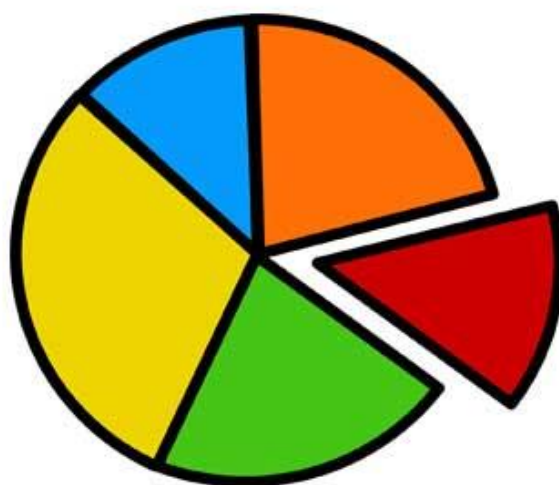


BIOSTATISTIKA PRAKTICKY

aneb

JAK NA PRAKTICKOU ČÁST KVALIFIKAČNÍ PRÁCE TAK, ABY MĚLA HLAVU A PATU

Poznámky k přípravě výzkumu, sběru a statistickému zpracování dat zejména pro potřeby tvorby studentských absolventských a kvalifikačních prací, ale i pro další zájemce o triviální biostatistiku.



MUDr. Ondřej Franěk

www.zachrannasluzba.cz

aktualizace 11. 10. 2018

PODĚKOVÁNÍ

Velké poděkování patří Tomovi Meinlschmidtovi za korekturu textu, pohlídání korektní typografické podoby a celkovou typografickou kultivaci tohoto veledíla. Protože text je živý a byl doplňovaný i po jeho korekci, platí, že co je dobře, je díky Tomovi, co je špatně, je moje chyba ☺.

Obsah

1	Úvod.....	5
2	Výzkumný problém a cíl práce – aneb kterým směrem se vydat?.....	6
2.1	Formulace problému	6
2.2	Data – kde vzít a nekrást...?.....	7
2.3	Typ výzkumu	8
3	Jdeme na to aneb jak vyrobit studii	11
4	Sběr dat.....	13
4.1	Jaký bude rozsah šetření?	13
4.2	Experimentální versus kontrolní skupina, randomizace	15
4.3	Na co si dát pozor.....	15
5	Zdroje chyb.....	18
5.1	Výběrová chyba	18
5.2	Chybné měření.....	19
5.3	Placebo efekt.....	20
5.4	Hawthornský fenomén	21
5.5	Osobní zájem/nezájem	21
6	Základy měření a zpracování biomedicínských dat.....	22
6.1	Typy dat.....	22
6.2	Ukazatelé polohy a variability.....	23
7	Stanovení hypotézy a hodnocení její platnosti.....	25
7.1	Co je to hypotéza?	25
7.2	Kdy se vůbec zabývat hypotézou	26
7.3	Kvalitativní stanovení rozdílu – je rozdíl dílem náhody?	26
7.4	Kvantitativní stanovení rozdílu – „jak moc“ je prokázaný rozdíl velký?	29
8	Statistické testování v typických příkladech.....	31

8.1	Cíl testování a základní informace	31
8.2	Nejjednodušší, nejobvyklejší používané statistické testy.....	32
8.3	Výběr vhodného testu v typických situacích.....	34
9	Praktický „checklist“ pro přípravu studie	38
10	Publikace.....	39
10.1	Standardní struktura publikace	39
10.2	Etika publikace	39
10.3	Jazyk a typografie publikace.....	40
11	Nejčastější chyby při přípravě výzkumu, ve zpracování a hodnocení dat	43
11.1	Nejčastější chyby v přípravě výzkumu.....	43
11.2	Typické chyby v dotaznících	44
11.3	Nejčastější chyby ve zpracování dat.....	46
11.4	Nejčastější chyby v prezentaci dat.....	46
12	Zdroje	51

1 Úvod

Dlouhé roky jsem měl tu povinnost (a čest) zabývat se studentskými kvalifikačními pracemi na úrovni vyšších a vysokých škol zdravotnického zaměření, a to jak v pozici vedoucího práce, tak v pozici oponenta. Během této doby jsem ale dospěl k přesvědčení, že metodologie sběru dat, stanovení hypotéz a jejich korektní vyhodnocení je dlouhodobě „velkou neznámou“ nejen pro drtivou většinu studentů, ale v řadě případů i pro jejich pedagogy, resp. vedoucí jejich prací. I přes všeobecný, celospolečensky pozorovatelný příklon k rozhodování založeném spíše na emocích a intuici, jsem ale stále přesvědčen o tom, že racionální hodnocení faktických údajů je jediná cesta, jak dosahovat skutečného poznání a pokroku.

Následující text je tedy pokusem o velmi stručnou pomůcku, které by měla pomoci metodicky správně (nebo alespoň přijatelně) navrhnout a vyhodnotit výzkumnou část závěrečné práce z hlediska sběru a zpracování dat.

Co se statistické části týče, považujte prosím tento text pouze za pootevřené okénko do světa statistiky. Jde vlastně o „kurs první statistické pomoci“, který si musí vystačit s řadou někdy i velmi zásadních zjednodušení a z celé problematiky statistického zpracování dat zahrnuje jen zcela titěrnou oblast. Mým úmyslem bylo celý text zaměřit velmi prakticky – proto je také nezbytná teorie doplněná řadou příkladů a „hovorových“ vysvětlení toho, k čemu jsou vlastně uvedené teoretické informace dobré. Tyto části textu jsou uvedeny *kurzívou*.

Přestože statistika zdánlivě nesouvisí češtinou a typografií, nenechme se mýlit – chceme-li opravdu přesně formulovat myšlenky, neobejdeme se bez správného použití jazyka, větných vazeb, diakritiky, interpunkce atd. Moje zkušenost je přitom taková, že i v tom bohužel drtivá většina studentů „plave“. Jakákoliv práce má přitom smysl jen tehdy, pokud je přesně jasné, jak „to“ autor myslel. Součástí materiálu je proto i kapitola, shrnující základní zásady pro tvorbu publikace, typografické úpravy textu atd.

PS: Prosím o shovívavost opravdové statistiky – jsem si vědom řady zjednodušení a nevyřčených předpokladů, ale i tak jsem přesvědčen, že pro 99 % členů cílové skupiny může být tento spis i tak zásadním přínosem.

2 Výzkumný problém a cíl práce – aneb kterým směrem se vydat?

Souhrn ve třech větách:

- *Než začnete psát práci, rozmyslete si, jaký **typ práce** budete tvořit: buď chcete jen něco konstatovat (je to tak a tak), nebo budete analyzovat a porovnávat rozdíly mezi dvěma skupinami a výsledkem bude konstatování, zda se skupiny v určitém parametru liší, nebo ne.*
- *Dopředu si rozmyslete **téma práce** a sledované parametry tak, aby jejich sběr byl reálně proveditelný a vyhodnocení mělo praktický smysl.*
- ***Nemějte velké oči** – definujte pokud možno malé množství jednoduchých a jasných parametrů – je lepší mít méně kvalitních dat, než množství nespolehlivých údajů.*

2.1 Formulace problému

Prvním (zdánlivě zcela samozřejmým) krokem při rozhodování o tvorbě práce by měla být **FORMULACE PROBLÉMU** – tj. „o čem to vlastně bude“ – co chceme prací zjistit, případně prokázat. Na základě toho bychom měli stanovit konkrétní **CÍL PRÁCE**.

Jde o zcela zásadní rozhodnutí, které předurčuje celou další metodiku výzkumu, zpracování a prezentaci výsledků.

Příklad:

FORMULACE PROBLÉMU: „V praxi se často setkáváme s tím, že svědci nehody nejsou schopni správně poskytnout první pomoc obětem dopravních nehod (...následuje literatura, která toto tvrzení dokládá...).“

CÍL PRÁCE: „Cílem práce je zjistit, zda zavedení zážitkové metodiky výuky první pomoci zlepšuje schopnost laiků prakticky poskytnout první pomoc.“

Častou chybou je, že cíl práce je příliš obecný, a tudíž nic neříkající („Cílem práce je zkoumat znalosti první pomoci.“)

2.2 Data – kde vzít a nekrást...?

Následujícím krokem autora je rozhodnutí, **kde a kdy bude sbírat data**. Sebelepší nápad není nic platný, pokud k němu potřebujeme data, která nemáme. Se stanovením cíle úzce souvisí i to, zda budeme vycházet z dat, která už existují (retrospektivní studie), nebo budeme data sbírat přesně podle vlastní metodiky (prospektivní studie). Tohoto rozhodování jsme ušetřeni, pokud my sami provádíme nějaký experiment (= intervenční studie) – taková studie je pochopitelně vždy prospektivní. Viz dále Sběr dat.

*Přestože to obvykle není předmětem kvalifikačních prací, pro jistotu se patří upozornit na to, že není možné přijít „jen tak“ s nápadem modifikujícím stávající odborné postupy. Studie, které se dotýkají jakéhokoliv aktivního vstupu do obvyklé léčby pacientů, podléhají poměrně složitému procesu **schvalování etickou komisí**, která je obvykle ustavena na pracovištích univerzitního typu.*

Vždy při zpracování dat pracujeme buď s **daty, která už existují** (např. jsou zadávána do počítače v rámci rutinního provozu organizace) – pak hovoříme o **RETROSPEKTIVNÍ** studii, případně budeme data sbírat nově (např. pomocí dotazníků, vlastního experimentu apod.) – pak jde o **PROSPEKTIVNÍ** studii. Obě možnosti mají svoje výhody a nevýhody:

Retrospektivní studie (data už existují)

- zpracování už existujících dat (registry, databáze...);
- vhodné pro pracování velkých souborů dat, zejména takových, kde jsou hodnoty zaznamenávány rutinně nebo automaticky (např. dojezdové časy záchranné služby, měření teploty v určitou hodinu, počty ošetřených pacientů apod.);
- výhoda = rychlé, laciné, konzistentně (stále stejně) zadávaná data, není „úmyslné zkreslení“; pokud je systémová chyba, je stále stejná (pokud se metodika časem nezměnila) a vcelku spolehlivě lze tedy hodnotit trendy (stoupá x klesá, zvětšuje se x zmenšuje se apod.), byť absolutní změřená hodnota nemusí být přesná; možnost zahrnutí velkého množství dat (za řadu let zpětně);

- nevýhoda = předem (někým jiným) daná metodika, která nemusí přesně respektovat naše potřeby (může obsahovat systémové zkreslení) a navíc se mohla v čase měnit (nasazení nového evidenčního sw. apod.); není možné ověřit, zda byla sbírána validní data; není možné získat doplňující informace; výsledná data jsou spíše orientační, trendová, přehledová; nejsou vhodné pro detailní zkoumání konkrétního jevu.

Prospektivní studie (data budu teprve sbírat v rámci studie – „na míru“)

- sběr a zpracování dat na základě námi navržené metodiky;
- vhodné pro zpracování menších dat, kde je požadované získání cíleně zaměřených údajů, které se jinak neshromažďují;
- výhody = metodika sběru může být přesně taková, jakou výzkumník požaduje; v průběhu sběru dat je stále stejná; je možné získat doplňující informace;
- výsledky jsou vhodné pro podrobné hodnocení absolutních hodnot i trendů;
- nevýhody = ekonomická a časová náročnost; často je potřeba velké množství spolupracovníků, které je nutné zaškolit a motivovat; možnost zanesení chyby, kterou je velmi obtížné až nemožné zpětně odhalit.

*Nejjednodušší formou prospektivního sběru dat v praxi závěrečných prací je **dotazníkové šetření**. Je ale nutné podotknout, že takové šetření může být ztížené velkým subjektivním faktorem a data takto vzniklá jsou obvykle přijímána velmi rezervovaně. Např. rozdáme všem studentům jedné třídy dotazník s pokynem, aby ho druhý den odevzdali vyplněný. Nikdy už nezjistíme, zda dotazník vyplnil opravdu každý z nich, nebo zda – z jakéhokoliv důvodu – všechny dotazníky ve skutečnosti vyplnil jen jeden nejaktivnější „zástupce třídy“. Jednou chybně sebraná data už budou vždy chybná a závěry z nich učiněné budou nesprávné, často aniž by to výzkumník vůbec tušil.*

2.3 Typ výzkumu

Na závěr je třeba rozhodnout, jakým způsobem budeme data zpracovávat. V zásadě je možné vydat se vydat dvěma směry – **vyjít z dat tak, jak jsou** (jak jsme je získali z již existujících databází) = observační studie nebo **provést experiment** (vnést do systému nový prvek – např. podat zkoumaný lék) a sledovat, jak se tento vliv projeví = experimentální studie:

Observační studie se dále dělí na studie popisné, které si kladou za cíl „jen“ konstatovat určitý fakt, a na studie analytické, které srovnávají více systémů mezi sebou.

