

UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V KOŠICIACH



Jaroslava KOPČÁKOVÁ

ÚVOD DO METODOLÓGIE VÝSKUMU

VYSOKOŠKOLSKÉ UČEBNÉ TEXTY

Košice 2024

UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V KOŠICIACH
Lekárska fakulta
CENTRUM MEDICÍNSKEHO VZDELÁVANIA



Jaroslava KOPČÁKOVÁ

ÚVOD DO METODOLÓGIE VÝSKUMU

Košice 2024

Úvod do metodológie výskumu

Vysokoškolský učebný text

Autorka:

Mgr. Jaroslava Kopčáková, PhD.

Centrum medicínskeho vzdelávania, Lekárska fakulta UPJŠ v Košiciach

Recenzentky:

doc. PaedDr. Petra Lajčiaková, PhD.

Filozofická fakulta Katolíckej univerzity v Ružomberku

PhDr. Tatiana Pethö, PhD.

Fakulta manažmentu, ekonomiky a obchodu Prešovskej univerzity v Prešove

Jazykové korektúry:

Bc. Emília Gajdošová

Táto práca bola podporovaná Agentúrou na podporu vedy a výskumu na základe zmluvy č. APVV-22-0078.

Tento text je publikovaný pod licenciou CC BY NC ND Creative Commons Attribution-NonCommercial-No-derivates 4.0 („Uvedte pôvod – Nepoužívajte komerčne - Nespracováajte“)



Za odbornú a jazykovú stránku vysokoškolského učebného textu zodpovedá autorka.

Dostupné od: 15.12.2024

Umiestnenie: www.unibook.upjs.sk

ISBN 978-80-574-0371-5 (e-publikácia)

Obsah

Úvod	4
1. VÝSKUMNÝ PROCES.....	5
1.1. Odlíšenie vedeckých metód poznávania od iných spôsobov poznávania	5
1.2. Rozpoznanie jednotlivých krokov výskumného procesu	7
1.3. Porozumenie procesu transferu konceptu do štatistickej premennej a rozlišovanie typov premenných	8
1.4. Porozumenie procesu transferu populácie do výskumnej vzorky a rozlišovanie spôsobu výberu vzorky	15
1.5. Formulácia výskumnej otázky	18
OTÁZKY	20
2. MERANIE VPLYVU	22
2.1 Popis spojitých premenných strednými hodnotami	22
2.2 Spôsoby úpravy údajov na štatistické premenné.....	23
2.3 Exponovanie skúmanej okolnosti	29
2.4 Popísanie frekvencie prípadov	30
2.5 Popísanie vplyvu expozície na výskyt prípadov, interpretácia	31
OTÁZKY	35
3 DIZAJN ŠTÚDIE.....	37
3.1 Klasifikácia dizajnov štúdií	37
3.1.1 Popisné (deskriptívne) dizajny štúdií.....	38
3.1.2 Analytické dizajny štúdií	41
3.2 Iniciatívy a odporúčania	49
OTÁZKY	51
4 VALIDITA ŠTÚDIE	57
4.1 Vnútoraná a vonkajšia validita štúdie	57
4.2 Rozlíšenie náhodnej chyby, systematickej chyby a vplyvu mäťúcich premenných.....	58
OTÁZKY	66
Zoznam použitej literatúry:	69

Úvod

Vážené študentky, vážení študenti,

dostáva sa vám do rúk vysokoškolský učebný text Úvod do metodológie výskumu, ktorý slúži ako podkladový a výučbový materiál. Jeho cieľom je pomôcť vám rozvinúť kompetencie dôležité pre vedeckú prácu.

Nasledujúci text je rozdelený do štyroch častí. Prvá časť je zameraná na vedecké metódy poznávania a na jednotlivé kroky výskumného procesu. V druhej časti prinášame informácie o meraní vplyvu, popisujeme spojité premenné strednými hodnotami, spôsoby ako upraviť údaje na štatistické premenné, venuje sa popísaniu frekvencie prípadov a vplyvu expozície na výskyt prípadov. Tretia časť venuje pozornosť rozličným typom dizajnov výskumných štúdií, klasifikácii dizajnov štúdie, výhodám a nevýhodám jednotlivých dizajnov štúdií a užitočným iniciatívam a odporúčaniam. Záverečná štvrtá časť prináša informácie o vnútornej a vonkajšej validite štúdie, identifikovaniu alternatívnych vysvetlení pozorovaných súvislostí ako je náhodná chyba, systematická chyba a vplyv mätúcich premenných a rozpoznaní najvýznamnejších zdrojov systematickej chyby.

Prednášky k jednotlivým častiam sú verejne dostupné prostredníctvom odkazu: <https://youtube.com/playlist?list=PLULtZO95jmNYFqLDuXoySmBZNAmXlKLNu&si=RL-Z7P3THk9le85t>.

Každá z kapitol prináša kľúčové poznatky sa danú oblasť, ktoré sú v priebehu textu ilustrované na praktických úlohách, ktoré študent priebežne rieši. Osvojenie si poznatkov za danú kapitolu si študent môže v závere overiť na pripravených otázkach. Želám vám všetkým, aby bol tento vysokoškolský učebný text príjemne a užitočne stráveným časom a naplnil svoj účel.

1. VÝSKUMNÝ PROCES

V tejto kapitole si predstavíme vedecké metódy poznávania a výskumný proces.

Vedecké poznávanie je len jedno z mnohých spôsobov poznávania a je len užučkým chodníčkom medzi obrovskými diaľnicami a viacprúdovými cestami, ktorými ten, ktorý konkrétny človek putuje za poznáním. Nie je teda ani najefektívnejším, ani najčastejším, dokonca ani nie najdôležitejším spôsobom poznávania. Ak by sme boli odkázaní len na vedecké poznávanie, nikdy by sme sa nedostali ani za učebné osnovy prvého ročníka, lebo by sme museli poctivo objavovať a preverovať každý jeden poznatok. V jednom ľudskom živote by sme mohli obsiahnuť len veľmi obmedzenú, temer nevýznamnú časť toho, čo potrebujeme ku každodennému životu, takže by sme boli dosť nepoužiteľní. Nikdy by sme nezmúdreli, lebo príbehy ľudí a ich skúsenosti sa nevtesnajú do pravidiel vedeckého poznávania. Poznávame vtedy, keď hľadáme ako vyhlíbiť jamu v piesku tak, aby v nej ostala morská voda pri prílive, ale aj keď učiteľka rozpráva o nervovej sústave dáždovky, aj keď sa túlame cudzími mestami a rozprávame s inými ľuďmi o ich skúsenostiach. Poznávame aj vtedy, keď sa pokúšame namaľovať obraz, či čítame román. Poznávame všade tam, kde nás zvedavosť núti hľadať odpovede na otázky, kde rozširujeme porozumenie dianiú okolo seba.

Vedecké poznávanie je iné v tom, že sa snaží dodržiavať isté pravidlá, aby predišlo chybám a snaží sa byť transparentné, aby sa mohlo neustále spochyňovať a opravovať.

Cieľom tejto kapitoly je pomôcť Vám *odlišovať vedecké metódy poznávania* od iných spôsobov poznávania. Ďalším cieľom je predstaviť Vám jednotlivé kroky výskumného procesu, vedieť rozpoznať jednotlivé kroky *výskumného procesu*, pomôcť Vám porozumieť akým spôsobom sa od zvedavej otázky dostaneme k *sformulovaniu výskumnej otázky*. V tejto kapitole nájdete aj ukážky, ako sa z nejakého konceptu dajú vytvoriť *štatistické premenné*, ktoré potrebujeme na to, aby sme mohli použiť štatistické metódy a vyjadrovať sa k súvislostiam medzi javmi či okolnosťami. A taktiež aj to ako môžeme nahradiť populáciu výskumnou vzorkou a predstavíme si rôzne spôsoby *výberu vzorky* vo výskumnom procese.

1.1. Odlíšenie vedeckých metód poznávania od iných spôsobov poznávania

Carl Edward Sagan, významný americký vedec, astronóm, astrofyzik, astrobiológ a popularizátor vedy a kritického myslenia, ktorý okrem iného pokazil radosť všetkým ufológom napriek tomu, že ho na počiatku kariéry existencia ufa nadchla, povedal nasledovné: „**Veda *nie je dokonalá*. Je to len nástroj. Ale je to zďaleka ten najlepší nástroj, aký mám k dispozícii, ktorý je schopný **sebakorekcie** a je aplikovateľný na všetko. Má dve pravidlá: Po prvé - **žiadne poznanie nie je nedotknuteľné**, sväté; všetky tvrdenia, predpoklady musia prejsť kritickou revíziou, overovaním; argumenty založené len na autorite sú bezcenné. Po druhé – **čokoľvek je v rozpore so skutočnosťou, musí byť vylúčené alebo zrevidované**.**“

Sagan vo svojom výroku popisuje jeden z poznávacích znakov vedeckej metódy a tou je sebakorekcia. Esenciálnou časťou každého vedeckého tvrdenia je ohraničenie za akých okolností a podmienok a do akej miery je možné spoľahnúť sa na dané tvrdenie a kde sú jeho limity. Esenciálnou časťou každého vedeckého poznávania je otvorenosť, transparentnosť, lebo bez nej nie je možné zachovať ďalší atribút a to je verifikovateľnosť. V gastronómii si šéfkuchár môže dovoliť zatajiť recept, ale vo vede je to jasná diskvalifikácia. Ako ste k danému poznaniu prišli musíte popísať tak presne, aby to mohli tí ďalší za vami zopakovať a dopracovať sa k rovnakým výsledkom. Inak je tu rozpor a poznanie musí podstúpiť revíziu.

Ak k tomu istému výsledku majú dôjsť viacerí, nemôže byť založené na individuálnych dojmoch, ale musí byť založené na objektívnych ukazovateľoch, na ktorých sa vieme zhodnúť.

Rozmýšľanie, uvažovanie, formulovanie hypotéz a teórii je nosnou zložkou vedy, ale vedecké poznávanie vyžaduje empirické overovanie. Môžeme mať veľmi silné, teoreticky podložené a expertmi formulované dôvody veriť, že niečo s niečím súvisí, ale bez empirického dôkazu to nie je dostačujúce. Na druhej strane chaotické, ojedinelé, anekdotické empirické skúsenosti nemajú s vedou nič spoločné a použiteľné sú len na skrátenie chvíle v čakárňach u lekárov. Poznávanie musí byť systematické a dobre organizované, aby zaručilo, že sa skontrolovali všetky možnosti.

Vedecká metóda poznávania sa teda od iných foriem poznávania podľa Charlesa Sandersa Peirca (Ferjenčík, 2010) líši minimálne v týchto 4 atribútoch:

1. systematické a organizované poznávanie (vedecké bádanie tvorí práca organizovaná podľa určitých pravidiel a princípov)
2. empirická povaha (vedecká metóda stavia na informáciách získaných skúsenosťami, kontaktom a manipuláciou s predmetom skúmania - tieto informácie nazývame dáta)
3. objektívnosť a verifikovateľnosť (objektívnosť znamená, že výsledok je nezávislý od toho, čo si praje výskumník, ale je daný dátami, teda aj iný výskumník by došiel opakovaním k rovnakému výsledku)
4. sebakorekcia (nepretržité približovanie k pravde, nie samotná pravda).

Cieľom vedeckej metódy poznávania je popísať, predvídať alebo vysvetliť javy okolo nás.

PRAKTICKÁ ÚLOHA 1:

Ktorá z charakteristík vedeckého poznávania chýba v nasledujúcich príkladoch?

- A) systematické a organizované poznávanie
- B) empirická povaha
- C) objektívnosť a verifikovateľnosť
- D) sebakorekcia

- Pitie zázvorového čaju lieči nádchu, lebo mne to vždy zaberie.
- Homeopatia je účinná, aj keď pri zriedení neostáva v roztoku žiadna účinná látka.
- Dinosaury boli zelené.
- Veľryby sú neviditeľné, lebo nepoznám nikoho, kto by nejakú videl.
- Červená na križovatke reaguje na moje auto, lebo je vždy červená keď príde.

RIEŠENIE PRAKTICKEJ ÚLOHY 1:

A, C Pitie zázvorového čaju lieči nádchu, lebo mne to vždy zaberie.

Skúsenosti s pitím zázvorového čaju majú asi mnohí a mnohí to majú ako rodinnú tradíciu. Možno im to poradil lekár alebo o tom čítali v nejakej knihe. To všetko sú veľmi užitočné spôsoby poznávania a transferu poznania. V tomto konkrétnom prípade však síce je prítomná empirická povaha, ale chýba systematické a organizované skúmanie a poznanie je založené na subjektívnej skúsenosti. To vôbec neznamená, že nie je možné postaviť výskumný plán, ktorý by overoval účinnosť pitia zázvorového čaju pri liečbe nádchy.

B, C Homeopatia je účinná, aj keď pri zriedení neostáva v roztoku žiadna účinná látka.

Ďalší príklad sa týka homeopatie. Homeopatia je účinná, aj keď pri zriedení neostáva v roztoku žiadna účinná látka. Ak tam ale neostáva žiadna účinná látka, nevieme empiricky overiť, či v prípade ak tam je, respektíve nie je k účinku dochádza alebo nedochádza. Chýba teda empirická povaha a chýba možnosť verifikovať toto poznanie.

C Dinosaury boli zelené.

V treťom príklade tvrdíme, respektíve by sme chceli vedieť, či boli dinosaury zelené. Zachovali sa nám z nich nejaké kosti, zo zubov vieme určiť, či boli bylinožravce alebo mäsožravce, ale farbu ich kože už asi nedokážeme overiť. Kože sa asi v takom dobrom stave nezachovali a vrátiť sa späť v čase zatiaľ nevieme.

A Veľryby sú neviditeľné, lebo nepoznám nikoho, kto by nejakú videl.

V niektorých situáciách je veľmi dôležité dôverovať len vlastnej skúsenosti. Veľryby sa však nezmestia do domácich akvárií a drvivá väčšina z nás, žijúcich uprostred kontinentu, veľryby na vlastné oči ani nevidela. Vedecké poznávanie by však malo systematicky preskúmať skúsenosti ľudí súčasnej i minulých generácií a nespoliehať sa len na vlastné zážitky.

A Semafor na križovatke reaguje na moje auto, lebo je vždy červená keď príde.

Nuž a posledné tvrdenie Semafor na križovatke reaguje na moje auto, lebo je vždy červená, keď príde na križovatku. Keďže sa to opakuje a môžu to potvrdiť aj spolujazdci, mohlo by to spĺňať kritérium objektívnosti a verifikovateľnosti a určite to má aj empirickú povahu. Ale chýba tam systematické a organizované skúmanie, čo by v tomto prípade znamenalo urobiť si výskumný plán a v rôznych časoch s rôznymi autami skúšať spustiť červenú na semafore.

1.2. Rozpoznanie jednotlivých krokov výskumného procesu

Už vieme čím sa líšia vedecké metódy poznávania od iných druhov a spôsobov poznávania a v nasledujúcej časti si povieme o jednotlivých krokoch výskumného procesu.

Výskumný proces začína **identifikovaním výskumného problému**, teda definovaním toho čo a prečo to chceme skúmať. Výstupom je teda zdôvodnená výskumná otázka. Druhým krokom je hľadanie spôsobu ako zodpovedať túto otázku čo najlepším spôsobom, teda čo najhodnovernejšie, za daných časových, personálnych a finančných možností a najmä pri zachovaní etických štandardov. Výstupom druhého kroku je vypracovanie **výskumného plánu/protokolu**. Nasleduje **zber údajov a vytvorenie databázy**. Keď už máme vytvorenú databázu, môžeme pristúpiť k **analýze a interpretovaniu údajov**. Len malinký zlomok z analýz, ktoré je obvykle potrebné vykonať, sa nakoniec aj reportuje vo

výsledkoch. Nie je to z toho dôvodu, že by boli dané výsledky chybné, alebo zbytočné, ale preto, že aj pri analyzovaní údajov je potrebné systematicky prehľadávať a overovať všetky možnosti, je potrebné zisteniam rozumieť. Po tejto fáze nasleduje **reportovanie zistení, hodnotenie výskumu a zhodnotenie validity štúdie**.

Identifikovanie výskumného problému nie je len kombinačná hra, kde skúšame, čo s čím by sme mohli spojiť a porovnať. Podnet by mal vychádzať z praxe, skúsenosti z prvej línie, skutočných potrieb reálneho života. Mať živý kontakt s praxou. Počúvať ľudí v prvej línii je preto výborný štart. Poznanie potrieb praxe je preto jedným zo základných pilierov pri identifikácii výskumného problému. Paralelne s vnímavosťou na potreby reálneho života je dôležitá poctivá práca s vedeckou literatúrou. Skôr než sa pustíme do formulácie výskumnej otázky, potrebujeme zistiť čo už vieme, akým spôsobom k tomuto alebo podobnému problému pristupovali iní a čo sa na tej ceste naučili, čo by sme mohli využiť. Poznanie súčasného stavu vedeckého poznania je preto kľúčové, následne je dôležité a potrebné si ujasniť ako konceptualizovať, pomenovať, ohraničiť dôležitých hráčov, naše budúce premenné. Musíme si ujasniť v akých súvislostiach, vzťahoch sú nami vybraní hráči a či nám stále niekto nechýba. Musíme si vytvoriť teóriu, ako to asi funguje – na základe toho, čo vieme z prvej línie, na základe toho, čo sme našli v literatúre, na základe logického uvažovania a ideálne si tie vzťahy zaznamenať/zakresliť do modelu. Potrebujeme si ujasniť o akej skupine, akej populácii chceme vynášať súdy, teda na koho chceme, aby naše zistenia boli uplatniteľné. Ak to zhrniem, je to veľa počúvania, čítania, diskutovania, kreslenia a najmä uvažovania. Samotným výsledkom je obvykle jedna veta, sformulovaná výskumná otázka, ktorá vládne celým výskumným procesom.

