného**.
- V kontrolovaných klinických studiích je srovnatelnost zajištěna **randomizací** (do určité míry).
- U studií bez randomizace **je nutné se tématu srovnatelnosti skupin věnovat**.
- Metody adjustace, matching, propensity scores.



# Klíčové aspekty – spolehlivost

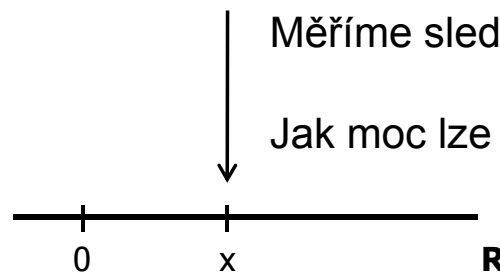
- Ve většině studií nás zajímá **kvantifikace sledovaného efektu nebo charakteristiky** ve formě jednoho čísla, tzv. bodového odhadu.
- Bodový odhad je však sám o sobě nedostatečný.
- Je nutné ho doplnit **intervalovým odhadem**, který odpovídá pozorované variabilitě sledované veličiny a odráží spolehlivost výsledku.

# Klíčové aspekty – spolehlivost

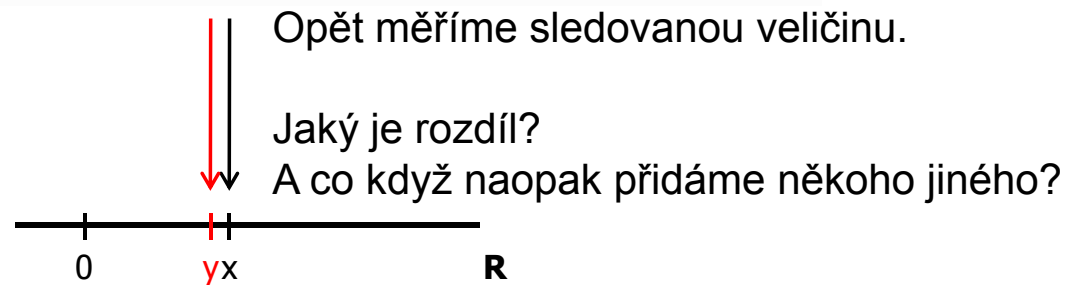


Měříme sledovanou veličinu a následně spočítáme odhad.

Jak moc lze bodový odhad zobecnit na cílovou populaci?

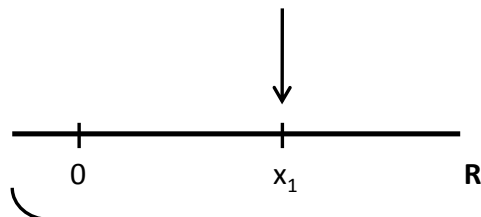


# Klíčové aspekty – spolehlivost

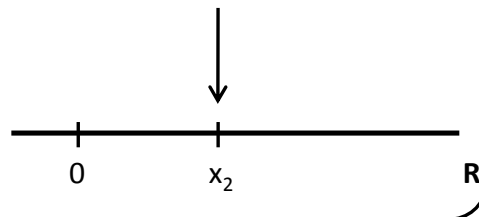
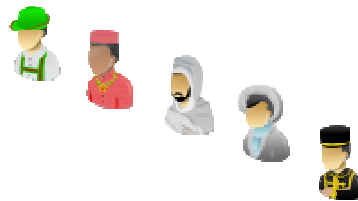


# Klíčové aspekty – spolehlivost

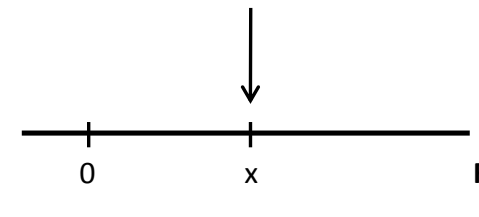
Výběr číslo 1



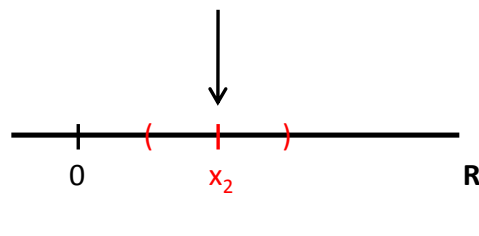
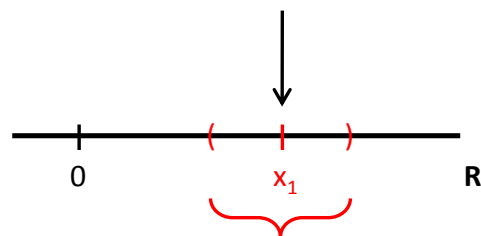
Výběr číslo 2



Celá cílová populace



Pracujeme-li s výběrem z cílové populace, je třeba na základě variability pozorovaných dat spočítat tzv. interval spolehlivosti pro bodový odhad.



Interval spolehlivosti na základě výběru číslo 1.

Umíme-li „změřit“ celou cílovou populaci, nepotřebujeme interval spolehlivosti, protože jsme schopni odhadnout sledovaný parametr přesně – v praxi je tato situace nereálná.

# Klíčové aspekty – významnost

- Analytické výsledky studie nemusí odpovídat realitě a skutečnosti.  
**Statistická významnost jednoduše nemusí znamenat příčinný vztah!**
- Statistická významnost pouze indikuje, že pozorovaný rozdíl není náhodný (ve smyslu stanovené hypotézy). Stejně důležitá je i **praktická významnost, tedy významnost z hlediska lékaře nebo biologa.**
- Statistickou významnost lze ovlivnit velikostí vzorku.

# Klíčové aspekty – významnost

## Praktická významnost

		Praktická významnost	
		ANO	NE
Statistická významnost	ANO	OK, praktická i statistická významnost je ve shodě.	Významný výsledek je statistický artefakt, prakticky nevyužitelný.
	NE	Výsledek může být pouhá náhoda, neprůkazný výsledek.	OK, praktická i statistická významnost je ve shodě.



Statisticky nevýznamný výsledek neznamená, že pozorovaný rozdíl ve skutečnosti neexistuje! Může to být způsobeno nedostatečnou informací v pozorovaných datech!

**MUNI  
MED**

**Děkuji za  
pozornost**