Experimentální studie se dělí podle toho, zda zkoumaná osoba a/nebo výzkumník vědí, kdo je předmětem intervence a kdo není.

Souhrnná tabulka dělení studií

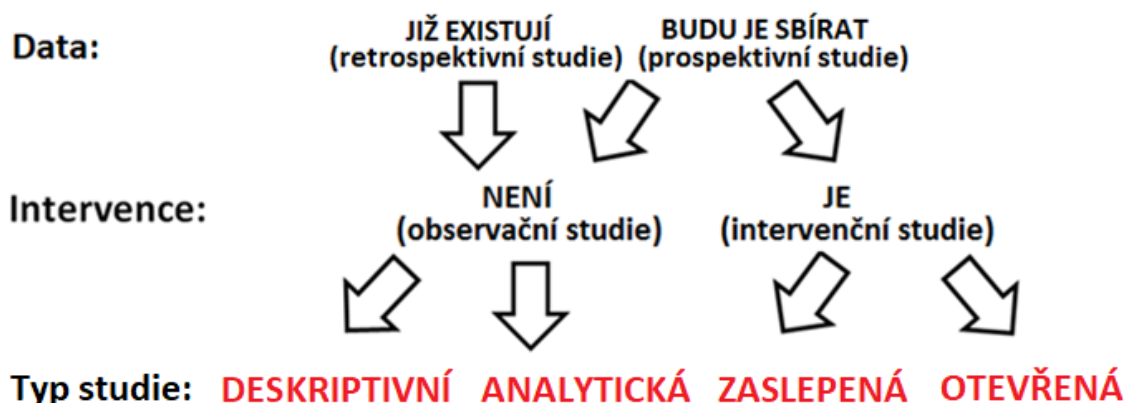
A) OBSERVAČNÍ STUDIE = popisují realitu tak, jak je, bez „umělého“ zásahu.

- **DESKRIPTIVNÍ (popisné) studie** = cílem je prostý popis systému
- **ANALYTICKÁ studie** = cílem je srovnání dvou (a více) systémů a vyhodnocení toho, zda se liší

B) EXPERIMENTÁLNÍ (INTERVENČNÍ) STUDIE = zkoumají vliv experimentálního zásahu. Cílem je zkoumání vlivu vnesení experimentálního prvku do systému a zhodnocení toho, zda tento prvek má měřitelný dopad na funkci systému.

- **OTEVŘENÁ studie** = subjekt i výzkumník vědí, kdo je předmětem výzkumu
- **ZASLEPENÁ studie** = subjekt a/nebo výzkumník nevědí, kdo je předmětem výzkumu

Určení typu studie podle času sběru dat a případné intervence



V každé práci musí být typ studie uveden – postupným vypsáním jednotlivých kritérií v pořadí podle tohoto schématu: – tj. např. „**RETROSPEKTIVNÍ OBSERVAČNÍ ANALYTICKÁ** studie“, „**PROSPEKTIVNÍ INTERVENČNÍ ZASLEPENÁ** studie“ apod.

Příklady:

Popisná (deskriptivní) práce = popisují vlastnosti určitého systému. Např. sledují vzdálenosti trvalého bydliště studentů od sídla školy. Výsledkem je konstatování skutečnosti – „je to takto“. Nic více. Z hlediska metodiky je popisná práce nejjednodušší.

Analytická práce = sledují rozdíly ve výskytu určitého jevu (např. tělesné hmotnosti) u dvou (a více) **různých experimentálních skupin**, vybraných podle určitých kritérií – např. ženy vs. muži. Základním cílem tohoto typu práce je konstatování, zda se obě experimentální skupiny ve sledovaném parametru liší dostatečně na to, aby bylo možné vyloučit, že jde pouze o náhodu. To může být (ale nemusí) doplněno kvantifikací – „jak moc“ se liší.

Poznámka: Posedlost tvorbou hypotéz „za každou cenu“ často vede ke snaze vytvořit z popisné práce práci analytickou. V praxi to vypadá tak, že studentovi „vyjde“ nějaký popisný výsledek (např. že 25 % studentů bydlí ve vzdálenosti větší, než 30 km), načež následně vytvoří hypotézu „Předpokládám, že více než 20 % studentů bydlí ve vzdálenosti větší, než 30 km“, a pak vyhodnotí, že „hypotéza je pravdivá“. Kromě absurdního vytvoření hypotézy „podle toho, co vyšlo“ je zásadní chybou i to, že výsledek v experimentální skupině neporovnáme s žádným standardem či kontrolní skupinou, ale s číslem vysloveným „jen tak“, bez odůvodnění toho, proč je to např. právě 30 km.

Experimentální práce = dtto jako analytické práce s tím rozdílem, že nezkoumám „stav jaký je“, ale rozdíl mezi skupinami se snažím dosáhnout vlivem určitého faktoru, kterému jednu skupinu vystavím a druhou ne (např. měření změny krevního tlaku po podání určitého léku vs. placebo).

Za nejkvalitnější jsou považované **prospektivní, intervenční randomizované zaslepené** studie, které ale obvykle přesahují organizační, technické i finanční možnosti autorů závěrečných prací.¹

¹ **prospektivní** = budu data teprve sbírat podle předem stanovené metodiky; **intervenční** = budu porovnávat vliv určitého faktoru; **randomizovaná** = ze základního souboru budu náhodně vybírat pacienty, kteří

3 Jdeme na to aneb jak vyrobit studii

Souhrn ve třech větách:

- Před zahájením práce bychom měli mít jasno nejméně v tom, jaký je **CÍL** práce (o co jde, proč to zkoumáme) a jaká bude **METODIKA** (jaká data a jak budete zpracovávat).
- Pokud zkoumáte nějaký rozdíl mezi dvěma skupinami, musí mít práce navíc předem určenou **HYPOTÉZU** (v čem očekáváte rozdíl mezi skupinami) a stanoveno, jakým způsobem provedeme **TEST ROZDÍLU** mezi skupinami (zda se ten rozdíl opravdu projevil, nebo ne).
- **Hypotéza by měla mít praktický smysl** – měla by být podkladem pro možný praktický závěr (změnu, nebo potvrzení správnosti stávajícího stavu).

Po základním rozhodnutí o cíli studie a zdroji dat můžeme přikročit k praktické realizaci. Příprava studie zahrnuje nejméně následující kroky:

DESKRIPTIVNÍ studie:

- **Stanovení CÍLE práce** = definice toho, CO je předmětem práce – viz výše
- **Stanovení METODIKY** = přesná definice toho, JAK budeme sbírat a hodnotit data
- **SBĚR DAT**
- **DISKUZE** = interpretace dat v kontextu dosavadních poznatků a praxe.

U ANALYTICKÉ a INTERVENČNÍ studie:

- **Stanovení CÍLE práce** = definice toho, CO je předmětem práce – viz výše
- **Stanovení HYPOTÉZY** = definice očekávaného výsledku experimentu nebo porovnání sledovaných skupin.²
- **Stanovení METODIKY** = přesná definice toho, JAK budeme sbírat a hodnotit data
- **SBĚR DAT**
- **TESTOVÁNÍ DAT** = rozhodnutí o přijetí, nebo zamítnutí hypotézy³
- **DISKUZE** = interpretace dat v kontextu dosavadních poznatků a praxe.

dostanou × nedostanou intervenci; **zaslepená** = já ani pacient nebudeme vědět, kdo intervenci dostal a kdo ne.

² Stanovení hypotézy je vždy spojeno s porovnáním nejméně dvou sledovaných skupin, tj. nemá smysl u deskriptivní práce. Přesto jsou do toho studenti často tlačeni...

³ Testování hypotézy nemá smysl u popisné práce – kde nejsou hypotézy, nemůžeme nic testovat. Testování také není potřeba u těch analytických studií, kde jsme získali data od úplného vzorku sledované populace – příkladem jsou např. volby.

Příklad 1 (analytická studie - zkoumání už existující databáze):

Cíl práce: posoudit dosažitelnost záchranné služby v obci Dolní Lhota

Hypotéza: dojezdová doba záchranné služby do obce Dolní Lhota (DL) odpovídá legislativě

Metodika: Jde o retrospektivní studii dojezdových časů do obce DL za období od 1. 1. xxxx do 31. 12. xxxx, přičemž dojezdovým časem se rozumí interval mezi časem výjezdu ze základny a časem dojezdu na místo. Data byla získána z databáze výjezdů ZZS Horní Dolní. Zařazeny byly všechny výjezdy do katastru obce DL v daném období s výjimkou výjezdů, kde některý z časů nebyl zaznamenán. Atd...

U intervenčních (experimentálních) studií bychom si, v souladu s mezinárodně uznávanou metodikou „PICO“, měli být schopni odpovědět na následující otázky:

P (patient) = kdo je náš pacient, neboli na jakou populaci se náš výzkum vztahuje;

I (intervention) = jakou intervenci (experimentální zásah) provádíme;

C(comparison) = s čím budeme porovnávat výsledek (jiný postup, podání placebo apod.);

O (outcome) = jak budeme hodnotit výstup/účinek.

Příklad 2 (experimentální studie):

Cíl práce: posoudit účinnost nového léku AnitAU (AA) na tlumení bolesti u pacientů se zlomeninou krčku stehenní kosti v přednemocniční péči

Hypotéza: podání jedné tabletky AA vede ke zmírnění bolesti u pacientů se zlomeninou krčku stehenní kosti v přednemocniční péči

Metodika: prospektivní pseudorandomizovaná studie, zkoumající změny na numerické škále bolesti (0 = žádná bolest, 100 = zcela nesnesitelná bolest) 5 minut po podání jedné dávky AA. Do studie byli zařazeni všichni pacienti starší 18 let s podezřením na zlomeninu krčku v období od ... do ..., s výjimkou pacientů udávajících alergii na některou ze složek přípravku. V lichých měsících byl lék AA podáván na stanovišti A, zatímco na stanovišti B byla poskytována standardní léčba, a v sudých měsících tomu bylo naopak. U pacientů byla zaznamenána hodnota na numerické škále bolesti před a 5 minut po podání AA. Atd...

Souhrn metodiky podle PICO: P= pacienti se zlomeninou krčku, I = podání AA; C = standardní léčba; O = rozdíl hodnot na numerické škále bolesti před a 5 minut po podání léčby.

4 Sběr dat

Souhrn ve třech větách:

- *Zásadní rozhodnutí spočívá ve **stanovení rozsahu šetření** – zda jsme schopni získat data od všech existujících subjektů zájmové populace, nebo ne (což je obvyklejší).*
- *Pokud nejsme, je naprosto zásadní **vhodně definovat, jak vybereme ty, od nichž budeme data sbírat** tak, aby tento malý vzorek (výběr) mohl **reprezentovat** celou zkoumanou skupinu (= byl pro celou zkoumanou populaci **reprezentativní**).*
- *Pokud provádíme nějaký **experiment**, je nezbytné zajistit, aby rozhodnutí o zařazení objektu do experimentální, resp. kontrolní skupiny bylo **co nejvíce náhodné**.*

4.1 Jaký bude rozsah šetření?

- **Úplné šetření** = sledovaná hodnota se zjišťuje u všech členů zkoumané populace. Výsledky jsou nejpřesnější, ale pak bývá drahé a někdy nerealizovatelné. Naopak masový laciný sběr dat bez přísného hlídání metodiky = riziko vysoce nepřesného výsledku. Pokud je vůbec úplné šetření technicky realizovatelné, je problém to, jak spolehlivě pohlídat kvalitu dat (což opět naráží na ekonomickou náročnost).
- **Výběrové šetření** = sleduje se pouze vybraný vzorek, který ale musí být v klíčových charakteristikách věrnou „zmenšeninou“ celku (musí být „reprezentativní“). Zkoumaný vzorek je menší a sběr dat o to pečlivější. Při chybném zvolení metodiky výběru existuje ale riziko zanesení výběrové chyby.

Např. náklady na volby (= úplné šetření) jsou o několik řádů vyšší, než náklady na předvolební průzkumy (= výběrové šetření), přičemž výsledky se zpravidla „až tak moc“ neliší, ale – s určitou pravděpodobností – se lišit mohou. Vzorky respondentů používané pro tyto průzkumy jsou ale jedním z nejvíce střežených tajemství agentur.