1.3. Porozumenie procesu transferu konceptu do štatistickej premennej a rozlišovanie typov premenných

V nasledujúcej časti si postupne prejdeme ako hráčov a koncepty pretransformujeme na čísla a urobíme z nich štatistické premenné. Začneme úlohou, v ktorej si na konkrétnom príklade vyskúšame proces, ktorý sa vlastne opakuje pri každej výskumnej úlohe.

PRAKTICKÁ ÚLOHA 2:

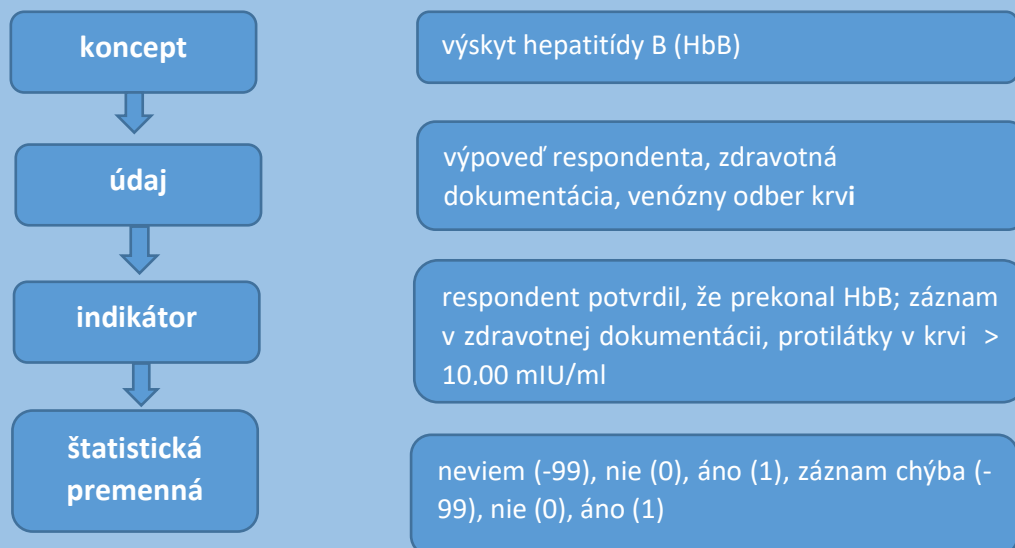
Skúste si predstaviť, že ste dostali za úlohu zistiť, aké okolnosti súvisia s výskytom hepatitídy B. Na hárok papiera si skúste napísať odpovede na nasledujúce otázky. Nemalo by Vám to zabráť viac ako 5 až 10 minút a je to väčšia zábava, ak môžete túto praktickú úlohu robiť v menšej skupine.

- Aké informácie budete potrebovať?
- Sú tieto informácie pozorovateľné, merateľné, zistiteľné a ak áno, ako by ste to urobili?
- Ako by ste zisťovali/ merali výskyt hepatitídy B? Odkiaľ budete mať informácie? Aký formát bude mať tento nameraný údaj?
- Aké okolnosti by mohli súvisieť s vyšším výskytom hepatitídy B? Odkiaľ by sa dali tieto informácie získať, respektíve, ako by sa dali zmerať? Mali by sme ešte niečo zohľadniť?

Kratučké zhrnutie informácii k hepatitíde B pre tých, ktorí toto ochorenie nepoznajú: Vírusová hepatitída typu B sa najčastejšie prenáša priamym kontaktom s nakazenou krvou alebo pohlavným stykom. Inkubačná lehota, teda čas od nákazy do prepuknutia prvých príznakov, je 4 týždne až 6 mesiacov. Po jej uplynutí má mnoho ľudí iba mierne príznaky podobné chrípke. Niekedy ochorenie prebehne úplne bez príznakov. Niektorí chorí sa však úplne neuzdravia, vírus v ich tele preživa ďalej a

choroba trvá v tzv. chronickej forme. Títo ľudia často o svojej chorobe nevedia, pretože nemajú žiadne príznaky. Sú však zdrojom nákazy pre okolie. Sami sú ohrození rozvojom cirhózy a rakoviny pečene.

RIEŠENIE PRAKTICKEJ ÚLOHY 2:



Ak by sme mali skúmať okolnosti súvisiace s výskytom hepatitídy B, potrebovali by sme v prvom rade vedieť, kto má a kto nemá hepatitídu B. Výskyt hepatitídy je teda jeden z konceptov, ktoré skúmame. Ako však určiť, kto je alebo nie je infikovaný? Aké údaje sú na to potrebné?

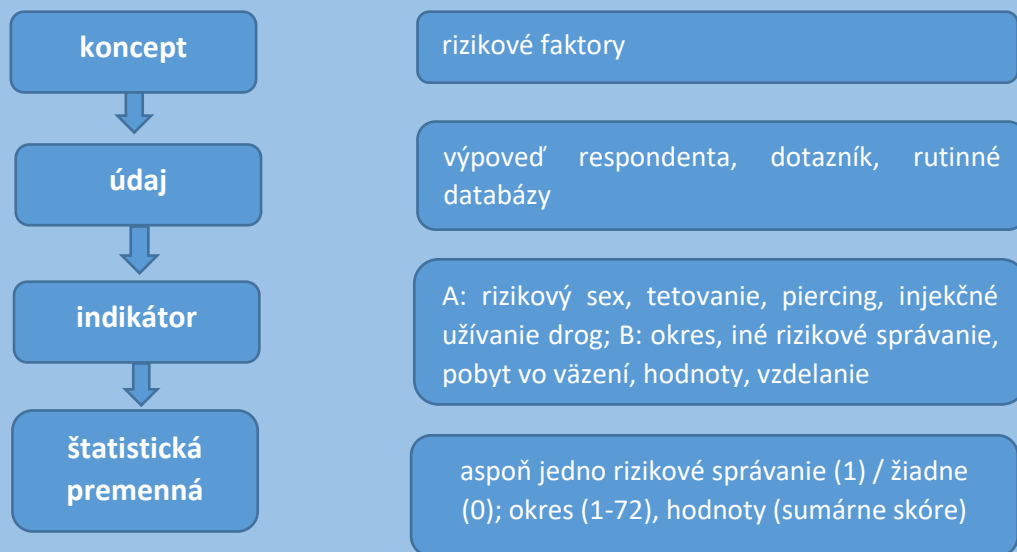
Respondentov sa môžeme opýtať: „Prekonal ste niekedy hepatitídu?“ Ich odpovede by mohli znieť áno, nie alebo neviem. . Môžeme sa na takýto údaj spoľahnúť? Čo ak si to nepamätajú alebo si pletú hepatitídu A a hepatitídu B? Čo ak o svojom ochorení ani nevedia? V takom prípade by sme mohli požiadať ich ošetrojúceho lekára, aby to dohľadal v zdravotnej dokumentácii. Tam by to predsa malo byť uvedené. Ak niekto príde s ťažkosťami, odošlú ho na infekčné oddelenie alebo do hepatologickej ambulancie, kde ho liečia a údaj zaznamenajú. Informácia sa ale nemusela dostať k ošetrovateľskému lekárovi, ochorenie mohol prekonať skôr, než začal navštevovať tohto konkrétneho lekára, karta sa mohla stratiť, pacient nemusel doniesť správy od hepatológa k všeobecnému lekárovi atď.

Chceme si byť absolútne istí, a tak sa rozhodneme, že odoberieme každému respondentovi krv a vyšetríme ju na prítomnosť protilátok v krvi. Ak bol organizmus konfrontovaný s týmto vírusom, bude mať protilátky. S protilátkami je to zložitejšie, ale pre účely tohto cvičenia to teraz skúsime ignorovať. Vyšetrenie protilátok si vyžaduje invazívne vyšetrenie, nie lacné testy, ale proste chceme mať istotu. Kto má protilátky, ten prekonal hepatitídu B. Je to tak ale určite? Protilátky sa v tele vytvoria aj pri očkovaní. Od roku 1987 sa očkujú zdravotnícki pracovníci, študenti zdravotníckych škôl a pacienti z rizikových skupín. Od roku 1998 sa očkujú všetky deti po narodení. Respondenti, ktorí boli očkovaní, teda budú v našom výskume figurovať ako tí, čo prekonal hepatitídu B, a to by nám mohlo značne skresliť zistenia. Máme teda prvý významný moderátor, premennú, ktorú musíme nejakým spôsobom zohľadniť, lebo kvôli nej neuvidíme skutočné súvislosti medzi expozíciou a dôsledkom. V tomto konkrétnom prípade by sme ich napríklad mohli vylúčiť zo súboru, lebo oni de facto nemôžu byť exponovaní vírusu, lebo majú bariéru v podobe očkovaním získanej imunity.

Rizikovou skupinou pre hepatitídu B sú mladí ľudia vo veku 15 až 24 rokov. To môžeme overiť tým, že rozdelíme populáciu do troch vekových skupín: 15 až 24 rokov, 25 až 49 rokov a 50 rokov a viac, a

následne porovnáme výskyt hepatitídy v týchto skupinách. Kde by bol výskyt najvyšší? V skupine, o ktorej sa predpokladá, že je riziková, teda 15 až 24 rokov? Obávam sa, že práve tam by mohol byť najnižší. Jednoduchý dôvod je ten, že 30-roční mali dvojnásobok času na nakazenie sa vírusom hepatitídy B v porovnaní s 15-ročnými. Aj keď vo vyššom veku, napríklad v 65 rokoch, už nie je toľko príležitostí na nákazu ako vo veku 25 rokov, tieto príležitosti sa už v minulosti mohli vyskytnúť a prejavíť svoje následky. Vek teda bude dôležitým moderátorom, ktorý musíme zohľadniť. V tomto konkrétnom prípade by sme ho buď zahrnuli do modelu, alebo zabezpečili, aby sme porovnávali skupiny s podobnou vekovou štruktúrou.

Druhým konceptom sú rizikové faktory:



Druhým konceptom sú teda rizikové faktory. Vírusová hepatitída typu B sa najčastejšie prenáša priamym kontaktom s nakazenou krvou alebo pohlavným stykom.

To nám umožňuje identifikovať možnú expozíciu. Nechránený pohlavný styk a najmä s osobami, ktoré sú rizikové. Tetovanie alebo piercing a najmä v podmienkach, kde sa nedodržia hygienické štandardy. Injekčné užívanie drog. V každom z týchto prípadov platí, že k infikovaniu môže dôjsť len od infikovaného, takže môžete sa nechať tetovať aj vo väzení nesterilnou ihlou a mať šťastie, alebo sa nechať tetovať vo vyhlásenom tetovacom štúdiu a mať smolu.

S veľkou pravdepodobnosťou budete mať šťastie na miestach, kde sa hepatitída B nevyskytuje alebo sa vyskytuje len málo, a naopak smolu na miestach, kde sa vyskytuje často. A na to by sme sa mohli pozrieť do rutinných štatistík či hlásení a následne identifikovať okresy, v ktorých je vyšší výskyt alebo najčastejšie miesta výskytu, napríklad vo väzeniach. V niektorých rómskych osadách je výskyt výrazne vyšší, čo má svoj dôvod. Ak sa infikuje niekto s dobrým prístupom k zdravotníckej starostlivosti, vysokou mierou zdravotnej gramotnosti, tak s veľkou pravdepodobnosťou rozpozná, že sa niečo deje, má sa s kým poradiť, vyhľadá lekára a podstúpi liečbu. Ak sa to isté stane niekomu z rómskej osady, ktorý má množstvo iných problémov, ktoré ho obťažujú viac, nemá sa s kým poradiť a k lekárovi je to 40 km a bude to pre neho veľmi nepríjemná skúsenosť, ktorú nepodstúpi – ochorenie prejde do chronického štádia, šíri sa ďalej, prenesie sa od partnera na partnerku a od matky na dieťa a stále sa nič nedeje na zastavenie šírenia. Teda aj vzdelanie a prístup k zdravotníckym službám môžu byť nepriamym rizikovým faktorom.

Dá sa tiež predpokladať, že ľudia, ktorí sú konzervatívnejší, religióznejší budú nie úplne ale predsa len o niečo menej náchylnejší k promiskuite, injekčnému užívaniu drog a asi aj návšteve tetovacích štúdií.

Teda aj otázka na hodnoty by mohla byť pre nás užitočná a mohol by to byť dotazník, z ktorého by sme získali informácie o sumárnom skóre. Na druhej strane, ľudia, ktorí majú sklony nadmerne piť, fajčiť, alebo experimentovať s marihuanou by mohli byť náchylnejší aj k iným druhom rizikového správania.

Pri každom z týchto konceptov bude množstvo okolností, ktoré budú moderovať ich súvislosť s výskytom hepatitídy B. Napríklad hýril v Thajsku, ale používal prezervatív; je veľmi religiózna, ale zamilovala sa do Anatolija a ten mal bujarejšiu mladosť; prešiel všetky väzenia, ale má strašný strach z ihly, takže nemá ani jedno tetovanie. Niektoré okolnosti sú veľmi výrazné a môžeme sa ich pokúsiť zaradiť do analýzy ako moderátory. V každom prípade, každý konkrétny prípad má svoj príbeh, šťastie alebo smolu, ale pre nás je dôležité nájsť zmysluplnú predikciu pre populáciu, veľkú skupinu ľudí, zaujímajú nás len zmysluplne významní hráči rizika.

Ak chceme čokoľvek merať, porovnávať, odhadovať musíme používať číselné kódy, číselné premenné. Tie môžu mať ale odlišný charakter, a to určuje, ako sa s nimi potom môže alebo nemôže pracovať.

Premenné môžeme rozdeliť na **kontinuálne** (spojité) a **kategorické** (diskrétne). Kontinuálne premenné používajú nejaké jednotky merania, napríklad kilogramy, metre, litre, pričom platí, že každý dielik na stupnici je rovnako veľký, ťažký, dlhý či objemný. Päťcentimetrový rozdiel medzi 165 a 170 cm je úplne rovnaký ako medzi 170 a 175 cm.

Kategorické premenné používajú kategórie, a keďže sú to tiež čísla, môžu v databáze vyzerať úplne rovnako ako kontinuálne premenné. Pacienta sa opýtame, ako sa má, a on odpovie „výborne“, zaznačíme 1; „veľmi dobre“ zaznačíme 2; „nie veľmi dobre“ zaznačíme 3; „zle“ zaznačíme 4. V tomto prípade však už neplatí, že rozdiel medzi 1 a 2, teda výborne a veľmi dobre, je rovnaký ako medzi 2 a 3, teda veľmi dobre a nie veľmi dobre. V tomto konkrétnom prípade môžeme nanajvýš povedať, že výborne bude asi viac ako veľmi dobre a veľmi dobre bude viac ako nie veľmi dobre, a to bude viac ako zle. Vieme teda určiť poradie. Táto premenná má viac kategórii, takže by sme ju mohli nazvať multinominálna, ale keďže vieme určiť poradie, je to aj **poradová premenná**.

Číslami však v databáze budeme mať označené aj premenné, kde sa nedá povedať, čo je viac alebo menej. Napríklad ženy budú označené 1 a muži 2. Alebo krvná skupina A, B, AB, O V týchto prípadoch nemá zmysel uvažovať o usporiadaní či poradí. Takéto premenné sa nazývajú **nominálne premenné**.

Skúsme si teda zhrnúť, čím sa medzi sebou líšia jednotlivé typy premenných.

Spojité (kontinuálne) premenné umožňujú kvantifikovať rozdiely medzi jednotlivými hodnotami, napr. telesná výška, telesná váha, hodnoty krvného tlaku, hodnota kreatinínu, dávka lieku.

Poradové (ordinálne) premenné umožňujú určiť poradie hodnôt, ale nie je možné presne určiť rozdiely medzi jednotlivými hodnotami, napr. typ vzdelania, máte sa výborne/dobre/zle.

U nominálnych premenných hodnoty určujú kategórie a nie je možné ich zoradenie, iba určenie ich početnosti v súbore, napr. krvná skupina, pohlavie, diagnóza, prípad/nepřípad.

Trochu zmätku vnášajú premenné, ktorým sa hovorí multinominálne, lebo samotný názov hovorí len o tom, že majú viac kategórii a medzi nimi v niektorých prípadoch môžeme urobiť poradie a niekedy nie. Typ vzdelania v zmysle učňovské, maturitné, vysokoškolské, je premenná, kde vieme určiť poradie (v dĺžke rokov vzdelania), ale typ vzdelania v zmysle inžinierske, zdravotnícke, pedagogické, prírodovedné, už nie je možné zoradovať a svojou povahou je to nominálna premenná.

PRAKTICKÁ ÚLOHA 3:

Určte typ premennej v nasledujúcich príkladoch. V niektorých prípadoch sa bude hodiť i viac označení.