Naprostá většina odborných šetření v biologii je výběrových právě proto, aby měla smysl a praktický dopad. Nelze např. zavést novu léčbu všem pacientům s danou nemocí současně bez ověření, zda je bezpečná. Vyzkouší se proto na určitém vzorku populace a výsledky (např. nový způsob léčby) se vztáhnou na celou populaci. V řadě případů je celá

populace je díky biologickému rozvoji nekonečně velká a nezachytitelná, tj. úplné šetření není možné.

Největší problém výběrového šetření je vhodná definice a výběr vzorku (zařazovací × vylučovací kritéria). Jakmile je vzorek vybraný, měl by být hodnocený celý, i když ve skutečnosti se někdo nakonec výzkumu nezúčastní nebo v jeho průběhu odstoupí (princip „**intention to treat**“). Ignorování „odpadlých“ pacientů, neodevzdaných dotazníků či nezodpovězených dotazů může narušit původně definovaný vzorek, a tudíž způsobit výběrovou chybu, byť původní výběr byl reprezentativní.

Např. testovaný lék je sice účinný, ale části pacientů působí silnou nevolnost. Ti ho v průběhu studie přestanou brát a odstoupí ze studie. V takovém případě nelze hodnotit pouze výsledek u těch, kterým nevolnost nepůsobí a doberou ho – šlo by o výběrovou chybu, výsledky by byly falešně lepší, než ve skutečnosti – i v reálném životě u některých pacientů lék „nebude fungovat“ ne proto, že by skutečně nefungoval, ale proto, že ho pacienti přestanou brát.

Obdobně dotazníková šetření spíše vyplní ti, kteří jsou se stávajícím stavem nespokojení, osoby nějak zainteresované apod. Nízká návratnost dotazníku proto také velmi diskvalifikuje věrohodnost výsledku, protože diskredituje reprezentativnost vzorku – i když ze 100 oslovených odpoví 20, výsledná zjištění by měla být vztahovaná k původnímu souboru 100 osob (5 odpovědí ANO zde neznamená 5 z 20, ale „nejméně 5“ ze 100).

Způsob zařazování do výběru, výběrová a vylučovací kritéria, je velmi vhodné důkladně zvážit a prodiskutovat v co nejširším týmu, ideálně se zkušeným biostatistikem. Dopředu je vhodné zvážit i to, jak naložíme s nepravděpodobnými, nebo dokonce zjevně chybnými daty (např. zadávající VS se splete a zadá dojezdový čas o hodinu nebo o den později než ve skutečnosti).

Metodika výběru musí být zcela exaktně popsána v případné publikaci a bývá jedním z nejpozorněji čtených částí každé publikace.

4.2 Experimentální versus kontrolní skupina, randomizace

V případě experimentálních studií, kdy je nutné objekty studie (např. vhodné pacienty) rozdělit do dvou skupin. Ta, u které provedeme sledovanou intervenci (např. podáme lék, jehož účinnost sledujeme) je **experimentální skupina**, zatímco ta, kde intervenci neprovedeme, je tzv. **kontrolní skupina**.

Abychom co nejvíce odstranili subjektivní vliv v rozhodování, kdo bude či naopak nebude zařazen do experimentální skupiny, používá se zpravidla nějaká forma „losování“ - tzv. **randomizace**. Tím se předejde podvědomé snaze ošetřujícího personálu poskytnout tu léčbu, kterou považuje za lepší, nadějnějším pacientům - čímž by se zcela zkreslil výsledek studie.

Randomizace může být buď „pravá“ (obvykle se k ní používá speciální počítač, který online rozhodne a zařazení/nezařazení pacienta do zkoumané skupiny), nebo lze použít tzv. „**pseudorandomizaci**“ - rozdělení pacientů podle určitého, předem daného klíče, u kterého nepřepokládáme vliv na prováděný experiment (např. v sudé datum se volí experimentální postup, v liché standardní apod.).

*Přes veškerou snahu o co nejlepší dodržení metodiky se někdy stane, že se výzkumník rozhodne pacienta zařadit (nebo naopak vyřadit) v rozporu s výsledkem randomizace. Tento tzv. „**crossover**“ představuje závažný metodický problém, při větším výskytu znehodnocující studii jako celek. Výskyt „crossoveru“ nemá žádné dobré řešení - je totiž chybou takové pacienty ve studii ponechat (nejsou do experimentální nebo kontrolní skupiny zařazení náhodně, ale podle určité vlastnosti), ale současně by byla chyba je ze studie vyřadit, protože ze základního souboru tak vypadává určitý typ pacientů a základní soubor tím pozbývá reprezentativnosti.*

4.3 Na co si dát pozor

V jednoduchosti je síla. V jedné studii je vhodné soustředit se na zkoumání co nejmenšího množství kritérií, ideálně jen jednoho. Velmi důležité (a jen zdánlivě

samozřejmě) je, aby zkoumané kritérium skutečně popisovalo to, o čem má být studie, resp. co nás zajímá.

Pokud chci zkoumat znalosti první pomoci, může jít např. o vhodný dotazník, ale také o komplexní hodnocení praktických schopností u simulovaných scénářů. Časová i finanční náročnost bude pochopitelně zásadně odlišná.

Kvalita dat je klíčová. Čím kvalitnější data máme, tím pevnější budou naše výsledky. Na druhou stranu kvalita vždy něco stojí – ať už vyjádřeno v penězích, nebo v množství vynaložené práce. Výsledkem je proto obvykle „vhodný kompromis“, zajišťující přiměřenou úroveň **validity** (námi zvolená metodika měří to, co chceme změřit) a současně **reliability** (při opakovaném použití metody dostaneme vždy stejný výsledek). Všechny známé kompromisy by ovšem měly být zmíněny v následné publikaci.

Praktické poznámky ze života:

Příprava sběru dat tak, aby „to doopravdy fungovalo“, je jednou z nejnáročnějších částí výzkumu a obvykle se nepodaří „na první dobrou“. Je proto velmi vhodné metodiku vyzkoušet na nějakém malém vzorku, a pak upravit podle praktických zkušeností – řada okrajových variant a možností nás obvykle při tvorbě metodiky vůbec nenapadne (= tzv. „pilotní studie“).

Zásadním požadavkem je výběr vhodné metriky – tj. toho, co jak vlastně budeme měřit požadovanou vlastnost. Typickým příkladem nevhodné metriky je použití dojezdové doby záchranné služby jako ukazatele kvality péče – ačkoliv pro drtivou většinu stavů nemá kratší dojezdová doba žádný reálný význam.

Nespoléhejte na to, že „za pět minut dvanáct“ prostě rozešlete dotazníky, a pak výsledky „nějak“ zpracujete. Počítejte s tím, že vám podobných jsou desítky a ochota „oblíbených“ cílových skupin odpovídat na neustále se opakující otázky je nízká. Vůbec nejnižší je pak ochota sepisovat odpovědi na volné (otevřené) otázky – sepisování jakýchkoliv textů „svými slovy“.

Nic není jasnějším projevem chybné přípravy studie než to, když měsíc před termínem uzavření práce přichází na určité pracoviště žádost o poskytnutí dat, které student nutně potřebuje, ale nemá k nim přístup, případně žádá o provedení šetření mezi zaměstnanci organizace. Je nutné si uvědomit, že data „nerostou na stromě“ – je nutné je (obvykle velmi pracně) sbírat a mají tudíž svoji faktickou cenu. V „ideálním případě“ je navíc žádost doprovázena komentářem typu „sháním NĚJAKÁ data na toto téma“, bez jasné představy, co s nimi vlastně bude dělat a jaký praktický dopad bude mít výsledek práce. Pokud se v organizaci sejde větší množství požadavků, navíc podložených ne zcela smysluplnými cíli práce, je možné s velkou pravděpodobností očekávat, že žadatel bude odmítnut. To není projev neochoty, ale zcela racionální rozhodnutí založené na zvážení poměru přínosu a materiálních i lidských nákladů. Autor práce se tak může ocitnout v nezáviděníhodné situaci.

5 Zdroje chyb

Souhrn ve třech větách:

- Jakákoliv závažná **metodická chyba** činí vaši práci nevěrohodnou.
- Veškerá metodika musí být nastavena tak, aby **známé zdroje chyb byly co nejvíce eliminovány** a ty, které nejsou eliminovány, musí být předmětem vysvětlení.
- Typickými zdroji chyb je **výběrová chyba** (nenáhodný výběr zkoumané populace), **chybné měření** (nepřesnosti v měření dané reálnými možnostmi, ale i omyly v záznamech dat apod.), „**placebo efekt**“ (vím, že jsem léčen, tak se mi daří lépe) „**hawthornský fenomén**“ (vím, že jsem zkoumán, tak pracuji lépe) a **osobní vliv výzkumníka** (je to moje (ne)oblíbená metoda, tak musím prokázat, že (ne)funguje).

Z různých důvodů je možné, že dojde k úmyslnému či podvědomému ovlivnění výsledku studie.

Nejobvyklejší důvody zkreslení jsou následující:

5.1 Výběrová chyba

Tento typ chyby je způsoben tím, že jednotliví členové sledované populace nemají stejnou šanci dostat se do vzorku. Bývá to častý problém u tzv. „pilotních studií“ – tyto studie mají sloužit k ověření proveditelnosti metodiky, nikoliv k ověření funkčnosti metody jako takové.

Příklady výběrové chyby z praxe:

- „Pilotní studie“ např. sledující možnosti určité operační metody, které se účastní jen ti lékaři, kteří o to mají zájem = pravděpodobně „ti schopnější“, kteří mají obecně méně komplikací, ale výsledek je vztažený na celou populaci potenciálních uživatelů;
- Rozdám dotazníky svým kamarádům = vypadnou respondenti, kteří nemají podobný názor, jako já (ti pravděpodobně nebudou mezi mými kamarády);

- Z telefonického průzkumu automaticky vypadává ten, kdo nepoužívá telefon (např. osoby s hendikepem sluchu);
- Novou (a drahou) léčebnou metodu podvědomě aplikují spíše pacientům, kteří se mi jeví jako „perspektivní“.

Výběrová chyba je „nevratná“ (data od subjektů, které se do výzkumu vůbec nedostaly, už nikdy neseženeme) a proto je nutné ji co nejvíce eliminovat (správná volba vzorku, striktní dodržení kritérií výběru, zaslepení studie apod.). **Nekontrolovaná výběrová chyba je jedním z největších „prohřešků“, které mohou jinak kvalitní studii zcela zásadně znevěhodnit.**

5.2 Chybné měření

Přestože v reálném světě měříme (hodnotíme) obrovské množství parametrů, ve skutečnosti jsou naše měření zatížena množstvím chyb. Ve studii bychom samozřejmě měli dbát na to, abychom nepřesnosti omezili na minimální dosažitelnou míru.

Typickými příčinami chyb měření jsou:

- **náhodná chyba** (chance) – např. měřím výšku pacienta a z praktických důvodů ji vždy zaokrouhlím na celé centimetry. Chyby jsou pravděpodobně náhodně rozvrstvené do „+“ i „-“, ukazatel polohy (např. průměr) bude odpovídat realitě.
- **systémová chyba** (bias) – např. měřím výšku pacienta, ale vždy v obuvi. Chyba je vždy „stejným směrem“ a výsledek nebude odpovídat realitě.
- **outlier** je hodnota, která se zcela zásadně vymyká ostatním naměřeným hodnotám. Jde o zachycení nějaké neočekávané, výjimečné situace (anebo třeba o chybu zadání), která je sama o sobě nevýznamná, ale zásadně zkreslí výpočet parametrů souboru (např. průměr).

Náhodná a systémová chyba jsou celkem běžné. Jejich příčinou je často nemožnost zpřesnit měření (obvykle z technických či ekonomických důvodů, ale také „placebo efekt“ a další subjektivní fenomény – viz výše). Lze se s nimi vypořádat –

počítat s nimi, a případně jejich důsledky vhodným způsobem eliminovat (např. opakovaným měřením, použitím vhodné korekce výpočtu apod.)

Riziko chyby dané zanesení outlierů je nejlépe eliminovat stanovením určitých hranic „pravděpodobného výsledku“, ve kterých se pohybujeme v praktickém životě, a hodnoty mimo tento interval zanedbáme (což je samozřejmě nutné popsat v metodice práce).

Outlierem je např. dojezdový čas záchranné služby v trvání 3 hodiny – v důsledku úplného zhroucení dopravy při sněhové bouři, přičemž obvyklé dojezdové časy do dané lokality se pohybují kolem 10 minut.

5.3 Placebo efekt

Objekt se chová jinak, pokud ví, že právě on je předmětem zkoumané intervence. Tam, kde lze intervenci simulovat (např. podáním stejné tabletky, ale bez obsahu účinné látky), je možné „placebo efekt“ eliminovat tzv. „**zaslepením**“ studie, kdy účastníci nevědí, zda jsou předmětem skutečné, nebo pouze simulované intervence:

- jednoduchý slepý pokus = sledovaná osoba to neví;
- dvojitý slepý pokus = sledovaná osoba ani intervent to neví;
- trojitý slepý pokus = sledovaná osoba, intervent ani hodnotitel to neví.