A) spojitá premenná

B) ordinálna premenná

C) dichotomická premenná

D) multinominálna premenná

E) nominálna premenná

- vek (v rokoch od 15 rokov do 65 rokov)
- krvná skupina (A, B, AB, O)
- pohlavie (muž, žena)
- vek (15-20 rokov, 20-25 rokov, 25-30 rokov)
- hodnotenie zdravotného stavu (vynikajúco, dobre, zle)
- telesná váha (v kg)
- vzdelanie (základné vs. maturitné a viac)
- typ liečby (injekčne, kvapky, tabletky, inhalátor)
- úmrtnosť (prežil, zomrel)

RIEŠENIE PRAKTICKEJ ÚLOHY 3:

vek (v rokoch od 15 rokov do 65 rokov): A) spojitá premenná

krvná skupina (A, B, AB, O): D) multinominálna premenná

poohlavie (muž, žena): C) dichotomická premenná, E) nominálna premenná

vek (15-20 rokov, 20-25 rokov, 25-30 rokov): B) ordinálna premenná

hodnotenie zdravotného stavu (vynikajúco, dobre, zle): B) ordinálna premenná, D) multinominálna premenná

telesná váha (v kg): A) spojitá premenná

vzdelanie (základné vs. maturitné a viac): C) dichotomická premenná

typ liečby (injekčne, kvapky, tabletky, inhalátor): D) multinominálna premenná, E) nominálna premenná

úmrtnosť (prežil, zomrel): C) dichotomická premenná

O type premenných, o type údajov je dôležité premýšľať už pri samotnom plánovaní výskumu. Typ údajov určuje aké analýzy bude možné alebo nebude možné použiť pri samotných štatistických analýzach. Spojité údaje nám dávajú najširšie možnosti a je ich možné aj neskôr v procese analyzovania zmeniť na kategorické. Vytvoriť z kategorických údajov spojité údaje už spätne nie je možné. Na druhej strane pri veľmi detailných spojitých údajoch sa nám môžu stratiť dôležité medzníky, v angličtine označované pojmom „cut offs“. Môžeme pacienta požiadať, aby nám na škále od 0 do 10 povedal, ako veľmi ho to bolí a získať krásnu spojitú premennú zameranú na milimetre presne. Môžeme sa ho opýtať, koľko minút denne strávil chôdzou, ale je otázne, či sa budeme môcť na takúto premennú spoľahnúť a či ju budeme vedieť interpretovať?

Na obrázku číslo 1 môžete vidieť ako naše premenné (údaje) vyzerajú v štatistickej aplikácii.

Na obrázku číslo 1 v jeho vrchnej časti je znázornená tabuľka - výstup zo štatistického programu SPSS, ktorý má v sebe vždy databázu nameraných údajov (označenie číslom 1).

V spodnej časti obrázka číslo 1 je znázornená tabuľka – databáza premenných (označenie číslom 2).

Napriek tomu, že je to databáza čísel, umožňuje zobrazovanie takzvaných štítkov (label), teda čo tie čísla znamenajú. Je to zobrazené na obrázku v tabuľke v pravom dolnom rohu (označenie číslom 3).

	IDENTITY	SEX	AGE	SRH	HEADACHE	HEIGHT	WEIGHT	BMI	NOTE
1	1	chlapec	11,5	Velmi dobré	Nie, ani raz	146,40	28,20	13,20	krátky čas administrácie
2	2	dievča	11,3	Dobré	Áno, raz	147,80	39,00	17,90	časový stres
3	3	dievča	11,4	Velmi dobré	-99,00	159,10	43,00	17,00	
4	4	dievča	11,0	Zlé	Áno, viac ako raz	154,50	45,40	19,00	problém pri administrácii
5	5	dievča	11,1	Vynikajúce	Nie, ani raz	148,10	38,80	17,70	
6	6	dievča	10,9	Nie veľmi d...	Nie, ani raz	142,90	42,30	20,70	
7	7	dievča	11,5	-99,00	Áno, viac ako raz	154,70	48,50	20,30	
8	8	dievča	11,3	Velmi dobré	Nie, ani raz	160,80	38,70	15,00	
9	9	dievča	11,2	Vynikajúce	Nie, ani raz	157,20	48,90	19,80	časový stres
10	10	dievča	11,8	Dobré	Áno, viac ako raz	149,50	44,80	20,00	časový stres
11	11	chlapec	11,6	Vynikajúce	Nie, ani raz	136,70	26,70	14,30	
12	12	chlapec	11,0	Dobré	Nie, ani raz	150,80	35,80	15,70	

Name	Type	Wi...	De...	Label	Values	Missing	
1	IDENTITY	Numeric	8	0	Identifikačné číslo	None	None
2	SEX	Numeric	8	2	Pohlavie	{1,00, chlapec}...	-99,00
3	AGE	Numeric	5	1	Vek v čase zberu údajov	None	-99,0
4	SRH	Numeric	8	2	Aké je tvoje zdravie v porovna	{1,00, Vynikajúce}...	-99,00
5	HEADACHE	Numeric	8	2	Bolela ťa hlava v uplynulom týždni	{1,00, Nie, ani raz}	-99,00
6	HEIGHT	Numeric	8	2	Výška	None	-99,00
7	WEIGHT	Numeric	8	2	Váha	None	-99,00
8	BMI	Numeric	8	2	Body mass index	None	-99,00
9	NOTE	String	40	0	Poznámka	None	None

Value Labels
Value: <input type="text"/>
Label: <input type="text"/>
1,00 = "Nie, ani raz"
2,00 = "Áno, raz"
3,00 = "Áno, viac ako raz"
Buttons: Add, Change, Remove
Buttons: OK, Cancel, Help

Obrázok číslo 1: Výstup zo štatistického programu SPSS.

V stĺpci HEADACHE (číslo 4) by sme mali mať jednotky, dvojky a trojky, ale my tam máme „nie, ani raz“, „áno, raz“, „-99“, „áno viac ako raz“. Ale v skutočnosti táto aplikácia pracuje len s číslami.

Každý riadok predstavuje jedného respondenta a je zviazaný, teda údaje medzi riadkami sa nemiešajú v exceli. Ak si pozriete riadok, pri ktorom je číslo 5, tak to je respondent číslo 1, je to chlapec, mal 11 a pol roka, svoje zdravie hodnotil ako veľmi dobré, hlava ho nebolela ani raz, bol vysoký 146 cm, mal 28 kíl a jeho body mass index (BMI) bol 13,2 a navyše o ňom vieme, že dotazník vyplnil extrémne rýchlo.

Každý respondent musí mať unikátne identifikačné číslo, aby v databáze nedochádzalo k duplikáciám. Takto dokážeme neskôr pridávať k jednotlivým respondentom ďalšie údaje, spojiť súbory, rozdeľovať ich na podsúbory a pridávať ďalších respondentov. Týmto spôsobom vieme zároveň chrániť osobné identifikačné údaje, ktoré by nikdy nemali byť uložené v databáze spolu s údajmi o respondentoch, ale mali by byť v samostatnom súbore. To je najjednoduchší spôsob, ako predísť úniku osobných informácií, čo je aj podmienka práce s údajmi. Asi nikto by nechcel nájsť svoje meno s výsledkami citlivého vyšetrenia u nejakého štatistika.

Každá premenná má svoj krátky názov a popis. Kým máte v databáze päť premenných a pracujete s nimi každý deň, asi si budete pamätať čo, ktorá skratka znamená. Keď tých premenných budú stovky, prípadne jedna premenná bude mať viac variácií, jediný spôsob, ako si v tom udržať poriadok, bude popísať ich priamo v databáze. Napríklad premenná „nadváha“ môže mať rôzne verzie: hodnota BMI, rozdelenie na tri kategórie podľa BMI alebo subjektívne hodnotenie nadváhy respondentom. Je dôležité popisovať premenné čo najpresnejšie hneď na začiatku. Napríklad, ak sa pozriete na riadok, pri ktorom je číslo 6, krátky názov je HEADACHE, popis je znenie otázky „Bolela tá hlava v uplynulom týždni?“

Keď sa vrátíme k číslu 3, máme tam nasledovný popis kódovania: číslo 1 znamená „nie, ani raz“, číslo 2 znamená „áno, raz“ a číslo 3 „áno, viac ako raz“. Spôsob ako sú premenné kódované určuje interpretáciu zistení, takže to nie je úplne jedno či je 1 „nie, ani raz“ a 3 „áno, viac ako raz“, alebo naopak. Rozumným pravidlom je, ak sa prípad označuje číslom 1 a neprípad číslom 0. V takom prípade sa exponovanie, výskyt rizikového faktora označuje 1 a nevýskyt 0. Dodržiavanie tohto pravidla potom výrazne zľahčuje čítanie výstupov z analýz.

Na obrázku vyššie pod číslom 7 je zároveň respondentka číslo 7 vo veku 11 a pol roka, ktorá neodpovedala na otázku aké je jej zdravie. To vieme podľa kódu -99, ktorý sme si zvolili na označovanie takýchto prípadov. Štatistická aplikácia nám umožňuje nastaviť viacero kódov, ktorými budeme označovať prípady, ktoré nezahrnieme do analýzy. Údaje môžu chýbať z viacerých dôvodov: napríklad pre chybu pri zápise (prázdne miesto), prípadne vieme, že respondent neodpovedal z dôvodu, že odpoveď nie je relevantná (napr. otázka „Ako ťažko sa Vám chodí?“ u pacienta, ktorý je na vozíku). So štatistickou aplikáciou dokážeme presne určiť, kedy a ktoré kódy sa majú zahrnúť do analýzy a kedy nie. Chýbajúce údaje musia byť označené, aby sa dalo odlíšiť, kedy došlo k chybe a kedy je to skutočne chýbajúci údaj, prípadne aký je dôvod chýbajúceho údaju.

Štatistická aplikácia umožňuje zapisovať aj textové premenné (označené ako „string“), ktoré nájdete pod číslom 8 a číslom 9. V tomto prípade je to premenná NOTE (poznámka), kde administrátor môže zapísať zvláštnosti pri administrovaní. Tieto informácie nám môžu potom pomôcť interpretovať niektoré zistenia. S textovými premennými ale štatistická aplikácia nevie pracovať. Ak by sme s nimi chceli pracovať, museli by sme z nich vytvoriť kategorickú premennú.

Zhrňme si, čo je dôležité vedieť o databáze skôr, ako zaklopete na dvere štatistikovi:

- každý respondent musí mať unikátne identifikačné číslo
- jeden riadok sa rovná jeden respondent a tento riadok musí byť zviazaný
- aj vďaka identifikačným číslam, štatistické programy umožňujú pridávať respondentov, premenné, vytvárať podsúbory, usporadúvať súbory, takže to nerobte v Exceli

- každá premenná musí mať názov a jednoznačný popis
- kategorické premenné sú prekódované do čísel a musia mať menovky, teda čo, ktoré číslo znamená
- spôsob kódovania je veľmi dôležitý
- chýbajúce údaje musia byť označené, aby sa dalo odlíšiť, kedy došlo k chybe a kedy je to skutočne chýbajúci údaj, prípadne, aký je dôvod chýbajúceho údaju
- štatistické programy umožňujú zapisovať aj textové premenné, ale nedokážu s nimi pracovať

1.4. Porozumenie procesu transferu populácie do výskumnej vzorky a rozlišovanie spôsobu výberu vzorky

V tejto časti si povieme, akým spôsobom uvažujeme o skupine ľudí, na ktorých by sa mali vzťahovať naše zistenia, teda o populácii, a ako môžeme vybrať výskumnú vzorku (výskumný súbor), aby nebolo nutné skúmať celú populáciu.

V niektorých typoch výskumu nie je problém získať údaje o všetkých členoch populácie a zozbierať dáta od tisícov respondentov. V iných prípadoch však býva výzvou zhromaždiť údaje čo i len od 25 respondentov, čo je často minimálny počet pre štatistické závery. Veľmi často sa zabúda na to, že vzorka bez populácie je nanič. Typickým príkladom je online prieskum, ktorý je rozvešaný po sociálnych sieťach. Aj keď získate údaje od tisícov respondentov, pokiaľ neviete, koho títo respondenti reprezentujú, nemožno na základe zistení robiť širšie závery. Uvedme si príklad, ako pristupovať k výberu vzorky. Na akú skupinu ľudí chcete, aby sa vzťahovali vaše zistenia a odhady? Akú populáciu reprezentujú alebo nereprezentujú pacienti na vašom oddelení, vo vašej ambulancii, či v školskej triede?

Predstavme si, že sme sa rozhodli skúmať zdravie detí v predškolskom veku, teda deti vo veku od troch do šiestich rokov. Na Slovensku ich máme približne 175 000. Vybrali sme deti hospitalizované na infekčnom oddelení a chceme skúmať rizikové faktory výskytu hnačky u detí predškolského veku. Povieme si, že ročne hospitalizujeme na našom oddelení asi 100 takýchto detí, čo je slušné číslo a máme ich po ruke, keďže u nás strávia aj tri dni.

V tomto veku hádam každé dieťa aspoň raz prekoná hnačku, a deti na našom oddelení sú z rôznych skupín, takže by mohli reprezentovať predškolákov. Problémom však je, či sa k nám dostanú všetky deti, ktoré trpia hnačkou? Pravdepodobne nie. Prvou podmienkou je, že rodič musí stav považovať za vážny. Niektorí prídu pri prvých príznakoch, zatiaľ čo iní prichádzajú až vo vážnom stave. Ak je to skutočne vážne, asi sa k nám ten pacient dostane.

Niektorí rodičia majú dostatočné vedomosti a skúsenosti, takže zvládnu aj dosť náročný stav doma a nemusia byť ich dieťa nikdy hospitalizované. V niektorých prípadoch lekár zvažuje, či to rodičia zvládnu a dieťa radšej hospitalizuje, než by riskoval, že príde ešte v horšom stave. Výsledkom je, že ľahké prípady hnačky sa v našej vzorke takmer nevyskytujú; stredne ťažké prípady máme vtedy, ak je nejaký problém v rodinnom zázemí alebo sú rodičia neskúsení či úzkostní. Ťažké prípady sú takmer vždy hospitalizované.

Teraz sme si popísali, akú skupinu vzorka hospitalizovaných detí reprezentuje a koho nereprezentuje.

Ako **populáciu** označujeme skupinu ľudí so spoločnými charakteristikami, na ktorých by sa mali vzťahovať výsledky výskumu a ktoré by sa mali dať **zovšeobecniť**. Spoločnými charakteristikami môže byť vek, miesto, kde žijú, diagnóza alebo fakt, že sa zúčastnili nejakej udalosti. Táto definícia môže byť

fixná, definovaná na trvalo (napr. pacienti po srdcovom infarkte), alebo **dynamická**, definovaná aspektom, ktorý sa môže meniť (napr. ľudia, ktorí ochoreli na chrípku).

Populácia, v ktorej chceme niečo odhadovať, nie je vždy tá istá ako populácia, z ktorej sa vyberá vzorka; táto sa označuje ako **náhradná** populácia. Napr. študuje sa niečo u potkanov, aby sa odhadovalo niečo u ľudí, alebo údaje z roku 2008 sa aplikujú na rok 2009.

Ak by napríklad boli v „populácii“ len rovnaké červené a rovnaké biele guľičky, potom stačí vybrať jednu červenú a jednu bielu guľičku, zmerať ich a spočítať počet bielych a červených guľičiek vo vrecúšku. V reálnom výskume však nikdy nevieme presne, čo vo vrecúšku (v populácii) je – koľko akých guľičiek tam je - a nedokážeme zohľadniť všetky atribúty rôznych druhov guľičiek.

Spôsobov výberu vzorky existuje viacero, postupne si ich v tejto časti popíšeme a uvedieme aj praktické príklady. Začneme uvažovaním, prečo skúmané parametre nemôžeme zmerať u všetkých v danej populácii. Napríklad zozbierať genetický materiál v celej populácii Islandu, urobiť biopsiu u všetkých pacientov, zmerať percento tuku u všetkých detí vo veku 11 rokov?

Dôvodov, prečo nedokážeme zmerať parametre u všetkých, je viacero. Napríklad, nie každý respondent musí zdieľať naše nadšenie z výskumu a nemusí súhlasiť s účasťou na výskume. Niektorí respondenti by aj súhlasili, ale nedostanú sa k nám, a ani my k nim. Niektoré vyšetrenia sú poriadne nepríjemné alebo obťažujúce a vážne tým nechceme obťažovať viac ľudí ako je nevyhnutné, respektíve sme celkom radi, ak aspoň niekto súhlasí s účasťou na výskume.

Niektoré vyšetrenia alebo procedúry môžu byť rizikové a určite ich nechceme vyžadovať od pacientov len kvôli výskumu, a tak nám ostanú vo vzorke len tí, ktorí by ich museli tak či tak podstúpiť. Niektoré vyšetrenia sú veľmi drahé, takže aj keby bol záujem zo strany výskumníkov a respondentov, finančný rozpočet dovolí iba istý počet vyšetrení. Alebo niektoré vyšetrenia sú aj relatívne lacné, ale vynásobené miliónmi obyvateľov predstavujú veľmi vysoké náklady.

Ďalej, ako vieme, že to, čo zistíme v istej špeciálne vybranej skupine, je možné zovšeobecniť na celú populáciu? Nebudeme merať daný jav u všetkých, ale ako vieme, že to, čo zistíme v istej špeciálne vybranej skupine, je možné zovšeobecniť na celú populáciu? Čo ak budeme mať vo vzorke len modrookých ľudí a u hnedookých ľudí to bude úplne inak? Pre vedcov je **reprezentatívny výber** jedinou podloženou procedúrou na to, aby výber jedincov mohli použiť na zovšeobecnenie vzhľadom k populácii. Reprezentatívny výber nám tak môže nahradiť skúmanie celej populácie.

Zlatým štandardom pre reprezentatívny výber je **náhodný výber**. Pri náhodnom výbere musí mať každý jedinec danej populácie rovnakú šancu byť vybraný do našej vzorky. Dá sa to urobiť tak, že každému z populácie pridáme lístok, schováme ho do klobúka a budeme losovať.