V řadě případů ale intervenci v praxi simulovat nelze (např. podání x nepodání defibrilačního výboje, provedení x neprovedení určité operace apod.)

Mechanismy „zaslepení“ jsou různé – příkladem může být např. postup, kdy tabletky testovaného léku i placebo (neúčinná látka) má stejnou podobu a jsou dodávány v identických baleních, lišících se jenom číslem šarže. Pouze podle tohoto čísla lze zjistit, co vlastně pacient dostal. Tabulku s určením „co je co“ ale nemá ani podávající lékař a až při vyhodnocení vlivu léku se na závěr „odtajní“, jaký pacient byl doopravdy léčen a jaký dostával pouze placebo.

Pokud by mohl být experiment nebezpečný, je nutné, aby výsledky průběžně kontroloval nezávislý bezpečnostní výbor, který se skládá z osob stojících zcela mimo výzkumný tým a svá zjištění udržuje v tajnosti. Pokud by ovšem zjistil např. významně větší množství

závažných komplikací (nebo naopak zjevně zásadně lepší výsledky) u léčených pacientů, je jeho právem studii zastavit.

5.4 Hawthornský fenomén

Objekt se chová jinak, pokud ví, že je zkoumán – zdravotník přesněji dodržuje předepsané postupy, pracuje pečlivěji, rychleji apod. To lze jej eliminovat skrytým sledováním (což může narazit na limity etiky), nebo delší dobou provádění studie (což časem otupí pozornost, ale může to být nákladné a nepraktické).

5.5 Osobní zájem/nezájem

Snaha podpořit „svoji“ metodu, vedená snahou získání odborné prestiže, komerčním zájmem apod. Lze ovlivnit tím, že autor „nápadu“ není přímo účasten tvorby metodiky, sběru a vyhodnocování dat, sestavením nezávislé dozorčí rady apod.

Ale i naopak – nezájem, nízká motivace může vést k tomu, že data jsou zaznamenávána „hala-bala“, „přibližně“, „aby se tam něco vyplnilo“, případně jsou výsledky rovnou vymyšlené.

Obvykle platí, že čím masověji sběr dat probíhá („povinné zadávání do počítače“), tím nižší kvalita dat. Velké, masově vedené databáze jsou proto vhodné spíše pro hrubé hodnocení trendů (předpokládáme, že chybovost se příliš nemění) a pro pilotní studie, kdy by hypotézy měly být ověřeny prospektivní, kontrolovanou studií.

Etická poznámka: není přípustný experiment založený na cíleném zhoršení stavu, u otevřených studií obvykle ani aktivní motivace účastníků k dosažení pozitivního výsledku apod. U intervencí s možnými závažnými negativními důsledky je nutná průběžná kontrola a předčasné zastavení studie, jakmile je statisticky jasný pozitivní/negativní vliv.

6 Základy měření a zpracování biomedicínských dat

Souhrn ve třech větách:

- Pro pozdější statistické vyhodnocení je dobré si uvědomit, jaký **typ dat** budu sbírat.
- Nejlépe se vyhodnocují **NOMINÁLNÍ** data (ANO × NE, ŽENA × MUŽ apod.) a **METRICKÁ** data (konkrétní číslo – např. výška postavy, nebo relativní číslo – např. podíl v %); naopak špatně se vyhodnocují **ORDINÁLNÍ** data (souhlasím – spíše souhlasím – nevím – spíše nesouhlasím – nesouhlasím apod.).
- Metrická data jsou často reprezentovaná pomocí **průměru**, což může být silně zavádějící v případě, že jednotlivé prvky souboru jsou velmi rozdílné (viz „já mám celé kuře a ty nic – v průměru máme oba půlku“).

6.1 Typy dat

Aby mohla být jakákoliv data vyhodnocena, je nutný jejich strukturovaný popis (zařadit jednotlivá měření / případy / odpovědi) do nějaké „škatulky“. K tomu je možné použít některý z následujících způsobů:

- **nominální** data (popis rovnocennými kategoriemi) – např. MUŽ × ŽENA, im. × iv., ANO × NE apod.
- **ordinální** data (popis stupňujícími se kategoriemi) – např. „ne -> spíše ne -> spíše ano -> ano“
- **metrická** data – absolutní čísla (např. tělesná výška pacienta) resp. relativní (poměrová) data (např. škála bolesti (0 % – 100 %))

Z hlediska následného vyhodnocení jsou nejlépe zpracovatelná nominální a číselná data, zatímco statistické zpracování ordinálních dat je obtížnější, obvykle spojené s poněkud problematickým převodem ordinální stupnice na relativní číslo.

6.2 Ukazatelé polohy a variability

Naprostá většina studií vychází ze zpracování dat od většího množství zkoumaných subjektů.

Zatímco pro nominální a ordinální data neexistuje žádné souhrnné vyjádření, pro **metrická data** (výsledek vyjádřený plynulou číselnou škálou – např. výška pacienta) je možné použít vhodný matematický postup, který „reprezentuje“ souhrnnou charakteristiku souboru. Každý soubor je možné (více či méně výstižně) popsat pomocí **ukazatelů polohy** a **ukazatelů variability**.

Ukazatelé polohy charakterizují polohu hodnot souboru na pomyslné číselné ose:

- **Průměr** = součet dělený počtem měření. **Výhoda** = **jednoduchý** výpočet, široce používaný, snadno pochopitelný. **Nevýhoda** = **může být silně zavádějící**, pokud nemá soubor rovnoměrné rozložení (viz níže).
- **Medián** = střední hodnota v souboru (nebo průměr dvou středních hodnot při sudém počtu měření). **Nevýhoda** – **složitější výpočet** (nutnost setřídít soubor podle velikosti), zásadní **Výhoda** = **lepší popis nehomogenního vzorku (eliminuje extrém)**. Použití např. pro dojezdové časy ZZS.
- **Modus** = nejčastější hodnota v souboru. Použití zejména pro nominální a ordinální stupnice.

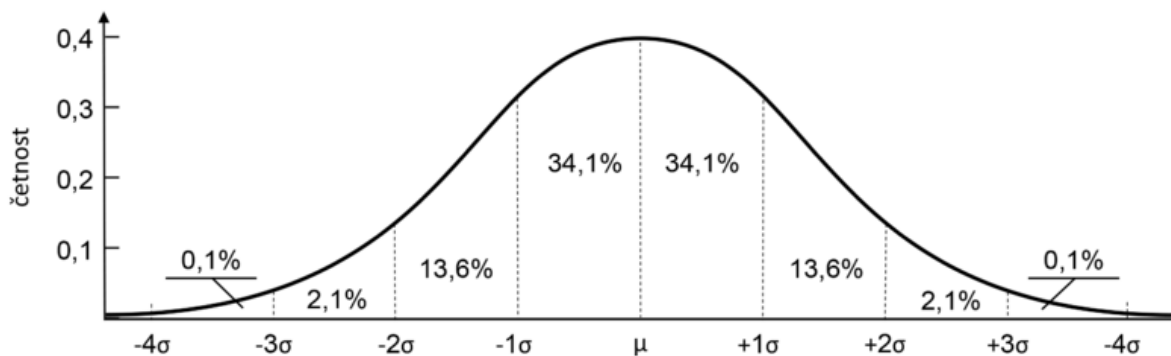
Ukazatelé polohy ovšem nevyprávějí nic o celkovém rozptylu (variabilitě) dat. Nedílným doplňkem učení „polohy“ musí být vždy určení „variability“:

Ukazatele variability charakterizují vnitřní variabilitu (šířku) souboru:

- **Směrodatná odchylka (SD)**. SD = „průměrná odchylka od průměru“, čím je větší, tím je soubor méně homogenní (více variabilní);
- **Rozsah** (od – do);
- **Percentil** (hodnota v určité proporci souboru – např. v 95 % rozsahu). Vhodná hodnota pro vyjádření určité „zaručené“ (nepodkročitelné) vlastnosti systému (např. 99 percentil = v 99 % případů je výsledek stejný nebo lepší).

Typicky používanou kombinací popisující data je použití průměru s uvedením směrodatné odchylky.

Obvyklou situací v biologii je to, že rozptyl dat je zhruba rovnoměrný kolem určité nejobvyklejší střední hodnoty = „normální“ (Gaussovo, náhodné) rozdělení.



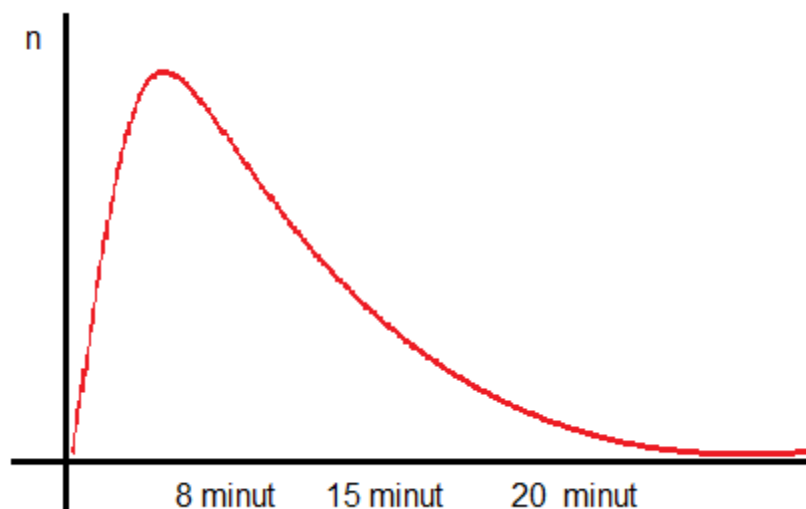
Graf 1: Normální rozdělení, průměr a směrodatné odchylky

Pro normální rozložení platí:

- průměr +/- 1 SD zahrnuje cca 70 % případů
- průměr +/- 2 SD zahrnuje cca 95 % případů
- průměr +/- 3 SD zahrnuje cca 99 % případů

Nenormální rozdělení

Ne vždy ale mají data „normální“ rozdělení – v praxi např. dojezdové časy záchranné služby (hodně hodnot kolem nejčastějšího času, ale pak poměrně velký rozptyl směrem k vyšším časům. **Použití průměru pro popis takového souboru není vhodné**, protože může být významně ovlivněn okrajovými hodnotami jedním směrem. Vhodnější je použití jiné metody (např. medián a 90. percentil apod.)



Graf 2: Nenormální rozdělení – dojezdové časy záchranné služby

7 Stanovení hypotézy a hodnocení její platnosti

Souhrn ve třech větách:

- Hypotézy je třeba **předem** definovat **u každé analytické nebo intervenční studie** (a naopak nemají smysl u popisných studií) přičemž jsou vždycky dvě – „nulová“ **hypotéza** říká, že zkoumaná intervence nemá žádný vliv, zatímco „**alternativní**“ **hypotéza** říká, že intervence vliv má.
- **Statistické testování** je matematický postup, který umožní určit, která z hypotéz platí.
- Testování určuje, s jak velkou **pravděpodobností** je vliv intervence **náhodný**, přičemž jako hranice pro rozhodnutí o platnosti (neplatnosti) hypotézy se používá míra pravděpodobnosti 5 % (tj. aby byla alternativní hypotéza považovaná za platnou (např. „lék funguje“), **rozdíl mezi kontrolní a experimentální skupinou musí být s pravděpodobností 95 %, způsobený experimentálním vlivem, nikoliv náhodou.**

7.1 Co je to hypotéza?

Hypotéza je myšlenka, jejíž platnost se předpokládá, ale zároveň je formulovaná tak, aby ji bylo možno potvrdit nebo vyvrátit.⁴

Ve vědeckém světě platí konvence, že:

- **základní hypotéza se označuje jako tzv. „nulová“ (H_0)**, a ta vždy říká, že zkoumaná intervence nemá žádný vliv (tj. případně zjištěné rozdíly **JSOU náhodné**, jev není závislý na provedené intervenci, změřené hodnoty jsou „normální“ apod.).
- **alternativní hypotéza (H_1)** = to, co platí, pokud H_0 neplatí – tj. rozdíly **NEJSOU náhodné**, jev JE závislý na intervenci, stav **NENÍ** normální apod.

Cílem studie je prokázat, zda platí nulová nebo alternativní hypotéza.

Hypotéza má formu oznamovací věty, které konstatuje vztah mezi dvěma situacemi, které se vzájemně vylučují.