Otázkou však je, prečo náhodný výber? Prečo musíme losovať z klobúka a nemôžeme vybrať toho, koho máme naporúdzi, kto je ochotný vyplniť dotazník alebo podstúpiť vyšetrenie? Odpoveďou je, že my nevieme a nedokážeme zohľadniť všetky atribúty, ktoré reprezentujú populáciu. Nevieme, koľko a akých guľičiek je vo vrecúšku. Ak to necháme na náhodu a každá guľička bude mať rovnakú šancu, s veľkou pravdepodobnosťou sa nám pri opakovaných výberoch, teda ak by sme znova a znova losovali istý počet guľičiek z toho istého vrecúška, podarí dostať veľmi podobné zloženie skupiny, kde budú zastúpené všetky druhy guľičiek.

Samozrejme, záleží na tom, koľko tých guľičiek je vo vrecúšku, koľko je druhov a koľko ich budeme losovať. Ak by sme sa snažili vybrať povedzme tri biele, dve červené a tri zelené, tak by všetky fialové a ružové mohli ostať neobjavené. Manipulovanie s výberom, teda snaha určiť, čo je reprezentatívne, vybrať biele, červené a zelené guľičky, môže byť spojené s chybou, tzv. skreslením pri výbere vzorky, ktoré môže výrazne ovplyvniť zistenia (anglicky selection bias).

Losovanie z klobúka nie je vždy možné a často je to dosť komplikované a náročné na logistiku aj financie. Predstavte si, že by ste chceli náhodne vyberať pacientov vedených v ambulancii, ale jeden, ktorého ste vybrali, tu bol práve včera a ďalší je objednaný až o dva mesiace. Vy potrebujete v nejakom ohraničenom čase zozbierať vzorky a odnieť ich do laboratória. Môžete si teda povedať, že zahrniete každého tretieho pacienta, ktorého ošetríte. Čo však ak vaša sestrička má pravidlo, že po dvoch ťažkých pacientoch vám dopraje pauzu a vyberie ľahšieho? Ak by ste brali každého tretieho pacienta, mali by ste vo vzorke len tých „oddychových pacientov“.

Alebo robíte prieskum v domácnostiach a nemôžete navštíviť všetky domácnosti, tak si poviete, že pôjdete do každej s párnym popisným číslom. Čo však s ulicou, kde sú na jednej strane honosné vily a na druhej bytovky so sociálnymi bytmi? A ako naschvál, všetky párne sú na jednej strane, teda nedostanete sa do žiadnej vily. Takémuto výberu hovoríme **systematický výber**, pretože vyberá respondentov podľa nejakého pravidla. Problémom môže byť, ak usporiadanie respondentov v zozname nie je náhodné a vnáša systematickú chybu to výberu vzorky.

Môže sa však stať aj to, že náhodný výber nám nebude stačiť, že nedokážeme urobiť reprezentatívny výber pomocou náhodného výberu. Môže sa to stať v prípade, keď máme veľmi malú vzorku a nechceme, aby sa stalo, že tam budeme mať len biele a žiadnu čiernu guľičku, ktorých bolo vo vrecúšku málo. Vieme, že v populácii sú aj zriedkaví respondenti, ktorí sú pre výskum dôležití a potrebujeme ich mať vo vzorke. Aby sa do vzorky dostal dostatočný počet iných, zriedkavých respondentov, potrebovali by sme veľmi veľký výberový súbor.

Uveďme si príklad: Košice majú približne 250 000 obyvateľov a v istom roku mali registrovaných 800 ľudí žijúcich na ulici. Akú veľkú vzorku by sme potrebovali, aby v nej pri jednoduchom náhodnom výbere bolo aspoň 15 z týchto ľudí? Na 1000 obyvateľov Košíc pripadnú približne 3 registrovaní bezdomovci. Pre slušnú analýzu by bolo dobré ich mať aspoň 15 vo vzorke, teda v ideálnom prípade potrebujeme skupinu 5000 respondentov. Reálne ich ale podľa zoznamu obyvateľov nenájdeme, lebo obvykle nemajú adresu. Veľmi ľahko ich ale nájdeme na miestach, kde sa zhromažďujú, takže v ich prípade je namiesto losovania lepšie urobiť si špeciálnu podskupinu vo vzorke. Situácia, ktorú sme popísali v príklade, sa volá **stratifikovaný výber**. Populácia sa rozdelí do skupín, v ktorých sa robí náhodný výber, pričom sa dbá na zachovanie proporcionálneho zastúpenia vo vzorke pri výbere alebo v analýzach.

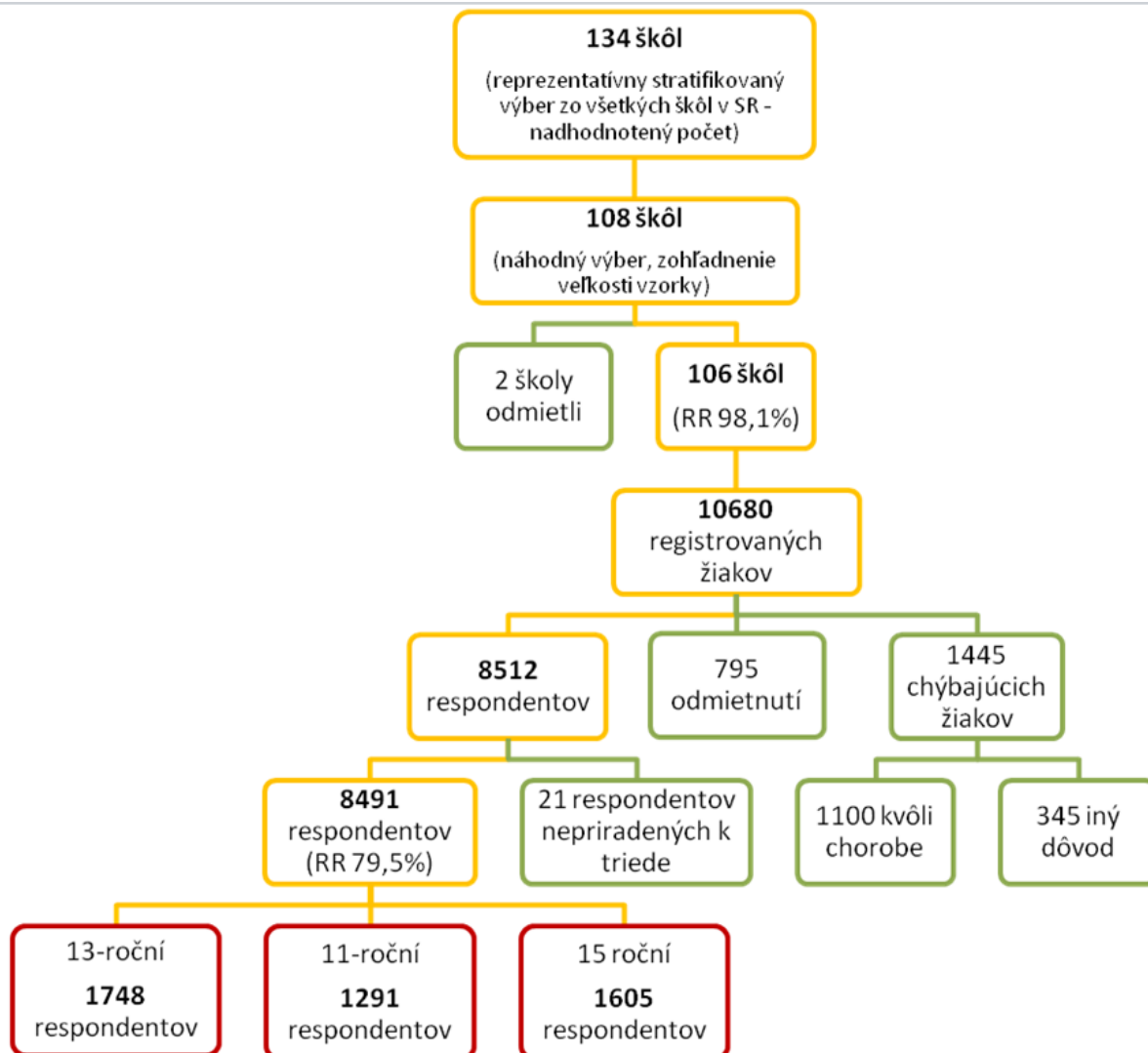
Iným spôsobom, ktorý nám môže uľahčiť náhodný výber, je **trsový výber**. O niektorých respondentoch vieme, že sú prirodzene zoskupení do trsov, teda do skupín, napríklad rajóny obvodných lekárov, školské triedy či pracovné kolektívy. Použitie náhodný výber na populáciu školákov by znamenalo zobrať zoznam všetkých žiakov všetkých škôl a vylosovať napríklad tisíc žiakov. Dvaja by boli zo základnej školy v Trstenej, traja v Bratislave na Vajnorskej, jeden v Ubli a tak ďalej. Bolo by to neuveriteľne logisticky i finančne náročné.

1000 žiakov však vieme vybrať aj tak, že náhodne vyberieme 50 škôl, čo by mohlo pokryť rôznorodosť populácie na celom Slovensku, a na každej škole náhodne vyberieme jednu triedu s 20 žiakmi. Losujeme teda trsy a nie jednotlivých respondentov.

Výber vzorky môže byť viacstupňový proces a na posúdenie reprezentatívnosti vzorky, jej silných aj limitujúcich aspektov, je potrebné zozbierať čo najviac údajov v procese výberu vzorky a popísať ich v metódach. Obrázok číslo 2 je príkladom popisu stratifikovaného výberu respondentov štúdie o zdraví školákov s názvom Health Behaviour in School Aged Children (HBSC) na Slovensku.

Ak pôjdeme zhora, prvým krokom bolo, že zo zoznamu všetkých základných škôl v danej krajine bolo náhodne vybraných 134 škôl. Z tohto zoznamu bolo oslovených 108 škôl, pričom 2 školy odmietli

spoluprácu. V texte je popísaný aj dôvod odmietnutia. V školách, ktoré súhlasili, bolo registrovaných 10680 žiakov v zodpovedajúcej vekovej skupine. 795 žiakov odmietlo účasť, 1445 nebolo v škole v čase zberu údajov, z toho 1100 kvôli chorobe a 345 z iného dôvodu. 8512 respondentov sa zapojilo do vyplňania dotazníkov, ale 21 dotazníkov sme nevedeli priradiť k žiadnej triede, a tak sme ich museli vyradiť.



Obrázok číslo 2: Popis výberu vzorky.

1.5. Formulácia výskumnej otázky

Na sformulovanie výskumnej otázky potrebujeme expozíciu, dôsledok, skresľujúce premenné, ich vzájomné vzťahy a populáciu, v ktorej by to malo platiť. **Výskumná otázka** skúma, či expozícia nejakému javu alebo okolnosti zvyšuje, respektíve znižuje riziko výskytu alebo závažnosti vybraného javu či okolnosti, a či sa na tomto vzťahu podieľajú ďalšie javy a okolnosti, ktoré daný vzťah sprostredkujú, ovplyvňujú, alebo len skresľujú.

Nie vždy je úplne ľahké usporiadať expozície, dôsledky a ďalšie premenné, pretože skutočné interakcie medzi javmi a okolnosťami sú veľmi zložité. V drvivej väčšine prípadov však výskumný dizajn, existujúce údaje a kapacity výskumníkov neumožňujú analyzovať všetko a už vôbec nie v jednej výskumnej otázke.

Preto je nutná istá miera redukcie a vysoká miera disciplíny, najmä v tom, čo v danom konkrétnom prípade predstavuje expozíciu, ktorej vplyv či súvislosť s predpokladaným dôsledkom budeme skúmať.

Ako príklad nás zaujíma, či nadváha a obezita súvisia s psychickou nepohodou, smútkom až depresiou. Nadváha a obezita nie sú okolnosti, ktoré sa menia zo dňa na deň, takže môžeme predpokladať, že respondenti sú už nejakú dobu exponovaní. V našom prípade teda bude expozíciou nadváha a obezita. Dôsledok, ktorý chceme skúmať, je psychická nepohoda. Výskumná otázka by teda znela: „Má nadváha a obezita vplyv na psychickú nepohodu?“ Zároveň vieme, že psychická nepohoda mohla a môže byť jedným zo spúšťačov nerovnováhy, ktorá viedla k nadváhe a obezite. Teda úlohu expozície a dôsledku by si tieto koncepty mohli aj vymeniť. V skutočnosti pravdepodobne dochádza k nepretržitej interakcii medzi týmito dvoma premennými. Rozhodnutie, ktorá premenná je expozícia a ktorá je dôsledok vo výskumnej otázke, teda nemusí byť vôbec jednoduché.

OTÁZKY

1. Ktoré atribúty musí spĺňať vedecká metóda poznávania? Označte všetko, čo sa hodí.
 - A) Zistenia nesmú byť spochybniteľné, ale musia mať definitívnu, nemennú povahu
 - B) Musí byť možné empiricky overiť zistenia
 - C) Skúmanie musí byť systematické a organizované
 - D) Musí byť možné opakovane overiť zistenia
 - E) Výskyt alebo nevýskyt daného parametra musí byť posudzovaný objektívne
 - F) Zistenia je možné kedykoľvek spochybniť, skontrolovať a v prípade chyby, rozporu aj opraviť
 - G) Tvrdenie musí byť logické, a v takom prípade nemusí byť empiricky overiteľné
 - H) Ak sa jav niekoľkokrát zopakoval, je možné uvažovať, že je to pravidlo a nie je nutné to ďalej systematicky skúmať
 - I) Názor skúseného experta je možné považovať za zistenie

2. Aký typ premennej je vek meraný v rokoch?
 - A) spojitá
 - B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna

3. Aký typ premennej je vek, ak respondentov rozdelíme do vekových skupín do 60 rokov, 61-75 rokov a nad 75 rokov?
 - A) spojitá
 - B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna

4. Aký typ premennej je vek, ak respondentov rozdelíme do vekových skupín do 60 rokov a nad 60 rokov?
 - A) spojitá
 - B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna

5. Aký typ premennej je krvná skupina?
 - A) spojitá
 - B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna

6. Aký typ premennej je hodnotenie zdravia, ak si respondenti vyberajú z nasledujúcich odpovedí (1) výborné zdravie, (2) dobré zdravie, (3) nie veľmi dobré zdravie?
 - A) spojitá

- B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna
7. Aký typ premennej je typ liečby, ak je na výber (1) injekčne, (2) kvapky, (3) tabletky, (4) inhalátor?
- A) spojitá
 - B) ordinálna (poradová)
 - C) kategorická, multinominálna
 - D) kategorická, dichotomická
 - E) kategorická, nominálna
8. O aký spôsob výberu vzorky z populácie ide v popísanom príklade? Lístky s menami všetkých pacientov dáme do klobúka a náhodne vytiahneme potrebný počet lístkov.
- A) jednoduchý náhodný výber
 - B) systematický výber
 - C) stratifikovaný výber
 - D) trsový výber
9. O aký spôsob výberu vzorky z populácie ide v popísanom príklade? Z abecedného zoznamu pacientov vyberieme každého tretieho.
- A) jednoduchý náhodný výber
 - B) systematický výber
 - C) stratifikovaný výber
 - D) trsový výber
10. O aký spôsob výberu vzorky z populácie ide v popísanom príklade? Rozdelíme zoznam pacientov na mužov a ženy. Lístky s menami mužov dáme do jedného klobúka a lístky s menami žien dáme do druhého klobúka a z oboch klobúkov vytiahneme potrebný počet lístkov.
- A) jednoduchý náhodný výber
 - B) systematický výber
 - C) stratifikovaný výber
 - D) trsový výber
11. O aký spôsob výberu vzorky z populácie ide v popísanom príklade? Zo zoznamu ambulancií náhodne vyberieme 10 ambulancií. Lekárov v týchto ambulanciách požiadame o zoznamy pacientov, z ktorých náhodne vyberieme potrebný počet pacientov.
- A) jednoduchý náhodný výber
 - B) systematický výber
 - C) stratifikovaný výber
 - D) trsový výber

2. MERANIE VPLYVU

V tejto kapitole sa naučíme popísať spojité premenné strednými hodnotami, ako sú priemer, medián a modus. Predstavíme si rôzne spôsoby, ako upraviť údaje na štatistické premenné, napríklad kategorizáciou, klastrovaním či vytváraním faktorov. Naučíme sa rozlišovať, kto bol a kto nebol exponovaný skúmanej okolnosti alebo javu, a kto je alebo nie je prípad. Taktiež si ukážeme, ako popísať frekvencie prípadov, teda incidenciu a prevalenciu. Na záver sa naučíme, ako popísať vplyv expozície na výskyt prípadov, teda vypočítať relatívne riziko, pomer šancí a vedieť interpretovať intervaly spoľahlivosti.

2.1 Popis spojitých premenných strednými hodnotami

Stredné hodnoty nám pomáhajú porovnávať spojité premenné a poskytujú informácie o tom, čo je typické, najčastejšie, alebo okolo ktorej hodnoty sú tie ostatné údaje rozložené. Okrem nich existujú aj ďalšie parametre, ktoré takúto premennú popisujú.

Na obrázku číslo 1 je výstup zo štatistického softvéru IBM SPSS Statistics, ktorý popisuje premennú vek. Vieme z neho prečítať, že vo vzorke bolo 8902 respondentov, žiadnemu nechýbal údaj o veku, priemerný vek bol 13,37 rokov, so smerodajnou odchýlkou $\pm 1,44$ roka. Najmladší respondent mal 10 rokov a najstarší mal 17 rokov. Približne polovica respondentov mala vek do 13,39, respektíve nad 13,39 rokov, štvrtina menej ako 12,15 alebo viac ako 14,57. Najčastejšie opakujúcou sa hodnotou bolo 11,641342.

Age of respondent continual (vypocítane z r

N	Valid	8902
	Missing	0
Mean		13,37111318
Std. Error of Mean		,015292487
Median		13,39767283
Mode		11,641342 ^a
Std. Deviation		1,442852451
Variance		2,082
Minimum		10,187543
Maximum		17,292266
Percentiles	25	12,15537303
	50	13,39767283
	75	14,57905544

Obrázok číslo 1: Výstup zo štatistického softvéru.

Teraz si bližšie popíšeme stredné hodnoty. Na obrázku číslo 2 je frekvenčná tabuľka výšky plátov v jednej firme, sú tu uvedené priemerné čisté zárobky zamestnancov. Táto firma má jedného šéfa s platom 6000 €, dvoch zástupcov s platom 3000 €, dvoch vedúcich dielní s plátmi 764 € a 724 €, a dvadsať zamestnancov s plátmi od 300 € do 654 €. Z dolnej tabuľky na obrázku číslo 2 vieme prečítať, že firma má celkovo 25 zamestnancov, pričom minimálny plat je 300 € a maximálny 6000 €. Priemerný zárobok je 908 € ± 1275 €.

Priemer (anglicky mean) vypočítame tak, že spočítame zárobky všetkých zamestnancov a vydáme ich počtom zamestnancov. **Medián** je hodnota, ktorá rozdeľuje súbor na polovice, teda polovica respondentov má nižší a polovica má vyšší plat ako je táto hodnota. V tomto prípade je to hodnota 500 €. **Modus** je hodnota, ktorá sa v súbore vyskytuje najčastejšie. Piati z dvadsaťpäť zamestnancov zarobili rovnakú sumu 300 €, teda modus je v tomto prípade 300.

V danom súbore teda poznáme priemer, medián i modus. Ktorá z týchto stredných hodnôt najlepšie popisuje platovú situáciu vo firme? Priemerný zárobok je cez 900 €, ale prevažná väčšina zamestnancov

má aspoň o 200 € nižší zárobok. Priemerná vzdialenosť hodnôt od priemeru, teda smerodajná odchýlka je 1275 €, ale maximálny rozdiel v zárobkoch je 5700 €, teda skoro 20 násobok minimálneho zárobku. Len traja zamestnanci majú nadpriemerný zárobok. Priemer nám v tomto prípade asi nepovie, s akým zárobkom máme počítať v tejto firme.

Najčastejšie majú zamestnanci zárobok 300 €, ale týka sa to len 5 zamestnancov, takže to tiež nebude dobrý ukazovateľ. Polovica zamestnancov má zárobok nižší ako 500 € a polovica vyšší ako 500 €. V tomto konkrétnom prípade ak by sme teda neašpirovali na miesto šéfa firmy, mali by sme očakávať plat v tejto firme okolo 500 €, lebo túto premennú v tomto prípade najlepšie popisuje medián. Ale v každom konkrétnom prípade to môže byť inak, preto odporúčam nespoliehať sa len na priemer, ale dobre si skontrolovať aj všetky ostatné parametre danej premennej.

	Frequency	Valid Percent	Cumulative Percent
300,00	5	20,0	20,0
450,00	4	16,0	36,0
453,00	1	4,0	40,0
455,00	2	8,0	48,0
500,00	1	4,0	52,0
521,00	1	4,0	56,0
531,00	1	4,0	60,0
534,00	1	4,0	64,0
541,00	1	4,0	68,0
623,00	1	4,0	72,0
654,00	2	8,0	80,0
724,00	1	4,0	84,0
764,00	1	4,0	88,0
3000,00	2	8,0	96,0
6000,00	1	4,0	100,0
Total	25	100,0	

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
VAR00001	25	5700,00	300,00	6000,00	908,3600	1274,49310	1624332,657
Valid N (listwise)	25						

Obrázok číslo 2: Výstup zo štatistického softvéru.

2.2 Spôsoby úpravy údajov na štatistické premenné

Akým spôsobom by sme pracovali s premennými, ktoré nameriame ako spojité (napr. krvný tlak, BMI, percento telesného tuku, telesná teplota) ak vieme, že klinicky zaujímavý je len istý rozsah, alebo len hodnoty nad nejakú konkrétnu hranicu? Ak v literatúre existujú odporúčania, *napr. štandardné hodnoty pre vekové skupiny*, alebo sú nejaké deliace kritéria (*anglicky cut off*) zaužívané, odporúča sa premenné kategorizovať na základe týchto štandardov. Spojitú premennú môžeme dichotomizovať, rozdeliť do dvoch kategórií, podľa strednej hodnoty, teda priemeru alebo mediánu. Môžeme ju tiež rozdeliť na tri kategórie, a to oddelením vrchného a spodného kvartilu, teda 25%. Môžeme vytvárať kategórie tak, aby sme zachovali dostatočný počet respondentov v každej skupine, teda 10 alebo 15. Alebo môžeme zvoliť intervaly v spojitej premennej, napríklad vytvoriť vekové kategórie po 5 rokov, teda 10 až 15, 15 až 20, 20 až 25 a tak ďalej.

Na obrázku číslo 3 je výstup zo štatistického softvéru. V prvom stĺpci sú hodnoty, v druhom stĺpci je uvedené koľko respondentov malo túto hodnotu premennej a v nasledujúcom je táto početnosť vyjadrená v percentách. „Valid percent“ je vyjadrenie v percentách po odrátaní chýbajúcich údajov. V tomto prípade máme jedného respondenta s chýbajúcim údajom. Posledný stĺpec je kumulatívne percentuálne zastúpenie, teda akoby spočítavané od prvého po posledný riadok. Ak si všimnete početnosti, tak prvú hodnotu 4,32 mal jeden respondent, druhú 4,58 jeden respondent a tretiu 5,44 dvaja respondenti. Keďže máme spolu 16 respondentov, ktorí vyplnili tento údaj, štyria respondenti

reprezentujú štvrtinu vzorky, a preto je v tomto poslednom stĺpci treťou hodnotou 25%. Je možné, že pre túto premennú existujú nejaké odporúčania deliť ju na tých, ktorí majú hodnotu 6 a viac. Na obrázku číslo 3 je svetlomodrá šípka označujúca hodnoty, ktoré by v takom prípade patrili do prvej kategórie (teda hodnoty nad šípkou) a do druhej kategórie (hodnoty pod šípkou). V prvej kategórii by sme mali 5 respondentov, ktorí by tvorili približne tretinu vzorky. Táto premenná má priemer 6,68, čo je znázornené na obrázku číslo 3 tmavou bodkovanou čiarou. Ak by sme túto skupinu rozdelili na kategórie podľa priemeru, 68,8% respondentov by patrilo do prvej kategórie a 33,2% do druhej kategórie. Mohli by sme použiť medián 6,50 (znázornené ako plná svetlomodrá čiara). V kategóriách by bolo po 50% respondentov.

Čo sa však stane, ak respondentov podľa nejakej hodnoty rozdelíme na polovice? Predstavte si, že je to vek od 30 do 70 rokov a medián je 50 rokov. Skúmate súvislosť veku s nejakým ukazovateľom zdravia. V jednej skupine teda budete mať 30 až 50 ročných a v druhej 51 až 70 ročných. V jednej aj druhej skupine budete mať pomerne veľa ľudí vo veku 50 plus mínus 10 rokov, teda 40 až 50 ročných a 51 až 60 ročných, ktorí môžu byť veľmi veľmi homogénnou skupinou, teda nelíšiť sa veľmi v zdraví. Okrem toho tam bude časť tých 30 až 40 ročných, ktorá na tom bude v priemer zdravotne lepšie a 61 až 70 ročných, ktorá na tom bude zdravotne horšie. Môže sa ale stať, že rozdiel nenájdete, lebo 40 až 60 ročných je viac a sú si výraznejšie podobní. Vy však z praxe viete, že medzi tými mladšími a staršími rozdiel predsa je. V takom prípade si vieme pomôcť oddelením vrchných a spodných kvartilov – čiarkované svetlomodré čiary - a rozdelením na tri skupiny – extrémna horná hodnota, hodnoty okolo stredu, extrémna spodná hodnota. Ak premennú kategorizujeme týmto spôsobom, máme oveľa väčšie šance zistiť, či sa rozdielne hodnoty tejto premennej líšia vo vplyve na inú premennú, v tomto prípade teda vzťah veku so zdravotnými ukazovateľmi.

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4,32	1	5,9	6,2
	4,58	1	5,9	12,5
	5,44	2	11,8	25,0
	5,64	1	5,9	31,2
	6,08	1	5,9	37,5
	6,32	1	5,9	43,8
	6,50	1	5,9	50,0
	6,52	1	5,9	56,2
	6,62	1	5,9	62,5
	6,68	1	5,9	68,8
	7,58	1	5,9	75,0
	7,88	1	5,9	81,2
	8,92	1	5,9	87,5
	9,62	1	5,9	93,8
	10,20	1	5,9	100,0
Total	16	94,1	100,0	
Missing System	1	5,9		
Total	17	100,0		

Obrázok číslo 3: Výstup zo štatistického softvéru.

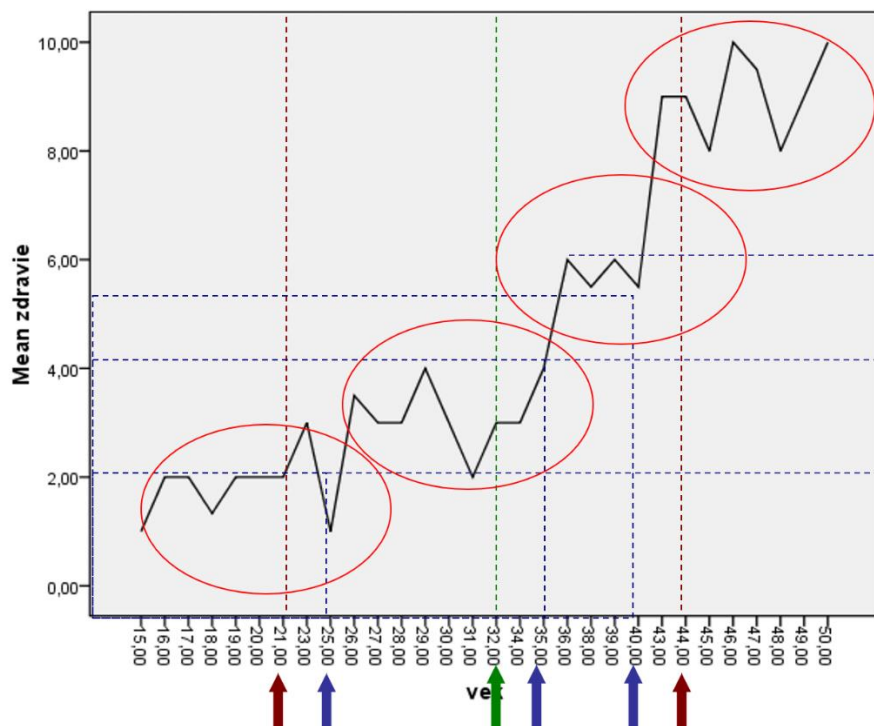
Ukážme si komplikovanejší prípad úpravy údajov na štatistické premenné. Obrázok číslo 4 znázorňuje súvislosť medzi vekom a ukazovateľom zdravia, pričom vyššia hodnota zdravotného ukazovateľa znamená horší zdravotný stav. Spojitú premennú, vek od 15 do 50 rokov, chceme kategorizovať, teda rozumne rozdeliť do skupín alebo kategórií. Jednou z možností je dichotomizovať túto premennú podľa priemeru. Priemerný vek je 32 rokov, takže môžeme rozdeliť respondentov na skupinu do 32 rokov a skupinu 33 rokov a viac. Pri pohľade na obrázok vidíme, že respondenti do 32 rokov nadobúdajú hodnoty zdravia od 1 do 4 a tí starší od 3 do 10 a dá sa predpokladať, že rozdiely v zdraví medzi týmito dvoma skupinami nájdeme. Keď však sledujeme vzťah medzi vekom a zdravím, zdá sa, že to nie je úplný obraz.

Skúsme to s kvartilmi (označené červenými šípkami). Rozdelíme respondentov na dolný kvartil - do 21 rokov, strednú skupinu - 22 až 44 rokov a horný kvartil - 45 a viac rokov. Respondenti do 21 rokov nadobúdajú len nízke hodnoty medzi 1 a 2. Respondenti vo veku 22 až 44 rokov nadobúdajú hodnoty medzi 1 až 9 a respondenti, ktorí mali 45 a viac rokov, hodnoty 8 až 10. Je tam jasná postupnosť a rozdiely v zdraví by sme pravdepodobne potvrdili.

Pri pohľade na strednú skupinu vidíme, že nie je úplne homogénna. Znovu sa zdá, že obraz o vzťahu medzi vekom a zdravím nie je úplný. V strednej skupine sú akoby dve podskupiny a aj tie hranice horného a dolného kvartilu by mohli byť trochu lepšie ohraničené. Teraz do hry vstupuje intuícia, skúsenosť z praxe, keď vám niekto, kto týchto ľudí dennodenne ošetruje povie: "Vieš, okolo veku 30 rokov sa to u tých pacientov láme". A vy začnete v dátach hľadať, či je tam skutočne nejaký zlom a skúsíte to overovať.

Prvý zlom môže byť vo veku 21 až 25 rokov. Nemáme žiadneho respondenta nad 25 rokov, ktorý by sa dostal pod hodnotu 2. Ďalší zlom je po 35. roku v našom súbore, lebo nemáme v súbore žiadneho respondenta staršieho ako 35 rokov, ktorý by mal hodnotu pod 5. A posledný zlom, ktorý tam môžeme vidieť je okolo 43 až 44 rokov, kde už nemáme respondentov s hodnotou zdravia pod 7. Takéto rozdelenie lepšie vystihuje vzťah medzi vekom a zdravím.

Tieto príklady ilustrujú, ako možno upravovať údaje na štatistické premenné pomocou **kategorizácie**.



Obrázok číslo 4: Výstup zo štatistického softvéru.

Ďalším spôsobom úpravy údajov na štatistické premenné je **klastrovanie**, teda zoskupovanie premenných. Len zriedka dokážeme expozíciu alebo dôsledok opísať jedinou premennou. Oveľa častejšie do hry vstupuje viacero faktorov a je dôležité zohľadniť ich čo najviac. Predstavte si, že chcete overiť súvislosť medzi rizikovým správaním a zdravím. Niektorí respondenti nezdravoedia, ale nefajčia a nepijú. Iní fajčia a pijú, ale športujú a zdravoedia. Zdravie je výsledkom všetkých týchto aspektov správania naraz a nie jednotlivo. Ak by sme to skúmali jednotlivo, tak by sa nám medzi konzumentmi

alkoholu objavili športovci aj nešportovci, medzi fajčiarmi zdravo aj nezdravo sa stravujúci a účinky týchto druhov správania na zdravie sú opačné, takže by sme v tom mali zmätok. V skutočnosti sa ale tieto typy správania na seba viažu, a teda by mohlo byť možné identifikovať typy respondentov, u ktorých sa rôzne druhy zdravého alebo nezdravého správania zhlukujú.

Holandská skupina autorov van Nieuwenhuijzen a kol. (2009) sa o to pokúsila v štúdiu, ktorá skúmala reprezentatívny výber holandskej populácie vo veku 12 až 40 rokov v roku 2005/2006, s využitím internetového dotazníka a rozhovorov. Overovali zhlukovanie rizikového správania, ako je konzumácia alkoholu, tabaku, ilegálnych drog, rizikové sexuálne správanie, nedostatok fyzickej aktivity, vynechávanie raňajok, konzumácia ovocia a zeleniny, spánkový režim, správanie v doprave, delikventné a agresívne správanie. Skúmali to zvlášť v skupine pubescentov vo veku 12 až 15 rokov (obrázok číslo 5), adolescentov vo veku 16 až 18 rokov (obrázok číslo 6), mladých dospelých vo veku 19 až 24 rokov (obrázok číslo 7) a dospelých vo veku 25 až 40 rokov (obrázok číslo 7).

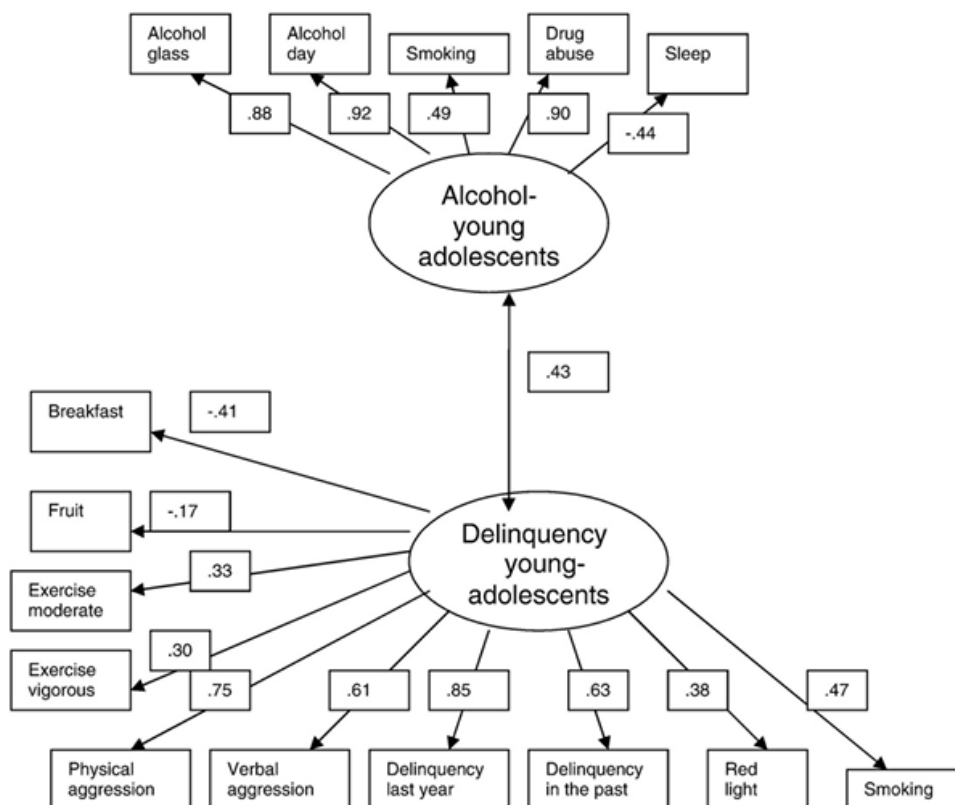


Fig. 3. Factor structure and loadings for 12–15 year old boys and girls ($N = 504$), Model 7: Risk Behavior Survey, The Netherlands, 2005–2006 ($N = 4468$).