⁴ „...je formulovaná tak, aby ji bylo možno potvrdit nebo vyvrátit“ je velmi významný – vyjadřuje racionální podstatu vědy. Nemá smyslu vědecky se zabývat něčím, co nejde dokázat (viz pověstné „Kolik andělů se vejde na špičku jehly?“).

Příklad:

CÍL STUDIE = Zjistit vliv podání léku L na hodnotu systolického krevního tlaku.

HYPOTÉZY:

Základní (nulová) hypotéza H_0 = Hodnota sTK se po podání léku L významně nezměnila.

Alternativní hypotéza H_1 = Hodnota sTK se po podání léku L významně změnila.

7.2 Kdy se vůbec zabývat hypotézou

Hypotézou se zabýváme u těch studií, která porovnávají dva (event. více) výsledků. Stanovit hypotézu naopak nedává smysl u DESKRIPTIVNÍCH studií – tj. tam, kde je cílem práce pouze popsat určitou skutečnost.

Příklad – srovnej:

- *„Cílem studie je zjistit průměrnou tělesnou hmotnost žáků ve třídě.“ Toto je DESKRIPTIVNÍ studie, nic neporovnává, pouze konstatuje. Zde nemá smysl vytvářet hypotézu, resp. žádná hypotéza neexistuje.*
- *„Cílem studie je zjistit průměrnou tělesnou hmotnost žáků ve třídě a porovnat ji s celostátním průměrem z roku 2015.“ Toto je ANALYTICKÁ studie a stanovení hypotézy zde má smysl.*
- *„Cílem studie je zjistit průměrnou tělesnou hmotnost žáků školy před a po zavedení dietních opatření.“ Toto je EXPERIMENTÁLNÍ studie a stanovení hypotézy zde má smysl.*

7.3 Kvalitativní stanovení rozdílu – je rozdíl dílem náhody?

Cílem vědecké práce je nejčastěji zkoumání určitého rozdílu mezi dvěma skupinami – „experimentální“ a „kontrolní“. Pokud bychom prováděli **úplné šetření** (např. výzkum, zda chlapci ve třídě 5. A jsou v průměru vyšší než dívky), snadno změříme všechny žáky a na závěr můžeme s jistotou konstatovat, zda průměrná výška chlapců a dívek je či není rozdílná. Statistické testování u úplného výběru není potřebné.

V biologii ovšem v naprosté většině případů pracujeme s **výběrovým šetřením**, tj. s určitým vzorkem celé populace. Vzhledem k přirozené variabilitě v biologii téměř nikdy neplatí, že stejná intervence má stejný důsledek u všech jedinců. Z tohoto důvodu

je nutné vždy počítat s určitou mírou pravděpodobnosti, že zjištěné rozdíly jsou dílem náhody.

V biologických výzkumech se obvykle stanovuje přijatelná pravděpodobnost chyby v podobě „**hladiny významnosti α (alfa)**“. **Běžně používaná hodnota $\alpha = 0,05$** , tj. 5 %. Znamená to, že za rozdílný lze výsledek považovat tehdy, pokud je s pravděpodobností nejméně 95 % nenáhodný (takže zatímco s pravděpodobností nejvýše 5 % jde o náhodu). Méně často se používá hladina např. 0,01, tj. spolehlivost testu vyšší, než 99 %.

Pro reálný test je vypočtena konkrétní spolehlivost „p“ a ta je porovnána s očekávanou hladinou významnosti. **Pokud je „p“ menší, než „ α “, zamítáme nulovou hypotézu, tj. zjistili jsme statisticky významný rozdíl.** „**Síla testu**“ (spolehlivost) je potom definovaná jako 1-p, takže např. pokud testem vyjde $p=0,02$, je síla testu 0,98 (= 98 %).

Zajímavé je, že neexistuje žádná jednoznačná odpověď na otázku, proč se nejčastěji považuje za limitní hladina pravděpodobnosti právě 5 %. Toto tzv. „Fischerovo kritérium“ pochází z dvacátých let minulého století a je založeno výhradně na konsenzuálním akceptování vědeckým světem. V poslední době se ozývají poměrně silné hlasy pro přehodnocení tohoto přístupu, protože by mohl být považovaný za brzdu vědeckého pokroku: zatímco studie síly 4,99 je „významná“ a může vést k velkým změnám v léčbě, výsledek 5,01 „jakoby ani nebyl“, přestože z praktického ohledu jsou oba výsledky víceméně totožné.

Faktory ovlivňující sílu testu:

- **Velikost výběrového souboru** – čím více pozorování náhodné veličiny máme k dispozici, tím větší má test sílu (závislost není lineární, ale odvíjí se od konkrétního použitého testu).
- **Velikost pozorovaného rozdílu** – čím větší je rozdíl ve sledované veličině, tím větší síla testu.
- **Variabilita dat** reprezentovaná rozptylem náhodné veličiny – čím větší je rozptyl sledované veličiny, tím menší je síla testu.
- **Hladina významnosti testu** – standardně testujeme nulovou hypotézu na hladině významnosti $\alpha = 0,05$. Snížíme-li hladinu významnosti, tedy zvolíme-li např. hladinu

$\alpha = 0,01$, bude obtížnější zamítnout nulovou hypotézu, což znamená, že se sníží síla testu. Naopak zvýšení hladiny významnosti (což je ale spojeno s vyšším rizikem získání falešně pozitivního výsledku) znamená zvýšení síly testu.

Poznámka:

V případě sledování výsledků intervence je obvykle velmi významné i to, jak „široké“ rozmezí zvolíme pro průkaz, že daná intervence má požadovaný efekt (pokud nejde pouze o binární měření – ANO × NE). Čím užší je rozmezí, tím jistější bude výsledek, ale také tím větší bude pravděpodobnost, že „nesprávně“ vliv intervence neprokážeme – a opačně. Používají se zde pojmy „chyba prvního typu“ (zamítneme nulovou hypotézu, přestože platí = intervenci prohlásíme za funkční, přestože nefunguje) a „chyba druhého typu“ (přijmeme nulovou hypotézu, přestože neplatí = nerozpoznáme funkční intervenci). Chyba 1. typu a chyba 2. typu spolu souvisí tak, že čím menší je pravděpodobnost jedné, tím větší je pravděpodobnost druhé.

Příklad:

Zkoumáme, zda podání daného léku vede ke snížení krevního tlaku. Čím menší rozdíl tlaků před a po intervenci si stanovíme jako průkaz úspěchu intervence, tím vyšší je pravděpodobnost, že prokážeme vliv nového léku, ale současně i toho, že skutečnosti výsledek nebude odpovídat realitě. Jinak řečeno, budeme-li za „úspěch léčby“ považovat snížení TK o 10 torr, pravděpodobně prokážeme úspěch léčby snáze, než pokud si stanovíme požadované snížení nejméně např. o 30 torr. Na druhé straně v prvním případě se může stát, že intervence ve skutečnosti nebude funkční a rozdíly budou dílem náhody.

Rozhodnutí o zamítnutí/nezamítnutí nulové hypotézy je vlastně rozhodnutím o tzv. **statistické významnosti** rozdílu dvou nebo více výběrových souborů ve sledované náhodné veličině. Praktická využitelnost pozorovaných hodnot odpovídá naopak tzv. **věcné (praktické, klinické, biologické) významnosti** výsledku, která ale nemusí vždy odpovídat významnosti statistické. Vzhledem k tomu, že testování statistických hypotéz vždy provádíme kvůli možnosti zobecnění z náhodného výběru na celou populaci, je ověření interpretační hodnoty výsledku minimálně stejně významné jako vlastní výpočet testu.

7.4 Kvantitativní stanovení rozdílu – „jak moc“ je prokázán rozdíl velký?

Kvantitativní stanovení rozdílu nemusí být vždy nutné – v řadě případů postačí rozhodnutí, zda má (nebo nemá) zkoumaná intervence významný vliv. V jiných případech ale může být i kvantifikace rozdílu zajímavá (např. k odpovědi na otázku „kolikrát daná intervence zvyšuje naději na přežití?“).

K danému účelu se používají nejčastěji dva parametry:

- **ODDS RATIO** (OR – poměr šancí) = odhad asociace mezi „příčinou“ a „následkem“ (např. diagnózou).

Typické použití OR je u studií typu „case-control“. Parametr udává sílu asociace mezi příčinou a následkem. **Pokud je OR > 1, tak je asociace mezi „Diagnózou“ a „Příčinou“**

Výpočet „ODDS RATIO“	Očekávaná příčina ANO	Očekávaná příčina NE
Následek ANO	a	b
Následek NE	c	d
OR = (a/c) / (b/d)		

- **RELATIVE RISK** (RR – relativní riziko) = odhad relativního rizika

Vyjadřuje, o kolik je vznik následku pravděpodobnější v jedné vůči druhé skupině. Pokud by byla pravděpodobnost stejná, platí, že RR = 1. Hodnota RR např. 1,67 znamená, že v jedné skupině je pravděpodobnost následku o 67 % vyšší, než ve druhé.

Je dobře připomenout, že v reálném světě může „následek“ nastat i bez intervence – např. pacient se může uzdravit i spontánně – bez léčby.

Příklad – výpočet zvýšení naděje na uzdravení pacienta při podání léčby ve srovnání s nadějí na uzdravení bez léčby:

Výpočet „RELATIVE RISK“	Léčba ANO	Léčba NE
Uzdraven ANO	a	b
Uzdraven NE	c	d
RR = (a/(a+c)) / (b/(b+d))		

Kvalitativní stanovení rozdílu přesahuje náplň této publikace a nadále se mu nebudeme věnovat.

8 Statistické testování v typických příkladech

Souhrn ve třech větách:

- *Základní testy jsou dostupné na internetu online a jejich použití je velmi snadné i pro neškoleného zájemce.*
- *Pokud chcete otestovat rozdíl v **NOMINÁLNÍCH** datech, použijte chí-kvadrát test, pokud chcete otestovat rozdíl mezi **ČÍSELNÝMI** daty, použijte T-test a pokud chcete otestovat závislost mezi číselnými řadami, použijte **Pearsonův koeficient**.*
- ***Pokud pracujete s procenty**, samotné porovnání dvou číselných hodnot procent nestačí – musíte znát velikost vzorku, ze kterého byla procenta vypočtena (tj. zdrojová data pro výpočet procent) a poté použít vhodný (obvykle chí-kvadrát) test.*

8.1 Cíl testování a základní informace

Cílem statistického testování je určení toho, zda **rozdíly ve zkoumaných datech jsou nebo nejsou dílem náhody** – jinak řečeno, zda platí hypotéza H_0 , nebo H_1 . Z toho vyplývá, že statistické testování nemá smysl **u popisných studií**, kde výsledek pouze konstatujeme.

U analytických a intervenčních studií, kdy je nutné vzájemně porovnat dva či více výsledků vzájemně, se používá speciálních matematických postupů, „**testů**“ (např. T-test, chí-kvadrát test, Fischerův test apod.).

Poznámka:

Samotné vyjádření případného výsledku v procentech (%) je statisticky nehodnotitelné, pokud neznáme absolutní čísla. Tj. např. sdělení „75 % respondentů podporuje určitou myšlenku“ nelze interpretovat tak, že tuto myšlenku podporuje **většina** respondentů, pokud neznáme jejich konkrétní počet. Např. pokud by respondenti byli pouze 4, o validní (nenáhodný) výsledek by nešlo, zatímco pokud by jich bylo 100, rozdíl by validní byl.

Vždy je více než vhodné **konzultovat design studie a vhodnou metodu vyhodnocení výsledků s biostatistikem**. Přesto se můžeme, ať už ze zájmu nebo při nedostupnosti konzultace, pokusit o jednodušší testování sami.

Provedení statistického testu je poměrně složitá matematická úloha. Naštěstí v současnosti existují webové testy, které – po zadání korektních hodnot – provedou potřebné výpočty za nás.

8.2 Nejjednodušší, nejobvyklejší používané statistické testy

A) TESTY PRO NOMINÁLNÍ DATA

Zkoumám data, definovaná určitými konkrétním nečíselným popisem (typu ANO × NE, SPRÁVNÁ ODPOVĚĎ × CHYBNÁ ODPOVĚĎ, PLATÍ × NEPLATÍ, NEMOCNÝ × ZDRAVÝ...). Zajímá mne, zda se v jedné (zkoumané) skupině odpovědi liší od druhé (kontrolní) skupiny.

- **Chí-kvadrát test dobré shody** – pokud chci otestovat, jednotlivé možnosti v souboru vyskytují s očekávanou pravděpodobností (např. muži/ženy zaujímají rovnoměrně 50 %).