Obrázok číslo 5: Zhlukovanie správania u pubescentov vo veku 12-15 rokov (prevzaté z van Nieuwenhuijzen M, Junger M, Klein Velderman M, Wiefferink KH, Paulussen T.W.G.M., Hox J, Reijneveld SA: Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Preventive Medicine* (2009), 48:572–578.)

V skupine pubescentov autori potvrdili širší zhluk správania porušujúceho normy (delinquency) a zhluk, ktorý obsahoval konzumáciu alkoholu, tabaku a ilegálnych drog, respektíve experimentovanie s konzumáciou týchto substancií (obrázok číslo 5). V skupine adolescentov potvrdili autori tejto štúdie tri zhluky (obrázok číslo 6). Zdravie podporujúce správanie sa vyčlenilo ako samostatný zhluk. Patrí tú

pravidelné raňajkovanie, konzumácia ovocia a zeleniny, ľahká alebo stredne náročná fyzická aktivita. Konzumácia alkoholu sa ocitla v zhluku spolu s rizikovým sexuálnym správaním, nebezpečným dopravným správaním a absenciou pravidelnej fyzickej aktivity. Zhluk normy porušujúceho správania, teda fyzicky a verbálne agresívneho správania, delikvencie a užívania drog zase prekvapujúco obsahoval i vynechávanie raňajok, spánkový režim a fajčenie.

V dospelosti dochádza ku kryštalizácii správania, a už nie je rozdiel v zhlukoch medzi 19 až 24 ročnými a 25 až 40 ročnými (obrázok číslo 7). Zdravý životný štýl a zdravie podporujúce formy správania tvorili jeden zhluk, do ktorého sa prekvapujúco zaradilo i nefajčenie. Zhluk delikventného správania sa vykryštalizoval a obsahoval iba formy správania porušujúceho normy. Konzumácia alkoholu vytvorila samostatný zhluk, do ktorého sa zaradilo i rizikové sexuálne správanie.

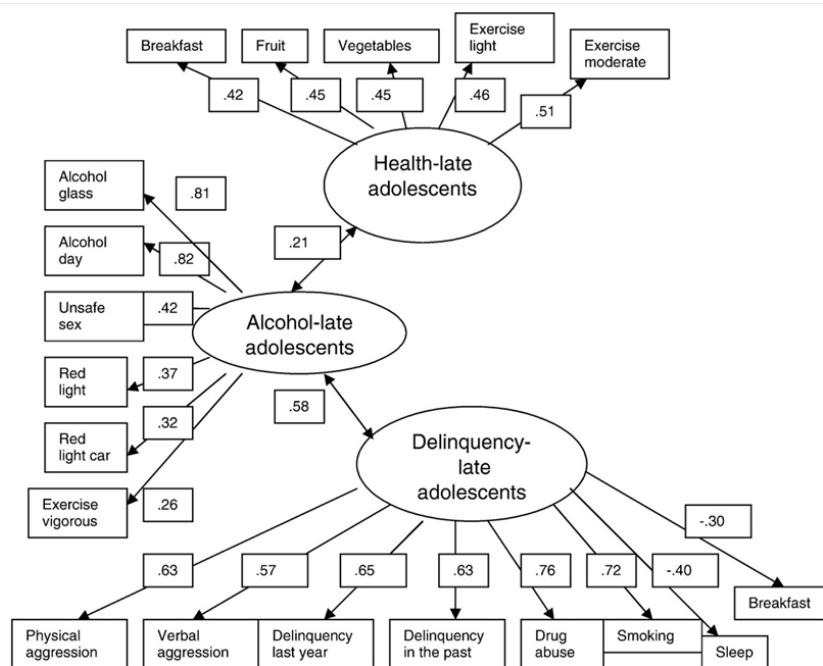


Fig. 2. Factor structure and loadings for 16–18 year old boys and girls ($N = 394$), Model 6: Risk Behavior Survey, The Netherlands, 2005–2006 ($N = 4468$).

Obrázok číslo 6: Zhlukovanie správania u adolescentov vo veku 16-18 rokov (prevzaté z *van Nieuwenhuijzen M, Junger M, Klein Velderman M, Wiefferink KH, Paulussen T.W.G.M., Hox J, Reijneveld SA: Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. Preventive Medicine (2009), 48:572–578.*)

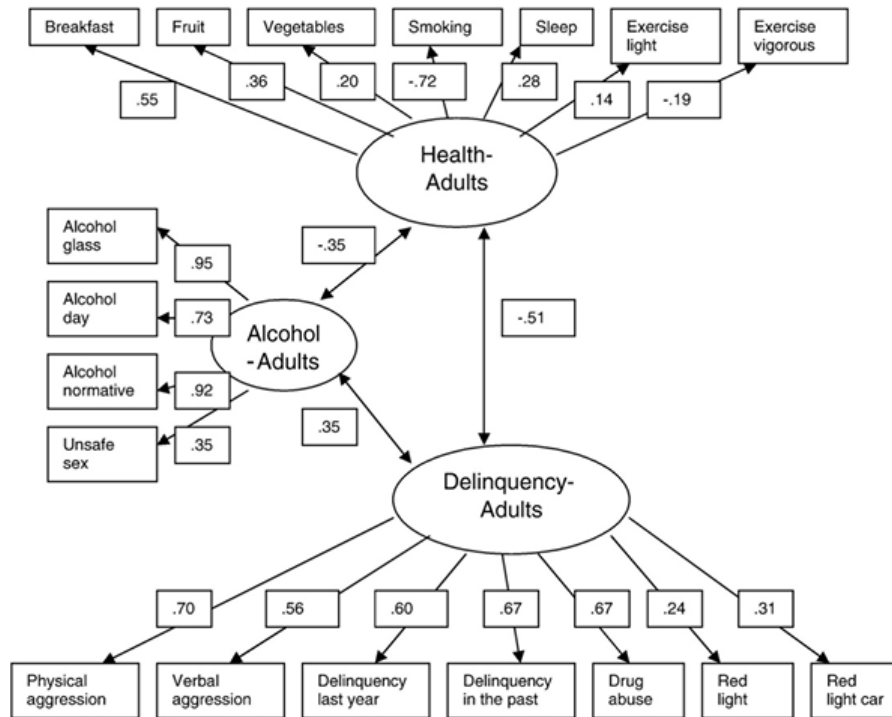


Fig. 1. Factor structure and loadings for 19–40 year old men and women ($N = 3497$), Model 3: Risk Behavior Survey, The Netherlands, 2005–2006 ($N = 4468$).

Obrázok číslo 7: Zhlukovanie správania v skupine dospelých vo veku 19-40 rokov (prevzaté z *van Nieuwenhuijzen M, Junger M, Klein Velderman M, Wiefferink KH, Paulussen T.W.G.M., Hox J, Reijneveld SA: Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. Preventive Medicine (2009), 48:572–578.*)

Iným spôsobom, ako zredukovať väčší počet premenných na menší počet premenných, je **faktorová analýza**. V tomto konkrétnom prípade (obrázok číslo 8) dvanásť otázok o psychickej pohode nahradila faktorová analýza dvoma spojenými premennými, tzv. faktorovými skóre. Prvú premennú, ktorá vyjadruje smútok a úzkosť, najviac ovplyvnilo prvých sedem otázok. Druhá premenná, na ktorú viac prispeli zvyšné otázky, vyjadruje skôr sociálnu nespôsobilosť. Faktorová analýza využíva zložitejšie matematické a štatistické postupy a v prípade, že potrebujeme zredukovať väčší počet spojených premenných je tu takáto možnosť. Faktorové skóre nie je súčtom položiek, sú to dve úplne odlišné premenné. Spodná tabuľka na obrázku číslo 8 popisuje faktorové skóre 1, ktoré nadobúda hodnoty od -1,98 do 3,37, pričom priemer je 0 a smerodajná odchýlka je 1.

Rotated Component Matrix^a

	Component	
	1	2
Mal si v poslednej dobe pocit, že si nešťastný, alebo smutný?	,771	
Mal si v poslednej dobe pocit, že nedokážeš prekonať ťažkosti?	,719	
Mal si v posl. čase starosti, kvôli ktorým si nemohol pokojne spať?	,705	
Strácal si v poslednej dobe sebadôveru?	,685	
Mal si v poslednej dobe pocit, že si pod neustálym tlakom?	,682	
Rozmýšľal si v poslednej dobe o sebe ako o bezcennom človeku?	,665	
Cítil si sa v poslednej dobe šťastný?	,486	,471
Mal si v poslednom čase radosť z toho, čo robíš?		,661
Mal si v poslednej dobe pocit, že to, čo robíš, má zmysel?		,631
Vedel si sa v poslednom čase samostatne rozhodovať?		,600
Vedel si sa v poslednej dobe postaviť zoči-voči svojim problémom?		,583
Vedel si sa v poslednej dobe sústrediť na veci, ktoré si robil?	,326	,375

	N	Minimum	Maximum	Mean	Std. Deviation
REGR factor score 1	3575	-1,98168	3,36931	,0000000	1,0000000
REGR factor score 2	3575	-2,98160	5,07941	,0000000	1,0000000
Valid N (listwise)	3575				

Obrázok číslo 8: Výstup zo štatistického softvéru.

2.3 Exponovanie skúmanej okolnosti

V tejto časti sa pokúsime vysvetliť, ako rozlíšiť, kto bol alebo nebol exponovaný skúmanej okolnosti či javu, a kto je alebo nie je prípad. Ak skúmame súvislosť medzi expozíciou a jej dôsledkom, potrebujeme u každého respondenta vedieť identifikovať, či bol alebo nebol exponovaný, prípadne ako veľmi.

Predstavte si usmievať sa chlapca, ktorý pózuje pred rozbitou bytovkou a haldami odpadkov – je podľa vás exponovaný chudobe? V akom veku a ako dlho tam musí žiť, aby to malo na neho nejaký vplyv? Pre niektoré dieťa môže byť aj jeden deň na takomto mieste celoživotnou traumou, iné to môže považovať za normálne a schované v bubline svojej rodiny žiť v normálne zariadenom byte a bude minimálne exponované tomu, čo sa deje vonku.

Čítanie rozprávok deťom je expozícia, ktorá môže výrazne stimulovať kognitívny aj emocionálny rozvoj detí a viesť k dramatickým pokrokom aj u detí so znevýhodneného prostredia. Stačí pritom tak málo – čítať im rozprávky. Program občianskeho združenia Cesta von, s názvom Omama, na Slovensku prináša celý program rozvíjajúcich intervencií do rómskych osád a už dnes vieme povedať, že aj v tých najhorších podmienkach dokážu v predškolskom veku urobiť za krátky čas niekoľkoročný rozdiel v kognitívnom vývine. Deti, ktoré by vo veku 6 rokov nezvládli zaškolenie, to zvládnu bez problémov. Najviac sú exponované deti, s ktorými sa priamo pracuje, a tých je len niekoľko desiatok. Nepriamo sú ale exponovaní aj ich súrodenci, lebo sú často prítomní na lekciách a pri spoločných hrách veľa zdieľajú.

Skôr než niekoho zaradíme do skupiny exponovaných, mali by sme sa zamyslieť nad tým, či je alebo bol skutočne vystavený daným podmienkam.

PRAKTICKÁ ÚLOHA 1:

Vráťme sa k príkladu s ľuďmi žijúcimi na ulici v Košiciach a pridajme k nim porovnanie s Trebišovom. V Košiciach bolo registrovaných 800 ľudí žijúcich na ulici, z ktorých štyria v zime umrzli, zatiaľ čo v Trebišove mali registrovaných 50 takýchto ľudí a reportovali iba jedno takéto úmrtie. V ktorom meste bola situácia horšia?

RIEŠENIE PRAKTICKEJ ÚLOHY 1:

Košice: $4/800 = 5/1000$

Trebišov $1/50 = 20/1000$

Výskyt, teda prevalenciu prípadov možno vypočítať tak, že vydáme počet existujúcich prípadov počtom všetkých prípadov. V Košiciach to je teda 4 delené 800, čo je 0,005, teda 5 úmrtí na tisíc ľudí žijúcich na ulici. V Trebišove je to 1 delené 50, čo je 0,020, teda 20 úmrtí na tisíc ľudí žijúcich na ulici. Z výpočtu vyplýva, že v Trebišove sú na tom oveľa horšie.

Je to ale skutočne tak?

V Košiciach môže veľmi dobre fungovať organizácia, ktorá s týmito ľuďmi pracuje, a tak ich dokáže viac podchytiť, viac z nich je registrovaných. V Trebišove takáto organizácia nemusí byť a registrovaných je oveľa menej z nich. Expozícia v Košiciach a Trebišove je odhadovaná odlišným spôsobom, máme neporovnateľné menovatele v rovnici.

Problém môže byť aj s odhadovaním dôsledku, teda čitateľa v rovnici. V Košiciach vďaka dobrej spolupráci spomínanej organizácie s lekármi, môžu lekári zapisovať do úmrtného listu príčinu úmrtia podchladením, kým v Trebišove ju môžu zaradzovať raz ako podchladenie a raz ako kardiovaskulárne zlyhanie.

Ak zbierame údaje od viacerých poskytovateľov alebo z viacerých lokalít či krajín, je veľmi dôležité premyslieť si a prediskutovať, kto bude alebo nebude označovaný za exponovaného, a kto bude alebo nebude považovaný za prípad či neprípad.

2.4 Popísanie frekvencie prípadov

Na zhodnotenie frekvencie prípadov používame prevalenciu a incidenciu.

Prevalencia (výskyt) je počet existujúcich prípadov vydelený počtom všetkých členov populácie v danom čase, ktorých sa to týka. Konkrétne príklady, ktoré ilustrujú prevalenciu sú nasledovné: Zo zistení štúdie *Health Behaviour in School-aged Children* na Slovensku z roku 2018 vyplýva, že tretina detí vo veku 11 rokov trpí nadváhou a obezitou, pričom chlapci majú viac ako 25% a dievčatá viac ako 30% telesného tuku podľa odhadov zloženia tela. V roku 2010 predstavovala prevalencia cukrovky 2. typu v populácii pacientov starších ako 18 rokov na úrovni 5,48 %.

Incidencia je počet nových prípadov vydelený počtom všetkých členov populácie, ktorých sa to týka. Konkrétne príklady, ktoré ilustrujú incidenciu sú nasledovné: V roku 2018 bolo prvýkrát diagnostikované ochorenie diabetes mellitus 2. typu 18 177 pacientom. Miera incidencie predstavovala 333,5 osôb na 100 000 obyvateľov. V 26. týždni v roku 2020 bolo hlásených 6 440 akútnych respiračných ochorení teda incidencia 319,8 prípadov na 100 000 osôb.

Incidenca a prevalencia daného ochorenia súvisia. Prevalenciu si môžeme predstaviť ako hmotou v nádobe, do ktorej neustále pritekajú nové prípady tvoriace incidenciu. Zároveň z nej neustále unikajú uzdravení a zomrelí.

Na obrázku číslo 9 je znázornený príklad na prevalenciu a incidenciu. Vo Veľkej Británii v roku 2008 trpelo onkologickým ochorením 247 667 ľudí, kým v Belgicku 47 948 ľudí. V daný rok žilo vo Veľkej Británii 60 miliónov ľudí a v Belgicku 10 miliónov ľudí. Po vydelení čitateľa menovateľom a prevode na 100 000 ľudí vieme povedať, že onkologickým ochorením v roku 2008 trpelo vo Veľkej Británii 413 a v Belgicku 479 ľudí na 100 000 obyvateľov. Prevalencia je teda vyššia v Belgicku, napriek tomu, že prípadov je viac vo Veľkej Británii.

Čo sa týka incidencie (spodná časť obrázka číslo 9), v roku 2003 bolo nahlásených 43 000 nových prípadov rakoviny prsníka u žien. Británia síce má 60 miliónov ľudí, ale len 20 miliónov dospelých žien, ktoré sú pre tento výpočet incidencie relevantné. Incidencia teda je 215 prípadov na 100 000 žien.

Prevalencia

	Počet prípadov čitateľ	Počet obyvateľov menovateľ	čitateľ/ menovateľ	na 100 000 obyvateľov
UK	247667	60 mil	0.00413	413
Belgicko	47948	10 mil	0.00479	479

Incidenca

Nové prípady dg. rakoviny prsníka u žien v roku 2003 v UK čitateľ	Počet žien menovateľ	čitateľ/ menovateľ	na 100 000 obyvateľov
43000	20 mil	.00215	215

Obrázok číslo 9: Príklad na prevalenciu a incidenciu.

2.5 Popísanie vplyvu expozície na výskyt prípadov, interpretácia

V tejto časti si ukážeme, ako popísať vplyv expozície na výskyt prípadov, vypočítať *relatívne riziko a pomer šancí*, a ako interpretovať intervaly spoľahlivosti.