Příklad situací, kdy lze test využít:

- Zkoumání, zda se se mezi určitou profesní skupinou vyskytují rovnoměrně muži a ženy
- Zkoumání, zda určitá odpověď v dotazníku převažuje nad ostatními možnostmi.
- **Chí-kvadrát test pro kontingenční tabulku** – pokud chci otestovat, zda se určitý jev vyskytuje se stejnou (ovšem mě neznámou) pravděpodobností v různých zkoumaných skupinách.

Příklad situací, kdy lze test využít:

- Zkoumání, zda se určitá nemoc vyskytuje se stejnou pravděpodobností u mužů, jako u žen;
- Zkoumání, zda ve skupině očkovaných a neočkovaných dochází ke stejnému výskytu sledované nemoci;
- Zkoumání, zda studenti proškolení v poskytování první pomoci častěji uspěli při hodnocení zásahu u simulované nehody než studenti neproškolení.

B) TESTY PRO ČÍSELNÁ DATA (řady měření)

Zkoumám data, definovaná změřenou hodnotou, číslem (výška, váha, krevní tlak, reakční čas záchranné služby, výsledek laboratorního měření...).

- **T-test** – pokud chci testovat naměřenou sadu dat a zkoumat, zda se odchyluje od známé normy (jednovýběrový test), nebo od jiné sady dat (dvouvýběrový test).

Příklad situací, kdy lze test využít:

- Zkoumání, zda podání určitého léku vede k normalizaci hodnoty systolického krevního tlaku (jednovýběrový test – srovnání hodnoty sTK po léčbě se známou normou 120 mm Hg)
- Zkoumání, zda podání určitého léku vede ke snížení hodnoty systolického krevního tlaku ve srovnání s neléčenou populací (dvouvýběrový test)
- **Pearsonova korelace – korelační koeficient** – pokud u každého jedince měřím dvě kritéria a chci zjistit, zda jsou na sobě závislá

Příklad situací, kdy lze test využít:

- Zkoumání, zda délka ošetření pacienta závisí na délce praxe záchranáře;
- Zkoumání, zda hmotnost jedince závisí na jeho výšce apod.

Výsledkem výpočtu je kritérium r v hodnotě mezi 0 (žádná závislost) až +1 (závislost pozitivní = čím více A, tím více B) resp. -1 (závislost negativní = čím více A, tím méně B)

Zatímco výsledkem ostatních testů je hodnota „ p “ a výrok o přijetí či odmítnutí nulové hypotézy, pro vyhodnocení síly závislosti lze v nejjednodušším případě použít „metodu rozměrového efektu“. Arbitrálně stanové limity jsou následující:

- do 0,1 = žádná závislost
- 0,1 – 0,3 = malá závislost
- 0,3 – 0,5 středně silná závislost
- nad 0,5 = silná závislost.

8.3 Výběr vhodného testu v typických situacích

S výběrem testu může pomoci např. online průvodce na této webové stránce:

[http://www.socscistatistics.com/tests/what stats test wizard.aspx](http://www.socscistatistics.com/tests/what%20stats%20test%20wizard.aspx)

Příklady vhodných testů při typických situacích v jednoduchých studiích:

Situace 1

- **Problém:** Házím mincí a zkoumání, zda se rovnoměrně objevují obě možnosti (panna x orel)
- **Nulová hypotéza:** výskyt daných možností je náhodný
- **Vhodný test:** Chí kvadrát test dobré shody
- **Tabulka dat:**

	ZJIŠTĚNÝ VÝSKYT	OČEKÁVANÁ FREKVENCE
PANNA	N1	0,5
OREL	N2	0,5

Online test (až pro 5 kategorií) viz např. zde:

<http://www.socscistatistics.com/tests/goodnessoffit/Default2.aspx>

Situace 2

- **Problém:** chci otestovat, zda některá z odpovědí na otázku dotazníku převažuje.
- **Nulová hypotéza:** výskyt jednotlivých odpovědí je rovnoměrný
- **Vhodný test:** Chí kvadrát test dobré shody
- **Tabulka dat:**

	ZJIŠTĚNÝ VÝSKYT	OČEKÁVANÁ FREKVENCE
Odpověď A	N1	0,25
Odpověď B	N2	0,25
Odpověď C	N3	0,25
Odpověď D	N4	0,25

Online test (až pro 5 kategorií) viz např. zde:

<http://www.socscistatistics.com/tests/goodnessoffit/Default2.aspx>

Poznámka:

Tento test je možné provést i bez výpočetní podpory – kritické (hraniční) hodnoty lze vyčíst z dále uvedených tabulek (zdroj – Eurostat, $p=0,05$):

Rozmezí náhodné odchylky (hladina pravděpodobnosti 95%)

Očekávaný výsledek	Velikost vzorku (n)									
	10	50	100	150	200	400	800	1000	2000	4000
5%	13.5%	6.0%	4.3%	3.5%	3.0%	2.1%	1.5%	1.4%	1.0%	0.7%
10%	18.6%	8.3%	5.9%	4.8%	4.2%	2.9%	2.1%	1.9%	1.3%	0.9%
25%	26.8%	12.0%	8.5%	6.9%	6.0%	4.2%	3.0%	2.7%	1.9%	1.3%
50%	31.0%	13.9%	9.8%	8.0%	6.9%	4.9%	3.5%	3.1%	2.2%	1.5%
75%	26.8%	12.0%	8.5%	6.9%	6.0%	4.2%	3.0%	2.7%	1.9%	1.3%
90%	18.6%	8.3%	5.9%	4.8%	4.2%	2.9%	2.1%	1.9%	1.3%	0.9%
95%	13.5%	6.0%	4.3%	3.5%	3.0%	2.1%	1.5%	1.4%	1.0%	0.7%

Zdroj: Eurobarometer 2009

Tj. např. pro počet hodů 10 (viz sloupceček s číslem „10“ v záhlaví) musí být rozptyl od 50 % větší, než 31 %, tj. alespoň $50 + 31 = 81$ %, tj. v praxi 8× „panna“ a 2× „orel“ je stále ještě náhodná odchylka, resp. jinými slovy, není rozdíl mezi pravděpodobnostmi toho, že padne „panna“, nebo „orel“.⁵

⁵Z tabulky je mimo jiné na první vidět, jak moc záleží na počtu měření. Např. pro očekávaný výsledek 50 % (=model házení mincí) je při 10 hodech rozptyl +/- 30 % ve prospěch jedné strany mince stále ještě náhodný. Pro 100 hodů už je to pouze 8 % a pro 1000 hodů pouhých 3 %.

Situace 3

- **Problém:** Chci zjistit, zda je výskyt určité nemoci u očkovaných stejný, jako u neočkovaných.
- **Nulová hypotéza:** očkování nemá vliv na vznik určité nemoci
- **Vhodný test:** Fischerův exaktní test (pro menší množství měření – řádově desítky v každé kategorii), event. Chí kvadrát test pro kontingenční tabulku 2x2
- **Tabulka dat:**

	ONEMOCNĚL	NEONEMOCNĚL
OČKOVÁN	N1	N2
NEOČKOVÁN	N3	N4

Online Fischerův exaktní test pro tabulku 2x2:

<http://www.socscistatistics.com/tests/fisher/Default2.aspx>

Online CHI2 test pro obecnou tabulku 2x2 (2 kategorie a 2 varianty zkoumaného znaku):

<http://www.socscistatistics.com/tests/chisquare/Default2.aspx>

Situace 4

- **Problém:** chci zjistit, zda pacienti v určité skupině mají vyšší než normální systolický krevní tlak
- **Nulová hypotéza:** hodnota ve zkoumané populaci se neliší od známé normy.
- **Vhodný test:** jednovýběrový T-test
- **Tabulka dat:**

NAMĚŘENÁ HODNOTA sTK	ZNÁMÁ NORMA
Hodnota u pac. 1	Normální sTK
Hodnota u pac. 2	
Hodnota u pac. 3 atd.	

Jednovýběrový T-test online viz např.

<http://www.socscistatistics.com/tests/tsinglesample/Default.aspx>

Situace 5

- **Problém:** chci zjistit, zda pacientům po podání určitého léku pokles systolický krevní tlak.
- **Nulová hypotéza:** hodnota u léčených pacientů se neliší od hodnoty u neléčených
- **Vhodný test:** dvouvýběrový T-test
- **Tabulka dat (dvouvýběrový test):**

	LÉK PODÁN	LÉK NEPODÁN (ev. PLACEBO)
Naměřené hodnoty	Pac. test 1	Pac. kontrola 1
	Pac. test 2	Pac. kontrola 2
	Pac. test 3 atd.	Pac. kontrola 3 atd.

Dvouvýběrový T test online viz např.

<http://www.socscistatistics.com/tests/studentttest/Default.aspx>

Situace 6

- **Problém:** chci zjistit, zda vyšší dávka léku vede k vyššímu účinku (snížení sTK)
- **Nulová hypotéza:** dávka léku a síla účinku jsou nezávislé hodnoty
- **Vhodný test:** Pearsonův korelační koeficient
- **Tabulka dat:**

	Dávka léku (mg)	SNÍŽENÍ sTK (mm Hg)
Pacient 1	20	-10
Pacient 2	20	0
Pacient 3	40	-10
Pacient 4	40	-40
Pacient 5	60	-20
Pacient 6	60	-30 atd...

Pearsonův korelační koeficient – výpočet online:

<http://www.socscistatistics.com/tests/pearson/Default2.aspx>

9 Praktický „checklist“ pro přípravu studie

Základní kroky:

1. **Vznik potřeby** - studium literatury, pozorování, nápad
2. **CÍL PRÁCE** – co a proč budeme zkoumat
3. **HYPOTÉZY** (*pouze nejde o deskriptivní studii*)
4. **METODIKA** – jak budeme sbírat a vyhodnocovat data
 - a. **Co a jak budeme měřit**, aby data byla co nejvíce validní a reliabilní?
 - b. **Bude potřeba intervence?** Pokud ano:
 - bude potřeba souhlas etické komise (nestandardní intervence)?
 - budeme randomizovat (náhodně vybírat členy experimentální a kontrolní skupiny)?
 - bude možné studii zaslepit?
 - c. **Kolik dat budeme potřebovat**, aby byly výsledky validní („power analýza“)?
 - d. **Bude šetření úplné, nebo výběrové** – a pokud výběrové, jak budeme vybírat?
 - e. **Kde vzít data** – máme taková data, která budeme hodnotit (retrospektivní studie), nebo je budeme teprve sbírat (prospektivní studie)?
5. **SBĚR DAT**
 - a. Jak budeme uchovávat data – s ohledem na další zpracování?
 - b. Co dělat, pokud se data nepodaří zjistit? Co se zjevně nesprávnými hodnotami?
 - c. Primární data x odvozená data (faktická hodnota TK vs. normální x zvýšený). Lépe mít primární data (lepší kontrola konzistence).
 - d. Kombinovaná data -> více jednotlivých údajů
 - e. Desetinná tečka × čárka...
6. **ANALÝZA DAT** (*pouze nejde o deskriptivní studii*)
 - a. Výběr vhodného testu
 - b. Test platnosti hypotézy a konstatování zjištěného výsledku
7. **INTERPRETACE DAT** a srovnání se známými informacemi, diskuze o možných příčinách a důsledcích našich zjištění
8. **PUBLIKACE**

10 Publikace

Souhrn ve třech větách:

- **Standardní struktura každé práce** by měla být nejméně následující: **ÚVOD – METODIKA – VÝSLEDKY – DISKUZE – LITERATURA**.
- Kromě odborné stránky musí splňovat i **etické požadavky** – zejména nezkreslování výsledků, nekopírování cizích textů, řádné citace cizích myšlenek, nepřipisování autorů, kteří se na práci ve skutečnosti nepodíleli (často vedoucí pracoviště apod.) a deklaraci případného konfliktu zájmů (komerční podpory apod.).
- Dílo pak korunuje **bezchybné použití jazyka** – v jednoduchých a jasně srozumitelných větách.

10.1 Standardní struktura publikace

Základní struktura textu při publikaci výsledků studií by měla být následující:

Úvod – zasazení problému do kontextu systému, význam studie – co si od výsledku slibujeme.

Metodika – jaká data tvořila základní soubor, jak byla vybraná data pro studii (zařazovací a vyřazovací kritéria), jak byla data měřená, zpracovaná.

Výsledky – konstatování výsledků studie

Diskuze – obvykle nejdelší a nejzajímavější část. O čem svědčí výsledky, kde mohly být chyby, co chybí nebo přebývá, konfrontace s jinými podobnými studii a návrh vysvětlení případných rozporů, na co by bylo možné výsledky vztáhnout, jak by měl výzkum pokračovat dál.

Přehled literatury na dané téma, zejména té, na kterou se v textu odvoláváme nebo ze které vycházíme.