Na obrázku číslo 10 je znázornený výpočet relatívneho rizika a pomeru šancí.

Relatívne riziko (RR), pomer šancí (OR)

		prípade		RR	OR
		áno	nie		
expozícia	áno	a	b	$a/(a+b)$	a/b
	nie	c	d	$c/(c+d)$	c/d

$$RR = \frac{r_1}{r_2} = \frac{a / (a+b)}{c / (c+d)}$$

$$OR = \frac{O_1}{O_2} = \frac{a / b}{c / d}$$

Obrázok číslo 10: Výpočet relatívneho rizika a pomeru šancí.

Tabuľka na obrázku číslo 11 popisuje, koľko zo sledovaných pacientov trpelo (bolo exponovaných) multiorgánovému zlyhaniu a koľko z nich zomrelo. V súbore bolo 26 pacientov exponovaných multiorgánovému zlyhaniu, z ktorých 22 zomrelo a 22 pacientov, ktorí neboli exponovaní multiorgánovému zlyhaniu, z ktorých zomrelo 5.

V prvom kroku vypočítame prevalenciu u pacientov exponovaných a neexponovaných multiorgánovému zlyhaniu.

V druhom kroku tieto prevalencie vydělíme. Prevalenciu úmrtí u pacientov exponovaných multiorgánovému zlyhaniu vypočítame tak, že zomrelých (22) vydělíme počtom všetkých exponovaných pacientov (22+4, čo je 26). 22 delené 26 je 0,846, výskyt (prevalencia) úmrtia u exponovaných je 84,6%. Prevalenciu úmrtí u pacientov neexponovaných multiorgánovému zlyhaniu vypočítame tak, že zomrelých (5) vydělíme počtom všetkých exponovaných pacientov (5+17, čo je 22). 5 delené 22 je 0,227, výskyt (prevalencia) úmrtia u exponovaných je 22,7 %.

Relatívne riziko (RR) exponovaných vypočítame tak, že prevalenciu úmrti u exponovaných vydělíme prevalenciou úmrtí u neexponovaných. To je 0,846 delené 0,227. Relatívne riziko úmrtia u exponovaných pacientov je 3,7-krát vyššie než u neexponovaných pacientov. Zjednodušene by sme mohli povedať, že viac než 80% pacientov s multiorgánovým zlyhaním zomrelo, zatiaľ čo z pacientov bez tohto zlyhania zomrelo niečo cez 20%. Je to takmer 4-krát vyššie riziko úmrtia v jednej skupine oproti inej.

Výpočet **pomeru šancí (OR)** budú poznať tí, čo stávkujú a riešia, aké šance na výhru má jedno mužstvo oproti druhému. V našom prípade jedno mužstvo má multiorgánové zlyhanie a druhé nie. V prvom kroku vypočítame šance každého mužstva, ako pomer počtu výhier k počtu prehíer. V druhom kroku tieto šance vydělíme. V našom prípade to vyšlo tak, že výhrami sú úmrtia. Takže mužstvo z multiorgánovým zlyhaním má jednoznačne navrch, lebo pomer 22 ku 4 je určite väčší ako 5 ku 17. Pomer šancí je 18,7 alebo inými slovami tí, ktorí boli exponovaní multiorgánovému zlyhaniu mali skoro 19-krát vyššiu pravdepodobnosť úmrtia v porovnaní s pacientami, ktorí neboli exponovaní multiorgánovému zlyhaniu (obrázok číslo 11).

Relatívne riziko (RR), pomer šancí (OR)

		úmrtie		RR	OR
		áno	nie		
Multiorgán. zlyhanie	áno	22	4	22/(22+4)	22/4
	nie	5	17	5/(5+17)	5/17

$$RR = (22/26) / (5/22) = 0,846 / 0,227 = \mathbf{3,72}$$

$$OR = (22/4) / (5/17) = 5,5 / 0,294 = \mathbf{18,7}$$

Obrázok číslo 11: Výpočet relatívneho rizika a pomeru šancí.

Relatívne riziko, rovnako ako pomer šancí, popisuje vplyv expozície na dôsledok, respektíve súvislosť medzi expozíciou a dôsledkom. V tomto konkrétnom prípade je medzi nimi veľký rozdiel, ale vo väčšine prípadov to môžu byť dosť podobné čísla. Vyjadrujú ale iný parameter vzťahu medzi expozíciou a dôsledkom.

To, akým spôsobom popíšeme súvislosť expozície a dôsledku záleží aj od toho, ktorú skupinu si vyberieme za referenčnú, a podľa toho bude vyzeráť aj **interpretácia** vašich **výsledkov**. V nasledujúcom prípade (tabuľka číslo 1) sme sledovali správanie matiek a pubescentných dcér tesne po príchode domov. 30 zo 130 sledovaných matiek vošlo do kuchyne a hneď sa pustilo do umývania riadu. No iba 20 zo sledovaných 220 pubescentných dcér sa správalo podobne.

Ak budú referenčnou skupinou dcéry, ich pomer šancí bude menovateľ a budeme popisovať šance matiek oproti šanciam dcér. Teda matky mali 3-krát vyššiu šancu oproti dcéram, že umyjú riad v dreze hneď po príchode domov.

$$(30/100)/(20/100)=0,3/0,1=3$$

Matky mali 3-krát vyššiu šancu, že umyjú riad v dreze po príchode domov (OR=3,0).

Ak budú referenčnou skupinou matky, ich pomer šancí bude menovateľ a budeme popisovať šance dcér oproti šanciam matiek. Teda dcéry mali 3-krát nižšiu šancu oproti matkám, že umyjú riad v dreze hneď po príchode domov. Mohli by sme aj povedať, že majú o 70% nižšiu šancu.

$$(20/100)/(30/100)=0,1/0,3=1/3=0,33$$

Dcéry mali 3-krát nižšiu šancu, že umyjú riad v dreze po príchode domov (OR=0,33).

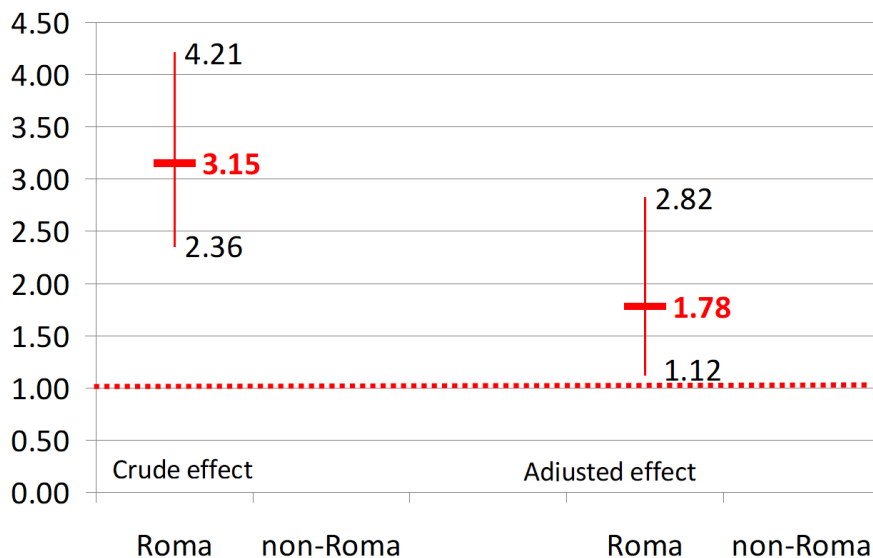
Ak chceme referenčné skupiny vymeniť, stačí vydeliť jednotku pomerom šancí. V tomto prípade, ak by sme chceli vedieť aké šance majú dcéry a vieme, že pomer šancí matiek je 3, pomer šancí pri vymenenej referenčnej skupine by bolo 1/3, čo je 0,33, teda 3-krát menej.

	umyli riad	neumyli riad
matky	30	100
dcéry	20	200

Tabuľka číslo 1: Príklad na výpočet pomeru šancí.

Poslednou témou, ktorej sa budeme v rámci tejto kapitoly venovať sú intervaly spoľahlivosti. Náš výskumný súbor je len náhodne vybranou skupinou guľičiek z vrecúška. Pomer šancí, ktorý sme vypočítali, sa týka len tohto výskumného súboru. My však chceme tie zistenia vzťahovať na celú populáciu. V tomto bode do hry vstupuje štatistika, odhadovanie. Vďaka intervalu spoľahlivosti vieme na 99% alebo 95% (v závislosti od zvolenej hladiny významnosti) povedať, že ak by sme z vrecúška vytiahli inú skupinu guľičiek a vypočítali pomer šancí, neboli by hodnoty nižšie ani vyššie, ako sú hranice intervalu spoľahlivosti.

Na obrázku číslo 12 je znázornený pomer šancí detí z rómskych osád v porovnaní s deťmi z majority pri reportovaní zdravotného problému. Jeho hodnota je 3,15, teda v našom výskumnom súbore mali rómske deti 3-krát vyššiu šancu oproti deťom z majority reportovať zdravotný problém. Ak by sme tento výskum zopakovali v skúmanej populácii, pomer šancí by bol prinajmenšom 2,36 a nanajvyš 4,21, čo by v každom prípade bolo viac ako 1. Znamenalo by to, že sa šance reportovať zdravotný problém detí z rómskej osady nelíšia od šancí detí z majority.



Obrázok číslo 12: Výstup zo štatistického softvéru.

Intervaly spoľahlivosti nám hovoria nie len o tom, či je naše zistenie štatisticky významné alebo len náhodné, ale hovoria tiež o spoľahlivosti tohto výsledku. Veľmi široké intervaly spoľahlivosti naznačujú, že pri opakovaní by sme mohli dospieť k veľmi rozdielnym výsledkom. Niektoré by boli v prospech detí z majority, iné v prospech detí z rómskej osady. Možno máme malý súbor. Možno je v hre ešte nejaký iný významný hráč, ktorého potrebujeme dať do modelu. To, čo chceme vidieť, je čo najmenší interval spoľahlivosti.

OTÁZKY

1. Ktorý z nasledujúcich výrokov popisuje prevalenciu a ktorý incidenciu?

	prevalencia	incidencia
V roku 2011 bolo v Košickom kraji hlásených 40 zhubných nádorov hrubého čreva na 100 000 mužov.		
V roku 2011 bolo v okrese Stropkov hlásených 132 zhubných nádorov prsníka na 100 000 žien.		
V Slovenskej republike bolo v roku 2012/2013 spolu hlásených 1 698 326 akútnych respiračných ochorení (ARO), čo predstavuje chorobnosť 54 631,2 na 100 000 osôb v starostlivosti hlásiacich lekárov.		
Každé tretie dieťa vo veku 11 rokov trpí nadváhou alebo obezitou.		
Arteriálnou hypertenziou trpelo v roku 2019 45% dospelaj populácie na Slovensku.		
V roku 2012 bolo na Slovensku hlásené, že chronickým ochorením trpí každý piaty človek.		

2. V sledovanom výbere je 200 chlapcov a 800 dievčat, pričom 40 chlapcov a 80 dievčat trpí nadváhou alebo obezitou. Je prevalencia nadváhy a obezity vyššia u chlapcov alebo u dievčat?

- A) chlapci
- B) dievčatá

3. Aká bola incidenciu cukrovky (na 100 000) u detí do 14 rokov na Slovensku v roku 2006, ak v tomto roku zaznamenali 176 nových prípadov cukrovky u detí, Slovensko malo 6 600 000 obyvateľov, z toho 1 000 000 vo veku do 14 rokov?

- A) 26,67 / 100 000
- B) 17,60/ 100 000
- C) 30,00/ 100 000
- D) 50/100 000

4. Aká bola incidenciu úmrtí na podchladenie v meste Košice za rok 2006 až 2008 v skupine ľudí žijúcich na ulici, ktorých bolo registrovaných 100. V roku 2006 bolo registrovaných 6 takýchto úmrtí a 14 sledovaných nebolo možné skontaktovať. V roku 2007 boli registrované 4 takéto úmrtia a 16 sledovaných nebolo možné skontaktovať. V roku 2008 boli zaregistrované 2 takéto úmrtia v sledovanej skupine.

- A) 120,00/1000
- B) 100,00/1000
- C) 420,00/1000
- D) 60/1000
- E) 50/1000

5. Aké relatívne riziko majú fajčiari oproti nefajčiarom, že budú trpieť astmou. V súbore je 100 fajčiarov a 100 nefajčiarov, pričom 40 fajčiarov a 20 nefajčiarov trpelo astmou.

- A) 2,68
- B) 2,00
- C) 0,40

6. Aké šance majú fajčiari oproti nefajčiarom, že budú trpieť astmou. V súbore je 100 fajčiarov a 100 nefajčiarov, pričom 40 fajčiarov a 20 nefajčiarov trpelo astmou.

- A) 2,68
- B) 2,00
- C) 0,40

7. V súbore sa zisťovalo aký pomer šancí majú fajčiari oproti nefajčiarom umrieť na rakovinu pľúc. Pomer šancí bol 1,23. Sú nasledujúce tvrdenia pravdivé?

	je to pravda	nie je to pravda
Šanca fajčiara, že umrie na rakovinu pľúc, je o 23% vyššia ako u nefajčiara.		
Šanca nefajčiara, že umrie na rakovinu pľúc, je o 23% vyššia ako u fajčiara.		
Šanca fajčiara, že umrie na rakovinu pľúc, je 1,2 krát vyššia ako u nefajčiara.		
Šanca nefajčiara, že umrie na rakovinu pľúc, je 1,2 krát vyššia ako u fajčiara.		

3 DIZAJN ŠTÚDIE

V tejto kapitole si predstavíme rôzne typy dizajnov výskumných štúdií, jednu z klasifikácii dizajnov štúdie, spoznáme výhody a nevýhody jednotlivých dizajnov štúdie a predstavíme si užitočné iniciatívy či odporúčania, ktoré nám môžu pomôcť overiť, či sme pri každom dizajne na nič nezabudli a všetko opísali v protokole.

Najprv si objasníme, čo je **dizajn štúdie**. Cieľom epidemiologického výskumu je získať **validné a presné informácie o príčinách, prevencii a liečbe ochorení**. Každý druh epidemiologického dizajnu jednoducho reprezentuje **odlišný spôsob získavania informácií** (Aschengrau, Seage, 2008). Každý zo spôsobov má svoje pravidlá, výhody a nevýhody. Ich zaradenie do typov nám pomáha rýchlejšie sa orientovať v metódach a interpretovať zistenia.

3.1 Klasifikácia dizajnov štúdií

Klasifikácie sú ako triediče na príbor v príborovom šuflíku alebo krabičky na skrutky. Je ich viacero druhov a podstatné je, aby sa v nich vyznal ten, kto s tým príborom alebo so skrutkami pracuje. V literatúre sa stretnete s rôznymi klasifikáciami dizajnov. Pre účely týchto učebných textov sme vybrali jednu z nich, ktorú budeme používať pri predstavovaní jednotlivých dizajnov.

Výskumné dizajny môžeme podľa Parikh a Concato (2006) rozdeliť na **popisné** (deskriptívne), **analytické** a **integratívne**.

Popisné dizajny majú ambíciu len popísať javy, prípady, série prípadov, incidenciu alebo prevalenciu nejakého javu. Patria tu prípadové štúdie (kazuistiky), série prípadových štúdií, incidenčné a prevalenčné štúdie.

Analytické dizajny majú ambíciu popísať súvislosti medzi expozíciou a dôsledkom. Realizujú to buď prostredníctvom pozorovania toho, čo by sa aj tak stalo (*observačné dizajny*), alebo do deja vstupujú a manipulujú s expozíciou (*experimentálne dizajny*). Observačné dizajny sa ďalej líšia podľa toho, či údaje o expozícii a dôsledku zbierajú v jednom čase, či počkajú na dôsledok alebo po expozícii pátrajú v minulosti.

Observačné štúdie delíme na prierezové, kohortové a prípadové kontrolné štúdie (tzv. case-control). K experimentálnym dizajnom zaraďujeme randomizované kontrolované štúdie a intervenčné štúdie.

Okrem toho existujú **integratívne dizajny**, ktoré sa snažia extrahovať zistenia integrovaním viacerých štúdií. Tu patria napr. metaanalýzy. V Európe máme napríklad desiatky štúdií o zlyhaní obličiek u darcov obličiek, ale každá z nich je realizovaná na malom súbore. Mohli by sme chcieť integrovať tieto zistenia, akoby ich spojili, a tým sa vyjadriť k danému javu s väčšou silou. To nám umožňujú metaanalýzy.

Podľa časového plánu niektoré klasifikácie rozlišujú **retrospektívne a prospektívne štúdie**. Príkladom prospektívnej štúdie je situácia, keď sledujeme skupinu tehotných žien, počkáme si na pôrod ich detí a následne tieto deti sledujeme až do dospelosti. Cieľom môže byť napríklad zistiť, či pôrodná váha a výška súvisí s reálnou telesnou váhou a výškou v dospelosti. Ak to isté urobíme naopak, vyberieme si skupinu dospelých a dohľadáme údaje o ich pôrodnej výške a váhe, to je príklad retrospektívnej štúdie.

3.1.1 Popisné (deskriptívne) dizajny štúdií

Deskriptívne štúdie si kladú za cieľ popísať a nie analyzovať. Patria sem prípadové štúdie (kazuistiky), série prípadových štúdií, incidenčné a prevalenčné štúdie.