10.2 Etika publikace

Výsledky nesmí být zkreslené a tendenčně upravené (zamlčené např. důležité potíže při sběru dat, které by mohly vést k zanesení výběrové chyby apod.)

Měly by být publikovány i negativní výsledky (ve studii nedošlo k potvrzení hypotézy) – toto je často problém u studií, které mají za cíl prokázat např. účinek léků

nebo přístrojů (proto je u „vážně míněných“ studií obvyklá veřejná registrace studie, kde je ještě před zahájením studie uvedena jak hypotéza, tak metodika).

Autor + spoluautoři: Autorství a spoluautorství je založeno na:

- a) podstatném příspěvku ke koncepci a designu studie anebo na získání dat anebo analýze a interpretaci dat;
- b) sepsání rukopisu nebo jeho kritické revizi z hlediska významného intelektuálního obsahu
- c) na konečném schválení verze určené k publikování. Autoři by měli splnit všechny tyto podmínky. **Hlavní autor** (ten, kdo píše rukopis) se píše jako první, vedoucí pracovní skupiny jako poslední, ostatní mezi ně podle významu, jakým přispěli k práci. **Autory nejsou** ti, kteří zajistili finanční příspěvek, sbírali data, prováděli obecný dohled nad prací, poskytli technickou pomoc (patří do „poděkování“).

Citace × kompilace: ve vědecké práci jsou přípustné citace (= podklad pro tvrzení, ze kterých studie vychází a která dále rozvíjí). Není přípustná kompilace, tj. použití celých pasáží textu, úvah apod. To přijatelné např. pro učebnice, učební texty (nutno ale respektovat autorská práva).

Měly by být uvedeny všechny externí zdroje financování, pokud byly použity (granty, sponzorské dary)

Konflikt zájmů by měl být deklarován vždy, pokud existují finanční nebo jiné faktory, které by mohly ovlivňovat objektivitu prezentovaných zjištění. Mít konflikt zájmů není samo o sobě nepřijatelné (je to naopak celkem běžné), je ale chybou o tom neinformovat.

Hrubé porušení etiky – vymyšlení a zkreslování dat; plagiátorství vč. netvůrčího kompilování; připisování spoluautorů;

10.3 Jazyk a typografie publikace

Kvalitní publikace musí být nejenom metodicky správně navržená a data musí být správně zpracovaná, ale měla by být prezentovaná i správným a srozumitelným jazykem, při dodržení typografických konvencí.

Moderní programy pro editaci textu (např. MS WORD) obsahují funkce, zvýrazňující pravopisné (červená vlnovka), stylistické (zelená vlnovka) i typografické (modrá čára) chyby.⁶

Přestože tyto funkce nepracují vždy zcela optimálně, prakticky vždy je výsledek lepší, pokud budeme jejich doporučení respektovat.

Nejobvyklejší prohřešky proti jazyku a typografii

- Dlouhá a složitá souvětí, strojená snaha o „úřední“ jazyk, nadměrné používání trpného rodu: srovnej: „Hmotnost jsme zjišťovali pomocí váhy EXTRA220.“ a „Za účelem zjištění hmotnosti, jak bylo plánováno, bylo provedeno vážení za pomoci váhy EXTRA220“.
- Nelogická a nesrozumitelná větná stavba – „Data jsme nesbírali o víkendu. Pouze v případě, že by to jindy nešlo.“
- Věty bez podmětu a přísudku, chybějící čárky, chyby ve shodě větných členů – „U studentů jsme zjišťovali tělesnou hmotnost. A také výšku, které jsme zaznamenali do počítače.“
- Pravopisné chyby (i/y); typicky chybné psaní slova „výjimka“ – správně se píše s dlouhým „ý“ a krátkým „i“, nikoliv opačně („vyjímka“).
- Chybné použití interpunkce – jeden, nebo druhý = vždy jen jeden z obou × jeden nebo druhý = jeden nebo i oba současně.
- Chybné použití uvozovek – správně bez mezer mezi uvozovkami a slovem, před textem jsou uvozovky dole, za textem nahoře.
- Chybný zápis interpunkčních znamének, zejména čárek, teček, dvojteček – správně vždy s mezerou ZA znaménkem; výjimkou je psaní času (15:45) a číslování kapitol (2.1.7).
- Chybné psaní mezery u znaku „%“ (15% = patnáctiprocentní, 15 % = patnáct procent)
- Chybné psaní data – bez mezer za tečkami – správně 15. 5. 2017.
- Nerespektování rozdílu mezi pomlčkou „-“ (delší, tenčí) a rozdělovníkem či spojovníkem (vypadá jako znaménko minus „-“). Rozdělovník je např. zde („chceme-

⁶ Většina „nesprávně označených chyb“ je ovšem ve skutečnosti zaviněna nedodržením typografických konvencí uživatele a počítač tudíž nepochopí, „co tím chtěl básník říct“.

li...“, „česko-slovenský“), pomlčka zde („data z let 2015–2017“). Pokud odděluje pomlčka víceslovné části věty, je ohraničena mezerou („Praha 6 – Ruzyně“). Číselný rozsah se spojuje pomlčkou bez mezer (str. 23–30).

Podrobný popis typografických zásad při sazbě práce najdete např. zde:

<http://orin.meinlschmidt.org/~znouza/skitch/pravidla2.pdf>

11 Nejčastější chyby při přípravě výzkumu, ve zpracování a hodnocení dat

Souhrn ve třech větách:

- *Nejvíce chyb vzniká v přípravné fázi – nerozmyšlením všech souvislostí (detailů) a nerespektováním reálných možností při sběru dat.*
- *Výsledkem rychlé a zbrklé snahy realizovat výzkum je **nenáhodný výběr** experimentální skupiny, **nesystematický sběr dat** a hypotézy definované „ex post“ podle toho, co vyšlo.*
- *Častou chybou je i **nesprávná interpretace dat** – zejména prezentace zjištěných rozdílů bez testování vlivu náhody, nerespektování rozdílu mezi korelací (společným výskytem) a kauzalitou (příčinnou souvislostí), vztahení výsledku platného pro úzkou vybranou skupinu na celou populaci apod.*

11.1 Nejčastější chyby v přípravě výzkumu

- **Cíl výzkumu je definovaný pouze obecně** (srovnej: „Cílem práce je zjistit situaci v LZS v ČR“ vs. „Cílem práce je zjistit rozdíly v počtu primárních zásahů LZS v denní a noční době“ apod.
- **Není stanovena smysluplná hypotéza** (Srovnej: „Zkoumaný lék má vliv na snížení hladiny krevního tlaku“ vs. „Ke snížení krevního tlaku dojde alespoň u 42,5 % pacientů“).
- **Není ověřena spolehlivost dat** u retrospektivních výzkumů (využití historických databází bez toho, aby bylo ověřeno, že data v nich zadaná opravdu odpovídají skutečnosti).
- **„Dábel se skrývá v detailech“** aneb detailní rozdíly v metodice vedou k zásadním rozdílům ve výběru (např. rozdíl mezi „výsledky resuscitací“ a „výsledky resuscitací u pacientů se spatřenou zástavou, pravděpodobně kardiálního původu“ a mnoho dalších).
- **Výběr vzorků nebo účastníků studie není náhodný**, ač by náhodný měl být (např. dotazníkové šetření na obecně společenské téma spořádané především mezi spolužáky, tj. lidmi s obdobným profesním zaměřením, věkem, sociálním zázemím; pacienty do studie „dodávají“ pouze někteří lékaři – pravděpodobně ti, kteří mají

o problematiku zájem, věnují se jí – a tudíž možná mají lepší výsledky, než „průměrný“ lékař... apod.).

- **Nerealistické očekávání** – složitý dotazník, náročný protokol studie + požadavek na plošné využívání = fiasko (buď se většina lidí nezúčastní, nebo dodá zkreslená data – jen aby „něco“ odevzdali).

11.2 Typické chyby v dotaznících

- **Nedbalý přístup k tvorbě dotazníku jako celku**
 - o otázky kladené bez logiky způsobem „co mě zrovna napadlo“. Každá otázka v dotazníku musí mít smysl – musím vědět, PROČ se na věc ptám a také to, proč se na ni ptám PRÁVĚ TAKTO a k jakému praktickému závěru bude odpověď dobrá;
 - o Nesrozumitelně formulované dotazy (Jakou školu jste absolvoval? – Vysokou...pražskou...zdravotnickou...?), negace negace („Nesouhlasím s tím, aby naše záchranka nepoužívala...“ – s čím tedy vlastně nesouhlasím?);
 - o Nedbalé formulace („však on už to respondent pochopí...“) – např. „Pracujete jako ARO, JIP nebo ZZS?“.
- **Nevhodný výběr oslovených respondentů** – ptáme se osob, které s věcí nemají osobní zkušenost, odpovídají podle informací „z doslechu“; dotazníky mi vyplní kamarádi – tedy osoby, které pravděpodobně sdílí můj pohled na svět apod.
- **Ignorování těch, kteří neodpověděli** – ochota odpovídat je přímo závislé na zájmu o změnu (nespokojení); odpovídají spíše aktivní osoby; počítačový dotazník nemusí jít otevřít počítačově méně gramotným respondentům apod. Je zásadní chyba nerespektovat neznámé odpovědi (těch, co neodpověděli) pokud nemáme ověřeno, že odpovědi máme od reprezentativního vzorku populace. Pokud např. ze 100 dotázaných na to, jak jsou spokojeni v zaměstnání, odpoví pouze 10 respondentů a z nich např. 8 není spokojeno, rozhodně výsledek nelze interpretovat tak, že 80% (8 z 10) respondentů je nespokojených. Ve skutečnosti je nespokojených (nejméně) 8% (o těch to víme). 2% nespokojená nejsou a o 90% nevíme vůbec nic.
- **Chybná volba typu otázky** – otevřené otázky mohou jít do hloubky, ale obsah je velmi ovlivněn osobou tazatele (schopností vyjádřit se) a jen velmi obtížně se vyhodnocují; strukturované otázky nemusí přesně vystihnout názor respondenta;

- **Není úplný výběr možných situací** – „Jak se dopravujete do práce – autem, MHD nebo pěšky?“ – chybí např. „kolo“, „brusle“, „letadlo“ – sice okrajové kategorie, ale mohou být zajímavé. Vždy by měla být volba „ostatní“ („jinak“, „jiné“, „nevím“, „nechci odpovídat“...);
- **Zaměření se na jednu situaci**, která autorovi přijde samozřejmá, ale neplatí univerzálně – „jste v tomto zaměstnání spokojenější, než v předchozím?“ – jak má odpovědět respondent, u kterého jde o první zaměstnání
- **Nejasná kvantifikace** – Jak často... – „zřídka“, „výjimečně“, „obvykle“ – každý si pod těmito pojmy v různé souvislosti představí něco jiného
- **Otázky, ve kterých chybí zásadní omezení** – „Chcete větší plat...?“ ...ale chybí – za cenu práce v sobotu i v neděli?... Máte na dovolené rádi zimu? (chybí omezení na jaké dovolené – na lyžích ano, u moře ne...)
- **Citlivé otázky** – „Kolikrát jste se dopustil zásadní odborné chyby?“ – nedávají validní výsledky (nikdo nechce přiznat chyby). Někdy může pomoci rozmezí („Jaký je váš plat, 20 000–30 000, 30 001–40 000 atd.“). Když už jsou citlivé otázky předmětem zkoumání, měla by být uvedena možnost „nechci odpovídat“
- **„Dvojhavňové“ otázky** – „Dáváte přednost kvalitnímu a drahému oblečení, nebo levnému a nekvalitnímu?“ – co když dávám přednost oblečení kvalitnímu, ale levnému? Ale zase naopak – existuje takové oblečení?
- **Škálové otázky bez definice škály** – „Ohodnoťte kvalitu výuky číslem od 1 do 5...“ – jde o „známku“, nebo „počet bodů“?
- **Chybná interpretace názoru jako obrazu reality** – otázky na „měkké“ nebo nepřesně definované skutečnosti se velmi špatně vyhodnocují a velmi špatně se validuje, že měří skutečně to, co od nich očekáváme. Viz otázky typu „Co si myslíte o...“. Správný závěr je „většina udává, že...“, zatímco chybný závěr je „je to tak, že...“. Např. dotazníkové šetření mezi záchranáři, nakolik jsou jejich výjezdy indikovány. Správná interpretace takového výzkumu je např. „záchranáři vnímají v průměru 20% výjezdů jako neindikovaných“, chybná interpretace je „20% výjezdů je neindikovaných“.
- **Široké používání otevřených otázek**, tj. s odpovědí „vlastními slovy“ respondenta. Tyto otázky jsou statisticky prakticky nevyhodnotitelné, slouží spíše pro dokreslení nějakého problému a jejich výsledek je zásadně modifikovaný vyjadřovací schopností respondenta.

11.3 Nejčastější chyby ve zpracování dat

- Ignorování statistického testování – výsledek je založený na pouhém srovnání hodnot – nerespektování vlivu náhody (v souboru A se jev vyskytoval v 26 % měření, zatímco v souboru B ve 32 % měření – závěr v souboru B je tedy více osob se sledovaným jevem je pravděpodobně chybný – záleží na rozsahu souboru);
- Použití nevhodného testu;
- Použití průměru u souboru, který nemá normální rozložení;
- Použití průměru bez uvedení směrodatné odchylky.

11.4 Nejčastější chyby v prezentaci dat

- **Nerespektování hypotézy a metodiky**, vztažení výsledku v určité speciální situaci na obecnou populaci (obzvláště v názvu práce)
 - **Klamavý název práce:** Srovnej: *„Kvalitní první pomoc může zachránit až 30 procent zraněných“* a *„Zaškrcení končetiny u úrazů nášlapnou minou vedlo v Africe v oblastech, kde je lékařská pomoc dostupná do 48 hodin, k poklesu mortality o 30 %“*, První titulky je po vědecké stránce naprosto nesprávný a manipulativní, vč. použití „kouzelného“ slůvka „až“, kterým autor obhajuje nevyřčenou skutečnost, že výsledku bylo dosaženo jen u vysoce specifické, v našich podmínkách prakticky neexistující populace.
 - **Extendování výsledku mimo metodiku a hypotézu studie:** Závěr studie: studie neprokázala, že lék A je účinnější, než lék B. **Chybná extenze:** Lék B je účinnější, než lék A (Může to tak být, ale nemusí – o tom studie nebyla).
Obdobně: Závěr studie: studie zabývající se podáním léku X v nemocnici prokázala, že je účelné jeho co nejčasnější podání. **Chybná extenze:** Lék X by měl být podáván už v přednemocniční péči PNP (Může to tak být, ale nemusí – může se např. stát, že podání vyžaduje nějaké přípravy a v PNP bude natolik zdržující, že se tím přínos léku X eliminuje).
 - **„Čím více, tím lépe“ resp. nezahrnutí nežádoucích účinků** – pokud je výsledkem studie je zjištění, že provedení „A“ způsobuje jev „B“, ještě to neznamená, že provedení „2xA“ způsobí „2xB“. Typicky podání jedné dávky

léku může být účinné, ale rozhodně automaticky neplatí, že „čím větší dávka, tím větší účinek“ – účinek dokonce může klesat, resp. vyšší dávky budou škodlivé. Obdobně např. daně nelze zvyšovat donekonečna, protože postupně převáží jejich „nežádoucí účinky“ (demotivace) a celkový efekt začne klesat.

- **Zahrnutí neověřeného předpokladu do hypotézy** (viz např. použití měření dojezdového času k hodnocení kvality záchranné služby – automaticky předpokládáme, že nižší čas lepší kvalita. To ale vůbec nemusí být pravda – nikdy nebylo nic takového prokázáno, ba naopak – kratší průměr může být spojený s prodloužením dojezdů u kritických stavů. Pokud měříme dojezdový čas, musí se hypotéza týkat dojezdového času, nikoliv „automaticky“ kvality záchranné služby); obdobně tvrzení „Zákaz kouření vedl ke snížení množství infarktů, protože na naší klinice jsme letos ošetřili méně pacientů s infarktem“ vyhází z nevyřčeného předpokladu, že dotyčná klinika ošetřuje stále stejný podíl pacientů s AIM (což ovšem vůbec nemusí být pravda – počet pacientů s infarktem ve skutečnosti nikdo nezměřil).
- **Nerespektování rozdílu mezi korelací a kauzalitou** („rozvedení trpí častěji depresí“ – víme, ale, zda trpí rozvedení depresí kvůli rozvodu, nebo nastal rozvod kvůli depresi? Obdobně „intubovaní pacienti mají horší prognózu“ – ale mají horší prognózu, protože jsou intubovaní, nebo jsou intubovaní, protože jsou v horším stavu, a tudíž mají horší prognózu?); „Cyklisté, kteří nosí přilbu, mají méně úrazů“ – je to tím, že je ochránila přilba, anebo tím, že přilbu nosí zodpovědnější cyklisté, kteří také méně riskují?
- **Chyby v prezentaci procent:**
 - o **Prezentace podílu bez znalosti základních souborů:** „Plných 10 % nehod zavinil řidič pod vlivem alkoholu“ – pokud ovšem neznáme podíl řidičů, jezdících pod vlivem alkoholu, nelze rozhodnout, zda alkohol k nehodám vede, nebo před nimi naopak chrání (*pokud hypoteticky pod vlivem alkoholu jezdí např. 50 % řidičů, ale zaviní jen 10 % nehod, tak alkohol naopak před nehodami chrání...*).

- **Procenta reprezentující příliš malý vzorek** (10 % z 10 měření vs. 10 % z 1 000 měření; faktický stav „1 ze 4“ je prezentovaný jako „plných 25 %“ apod.);
- Nerespektování rozdílu mezi **procentem** a **procentním bodem** (pokud úspěšnost léčby poklesla z 20 % na 15 %, nejde o pokles o 5 %, ale buď o pět procentních bodů, nebo o pokles o 25 % (pět z dvaceti));
- **Použití nevhodného typu grafu** (např. koláčový graf, jehož dílky nevyjadřují část z celku apod.)
- **Sčítání vzájemně nezávislých procent** (např. „Na záchrance tvoří 20% výjezdů úrazy, 10% dopravní nehody a 5% popáleniny. To je celkem 35%“. Samozřejmě přitom platí, že většina pacientů při dopravní nehodě či popálenině současně spadá do kategorie „úraz“ a ve statistice jsou tedy zahrnuti dvakrát.
- **Chybná interpretace odečtení a přičtení procent** – výsledkem poklesu nějakého čísla a jeho následného zvýšení o stejnou relativní část (procenta) není původní číslo, protože zvýšení vychází z nového (menšího) základu. Viz např. snížení platu o 20% a jeho následné zvýšení o 20% - při původním platu 100 Kč je plat po snížení 80 Kč (-20%), ale po následném zvýšení o 20% pouze 96 Kč (20% z 80 je 16). Analogicky i opačně – zvýšení a následné snížení o stejná procenta znamená ve výsledku vyšší hodnotu.

- **Úmyslné manipulace s daty**

- **Úmyslné zkreslení** (vymyšlení si) dat – viz např. tvrzení typu „80 % výjezdů záchranné služby není indikovaných“, aniž by proběhla jakákoliv studie na toto téma
- **Vynechání části populace**, kde výsledky „nejsou podle mých představ“, bez odpovídající úpravy metodiky a hypotézy
- **Prezentace dat zdůrazňující nebo naopak potlačující rozdíly** (graf s potlačenou nulou; vyjádření velkých hodnot celým zápisem x zkráceným apod. – 1.000.000 vs. 1 mil.);

12 Jak hodnotit studie aneb pár slov na téma medicíny založené na důkazech (evidence based medicine – EBM)

Souhrn ve třech větách:

- *Medicína založená na důkazech není slepá aplikace výsledku studií, ale propojení kriticky hodnocených vědeckých důkazů, zkušenosti zdravotníka a potřeb konkrétního pacienta.*
- *Není studie jako studie – za nejvěrohodnější zdroje poznání se jsou považované randomizované experimentální studie, následované ostatními experimentálními studii, analytickými studii, deskriptivními studii až po názory expertů a výsledky jednotlivých experimentů.*
- *Typ studie sám o sobě kvalitu nezaručuje – dobrá studie musí mít jasně definované parametry podle metodiky PICO.*

12.1 Medicína založená na důkazech

Ačkoliv se celé toto pojednání zabývá prací s daty, na závěr je nutné zdůraznit, že výsledky jakkoliv kvalitních studií nemají univerzální a trvalou platnost. Opak bývá zpravidla pravdou – ať už proto, že se mění vnější podmínky (kvalita dostupné péče, zdravotní stav obyvatel, vlivy životního prostředí, ekonomiky atd...), ale i proto, že výsledek studie byl ovlivněný nějakou další neznámou nebo nepopsanou chybou. Není žádnou výjimkou, že jeden a týž problém je v různém prostředí zkoumán několika studii, které dojdou k protichůdným závěrům. Žádná ze studií přitom není metodicky chybná – jen byly provedeny v různém prostředí, jehož vliv často nedokážeme přesně popsat a změřit. Je zkrátka nutné smířit se s tím, že výsledky jsou málokdy černobílé a trvale platné. Z tohoto důvodu vznikl pojem „**medicína založená na důkazech**“, což je celosvětově respektovaná metodika zabývající se tím, jak zacházet s výsledky studií.

Medicína založená na důkazech (Evidence-Based Medicine - EBM) „je systematický přístup k léčbě, **propojující nejlepší dostupné vědecké důkazy s klinickou zkušeností lékaře a s preferencemi a potřebami ošetřovaného pacienta**“.

V definici EBM je naprosto klíčové právě to, že „vědecké důkazy“ (výsledky studií) je **NUTNÉ** propojit se **zkušenostmi zdravotníka** a situací **konkrétního pacienta**, tj. ne je slepě aplikovat, ale vždy vážit všechny okolnosti. EBM bývá občas zaměňována za samotnou aplikaci nejlepších vědeckých poznatků v klinické praxi. Tento postup pak vystavuje celý koncept EBM kritice, že výhradní používání výsledků studií popisujících obecné tendence může při péči o konkrétního pacienta selhat.

12.2 Hodnocení kvality studie

Pro hodnocení kvality výsledku je znát především typ studie. Výsledky studií různých typů mají různou váhu v závislosti na tom, jak moc byla daná metodika „odolná“ vůči potenciálním chybám („Hayesova pyramida“). Pořadí metodické kvality je následující:

1. Randomizované experimentální studie
2. Prospektivní experimentální studie
3. Retrospektivní analytické studie
4. Deskriptivní studie
5. Názory expertů, výsledky jednotlivých experimentů

Typ studie ovšem sám o sobě pro kvalitní publikaci nestačí. Při posuzování zda nalezená publikace je kvalitní, je třeba zvažovat pět otázek, které by měly tvořit základ pro rozhodnutí, zda publikaci odložit nebo ji použít k ovlivnění praxe:

- Byla studie původní?
- Koho se studie týkala?
- Byl použit správný design?
- Zabránilo se systematické chybě (tj. byla studie dostatečně kontrolovaná)?
- Byla dost velká a trvala dost dlouho na to, aby přinesla věrohodné výsledky?

Po kritickém posouzení metodologické kvality publikace by každý měl být schopen říci několika větami, jaký druh studie byl proveden, u kolika osob, odkud osoby pocházely, jaká léčba či jiná intervence jim byla nabídnuta, jak dlouhé bylo období sledování (nebo, pokud šlo o přehledovou studii, jaká byla četnost odpovědí) a co bylo použito jako měřítko následku (PICO kritéria – viz kap. 3). V této fázi by se také mělo určit, jaké statistické testy, pokud vůbec nějaké, byly použity k analýze výsledků. Pokud jsou tyto

věci jasné, je snadnější porozumět výsledkům studie, hodnotit jejich význam a, je-li to na místě, i je případně odmítnout.

13 Zdroje

Masarykova univerzita: [Typy epidemiologických studií \(LF UK\)](#)

Sekera M. Vybrané kapitoly z metodologie. Masarykova univerzita, Brno 2012 ([online](#))

Wikiskripta: [Deskriptivní studie](#) [Analytické studie](#) [Experimentální studie](#) [Metodologie v epidemiologii](#)

Janout V. Medicína založená na důkazu ([online](#))

Holčík J., Komend M. e-Learningová učebnice matematické biologie ([online](#))

Šrámek J. Možnosti a limity psychosomatické medicíny ([online](#))

Dušek L. a spol. Analýza dat v neurologii ([online](#))

Grim T. (Spolu)autorství – triky a pověry ([online](#))

Knecht P. a spol. Etika vědecké práce a publikování pro mírně pokročilé ([online](#))
www.biostatisticka.cz

Sackett DL, Rosenberg WMC, Gray AJM a Hayes RB: Evidence based medicine: what it is and what it isn't. BMJ 1996;312(7023):71-72.)

DRDKOVÁ, Soňa a Marcel JOSÍFKO. Zdravotnická statistika. Státní pedagogické nakladatelství Praha, 1973.

Huff D. Jak lhát se statistikou. Brána Praha, 2013.

*** KONEC TEXTU ***