Prípadové štúdie, teda kazuistiky, popisujú prípad alebo série prípadov. V roku 2009 sa objavila prasacia chrípka. Prvé informácie, ktoré sme o tomto ochorení mali, boli prípadové štúdie, ktoré popisovali priebeh ochorenia u jedného alebo viacerých pacientov. Prekvapujúco postihovalo to skôr mladšie ako staršie ročníky a rizikovou skupinou sa zdali byť aj mladé tehotné ženy. Tie to prvé popisy prípadov a série prípadov boli jedinými zdrojmi informácii, o ktoré sme sa mohli oprieť, na základe ktorých sme mohli robiť hypotézy a pripraviť výskumné plány štúdií s komplikovanejšími dizajnami. Kontrolná skupina nebola k dispozícii, a tak **nebolo možné porovnávať exponovaných a neexponovaných** ani sa vyjadrovať ku kauzalite (teda potvrdiť, či je expozícia príčinou dôsledku).

Mohlo by sa stať, že prípadové štúdie by naznačovali vyššiu náchylnosť na ochorenie u futbalových fanúšikov alebo mladých matiek, pretože by ich v sérii prípadových štúdií bolo výrazne viac. V skutočnosti však mohli mať účastníci konkrétneho futbalového zápasu len smolu, že prvý exportovaný nakazený sa prišiel pozrieť na zápas, alebo prvá exportovaná nakazená bola lektorka na sérii kurzov v materských centrách.

Na druhej strane je to veľmi **nenáročný dizajn** – využíva to, čo máte k dispozícii, napr. pacientov, ktorých k vám priviezli a vy ich ošetrujete. Tento dizajn je ale veľmi užitočný v prípade **nových ochorení alebo zriedkavých komplikácií** a môže poskytnúť **indície** potrebné pre ďalší výskum.

Predstavme si napríklad sériu prípadov milovníkov piva, ktorí by trpeli častým močením. Ak by sme chceli overiť, či je príčinou práve pivo, museli by sme ich skúsiť porovnať s inou skupinou, napríklad milovníkmi vína alebo čistej vody, prípadne ich požiadať, aby namiesto piva v určených intervaloch pili rovnaké množstvo vody. To by už však bola analytická štúdia.

Medzi *výhody prípadových štúdií* patrí, že umožňujú popísať zriedkavé a novo sa vyskytujúce choroby či monitorovať *nezvyklé prípady*. Môžu taktiež poskytnúť potrebné *indície k ďalším štúdiám*. Medzi ich *nevýhody* patrí, že majú obmedzené implikácie vzhľadom ku *kauzalite* a môžu byť *mätúce* alebo môžu viesť *k falošným stopám*.

Príkladom prípadovej štúdie je kazuistika autorov Rosenberger a kol. (2005 a), znázornená na obrázku číslo 1. Popísali pacienta po transplantácii obličky, ktorý mal za sebou tri epizódy štikútania, ktoré ho veľmi obťažovali. Prekvapujúco, každá z týchto epizód mala úplne inú príčinu. Jednu z nich mali na svedomí steroidy a anestetiká, druhú vážne vredovité ochorenie pažeráka a tretiu bakteriálny zápal pľúc. Tento pacient mal teda riadnu smolu. Popis tohto prípadu však môže pomôcť iným lekárom pri diagnostike a liečbe takýchto pacientov, pri ktorých by im ani nenapadlo, že sa u jedného pacienta môžu zísť tri rôzne dôvody štikútania.

Incidenčné a prevalenčné štúdie popisujú **rozsah ochorenia** alebo **charakteristik ochorenia** v populácii. Sú dôležité pre **plánovanie** zdravotníckej starostlivosti a môžu byť i **východiskovým bodom** pre analytické štúdie. Obvykle sú prierezové (prevalencia), ale môžu byť aj longitudinálne (incidencia).

Príkladom prevalenčnej štúdie je výskum švédskych autorov Lötvald a kol. (2009), znázornený na obrázku číslo 2. Táto štúdia skúmala respiračné symptómy astmy v populácii vo veku 17 až 75 rokov v roku 1990 a 2008. Autori predpokladali, že výskyt astmy stúpa, ale ani po 8 rokoch nezaznamenali zmenu. Zmenil sa však výskyt respiračných symptómov.

Review > Transpl Infect Dis. 2005 Jun;7(2):86-8. doi: 10.1111/j.1399-3062.2005.00096.x.

A renal transplant patient with intractable hiccups and review of the literature

J Rosenberger¹, E Veseliny, L Bena, R Roland

Affiliations + expand

PMID: 16150097 DOI: 10.1111/j.1399-3062.2005.00096.x

Abstract

Intractable hiccups in transplanted patients may be caused by various medical conditions including infections. We report a case of a 44-year-old man who suffered from intractable hiccups after cadaveric kidney transplantation. We identified 3 different hiccup periods with different causes: 1) steroid and anesthetics use, 2) severe ulcerose herpetic and mycotic esophagitis, and 3) pleuropneumonia caused by nosocomial methicillin-resistant *Staphylococcus epidermidis* and pulmonary abscess requiring thoracic surgery.

PubMed Disclaimer

FULL TEXT LINKS



ACTIONS

Cite

Collections

SHARE



PAGE NAVIGATION

< Title & authors

Obrázok číslo 1: Príklad prípadovej štúdie.

Research | [Open access](#) | Published: 12 October 2009

West Sweden Asthma Study: Prevalence trends over the last 18 years argues no recent increase in asthma

Jan Lötvall , Linda Ekerljung, Erik P Rönmark, Göran Wennergren, Anders Lindén, Eva Rönmark, Kjell Torén & Bo Lundbäck

Respiratory Research 10, Article number: 94 (2009) | [Cite this article](#)

8898 Accesses | 133 Citations | [Metrics](#)

Abstract

Asthma prevalence has increased over the last fifty years, but the more recent changes have not been conclusively determined. Studies in children indicate that a plateau in the prevalence of asthma may have been reached, but this has not yet been confirmed in adults.

Epidemiological studies have suggested that the prevalence of asthma in adults is approximately 7-10% in different parts of the western world.

We have now performed a large-scale epidemiological evaluation of the prevalence of asthma and respiratory symptoms in adults between the ages of 16-75 in West Sweden. Thirty thousand randomly chosen individuals were sent a detailed questionnaire focusing on asthma and respiratory symptoms, as well possible risk factors. Sixty-two percent of the contacted individuals responded to the questionnaire. Asthma prevalence, defined as asthma diagnosed by a physician, was 8.3%. Moreover, the prevalence of respiratory symptoms was lower compared to previous studies. The most common respiratory symptom was any wheeze (16.6%) followed by sputum production (13.3%). In comparison with studies performed 18 years ago, the prevalence of asthma has not increased, and the prevalence of most respiratory symptoms has decreased. Therefore, our data argues that the continued increase in asthma prevalence that has been observed over the last half century is over.

Obrázok číslo 2: Príklad prevalenčnej štúdie.

PRAKTICKÁ ÚLOHA 1:

Na obrázku číslo 3 je zobrazený výskyt jednotlivých respiračných symptómov v roku 1990 (svetlé stĺpce) a v roku 2008 (tmavé stĺpce) zo štúdie už spomínaných švédskych autorov Lötvall a kol. (2009).

Skúste si obrázok dobre pozrieť a na hárok papiera napísať odpovede na nasledujúce otázky:

Ktorý symptóm sa vyskytoval najčastejšie v roku 1990 a ktorý v roku 2008?

Ktorý symptóm sa vyskytoval častejšie v roku 1990 ako v roku 2008?

Ktorý symptóm sa vyskytoval častejšie v roku 2008 ako v roku 1990?

Prečo mohlo dôjsť k zmene výskytu symptómov? Konkrétne, prečo sa v roku 2008 pacienti menej často sťažovali na to, že sa budia na záchvat kašľania. A prečo sa naopak v roku 2008 častejšie sťažovali na symptómy sennej nádchy/ alergickej rinitídy?

Respiratory Research 2009, 10:94

<http://respiratory-research.com/content/10/1/94>

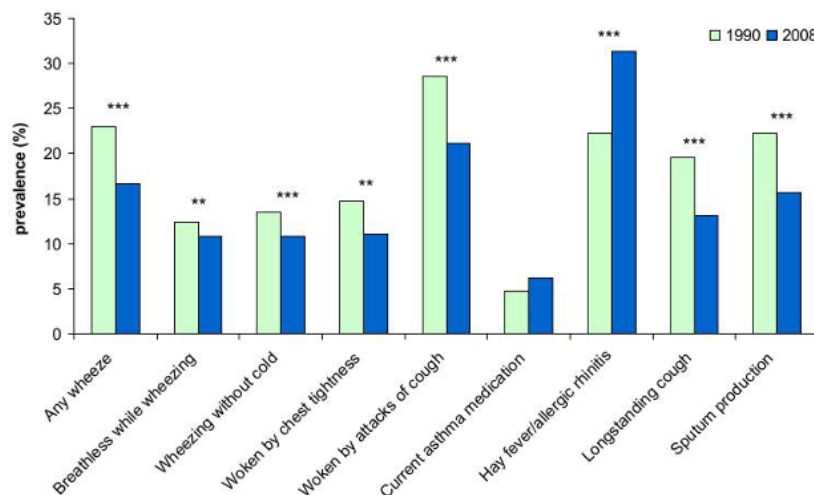


Figure 3
Comparison of the prevalence of respiratory symptoms using identical questions in the ECRHS Study (1990) and in the current study (2008) among 20-44 year old subjects living in the area of Hisingen, Gothenburg (* p < 0.05, ** p < 0.01, ***p < 0.001).

Obrázok číslo 3: Porovnanie výskytu respiračných ochorení v roku 1990 a 2008.

RIEŠENIE PRAKTICKEJ ÚLOHY 1:

Pacienti sa v roku 1990 najčastejšie sťažovali, že sa budia na záchvat kašľania.

V roku 2008 sa pacienti najčastejšie sťažovali na symptómy sennej nádchy/ alergickej rinitídy.

Z celého súboru sledovaných respiračných ťažkostí sú len dva také, na ktoré sa v roku 2008 pacienti sťažovali častejšie ako v roku 1990, pričom iba v jednom je rozdiel štatisticky významný, keďže hviezdičky nad stĺpcami označujú štatistickú významnosť. Pacienti sa v roku 2008 častejšie sťažovali na symptómy sennej nádchy ako pacienti z roku 1990.

Jedným z dôvodov, nad ktorými ste mohli uvažovať pri hľadaní dôvodu, prečo došlo k zmene výskytu symptómov, by mohla byť zmena liečby. Za 8 rokov skutočne mohlo dôjsť k novým objavom, zavedeniu lepšej, účinnejšej liečby a teda pacienti trpia vybranými respiračnými symptómami menej často.

Táto prevalenčná štúdia nám teda ponúka indíciu pre hypotézu, že zavedenie nových liekov znižuje výskyt respiračných symptómov, ale na jej overenie by sme museli použiť prinajmenšom analytický dizajn štúdie.

Prečo ale pacienti trpia častejšie príznakmi sennej nádchy? Mohli by ste uvažovať o tom, že absencia vážnejších symptómov umožní prejavíť sa miernejším symptómom, alebo že sa zhoršuje znečistenie ovzdušia, alebo vo Švédsku začali rásť rastliny s agresívnejším peľom. To všetko by nám mohlo poslúžiť ako indícia pre sformulovanie hypotéz analytických štúdií.

Zhrnieme si čo sme si povedali o incidenčných a prevalenčných štúdiách. Medzi ich hlavné *výhody* patrí, že sú dôležité pre plánovanie zdravotníckej starostlivosti a môžu poskytnúť potrebné *indície k ďalším štúdiám*. Na ich realizáciu často *postačujú rutinné databázy*. Medzi ich *nevýhody* patrí, že len popisujú, ale *neanalyzujú vzťah medzi expozíciou a dôsledkom*.

3.1.2 Analytické dizajny štúdií

Analytické štúdie analyzujú vzťah medzi expozíciou a dôsledkom, teda porovnávajú výskyt prípadov v skupine exponovaných a v skupine neexponovaných respondentov. Delíme ich na observačné a experimentálne dizajny štúdií.

V **observačných štúdiách** je expozícia prirodzeným javom. Môžeme napríklad porovnávať, ako rýchlo vyrastú rastliny na poli, kde nikdy neprší, kde prší občas a kde prší každý deň. Dážd' však nijakým spôsobom neovplyvňujeme, iba zapisujeme úhrnné zrážky.

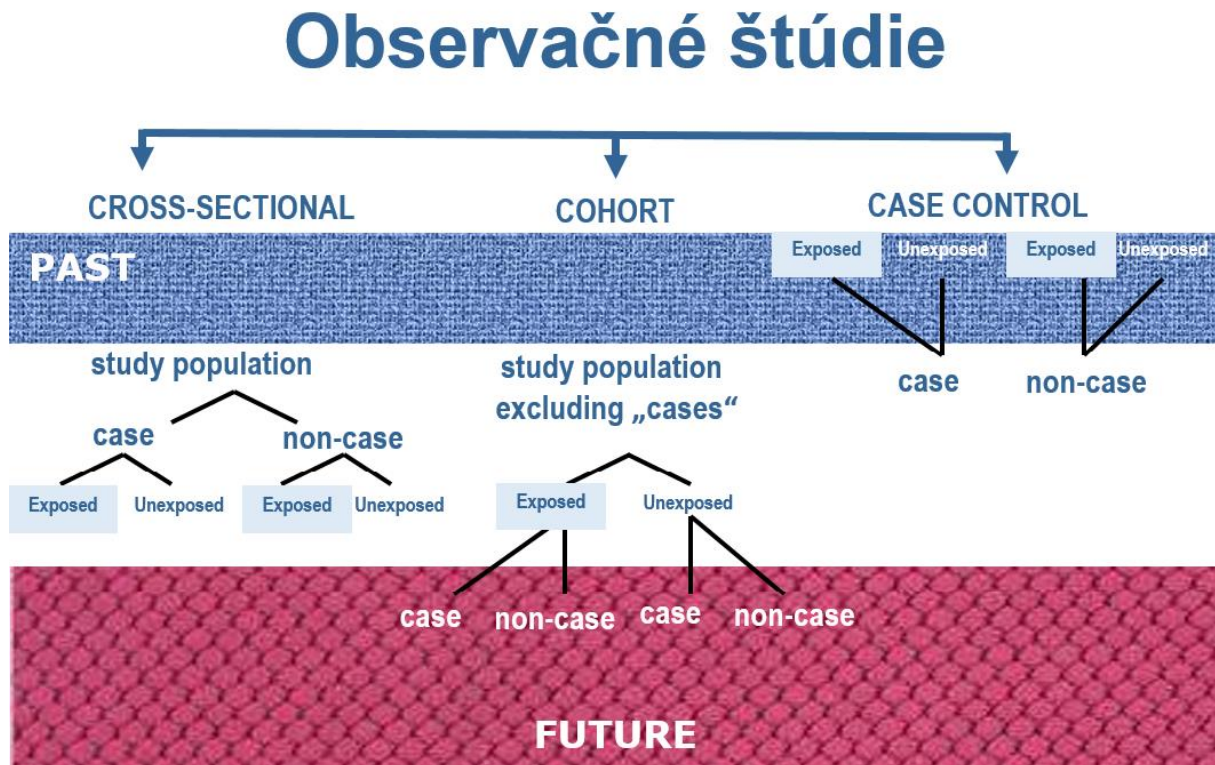
V **experimentálnych štúdiách** je expozícia plánovanou súčasťou experimentu, dôsledok aktívnej manipulácie. Kvetináče napríklad presťahujeme do skleníka a presne určujeme frekvenciu a výdatnosť zavlažovania.

Observačné štúdie ďalej delíme na prierezové (anglicky cross-sectional), kohortové (anglicky cohort) a prípadové kontrolné štúdie (anglicky case-control), ako je to znázornené na obrázku číslo 4. V **prierezových štúdiách** údaje o expozícii aj o dôsledku získavame v tom istom čase. Napríklad sa pýtame, ako často v uplynulom týždni pršalo, a zároveň zisťujeme, či majú príznaky respiračného ochorenia.

V **kohortových štúdiách** v prvom kroku vylúčime všetkých, ktorí už sú prípadmi a počkáme si, kým dôjde alebo nedôjde k expozícii. Následne ich roztriedime na prípady a neprípady. V tomto prípade by sme vylúčili všetkých už chorých na respiračné ochorenie. Ostatných by sme o týždeň skontrolovali, či majú alebo nemajú príznaky respiračného ochorenia, pričom by sme v tom týždni sledovali počasie.

V **prípadových kontrolných štúdiách** najprv vyhľadáme prípady, a potom k nim skúsime vybrať kontrolnú vzorku. Následne dohľadávame k obom skupinám informácie o expozícii. V takomto prípade by sme si v ambulancii všeobecného lekára zobrali lekárske záznamy všetkých liečených na respiračné ochorenie a skúsili k nim vybrať tých, ktorí nemali respiračné ochorenie a porovnávaním ich záznamov by sme hľadali, či boli alebo neboli exponovaní.

V rámci observačných štúdií teda prierezové štúdie zozbierajú všetky údaje naraz, kohortové štúdie si na informáciu o dôsledku počkajú a prípadové kontrolné štúdie vyhľadávajú informácie o expozícii v minulosti.



Obrázok číslo 4: Rozdelenie observačných štúdií.

Ako **príklad prierezovej štúdie** vidíte na obrázku číslo 5 štúdiu od autorov Rosenberger a kol. (2005 b). Dodržiavanie pokynov lekára je veľmi dôležité pre pacientov po transplantácii, aby im darovaná oblička vydržala funkčná čo najdlhšie. Približne polovička pacientov však až tak úplne pokyny lekárov nedodržiava. Autori sa teda rozhodli preskúmať faktory, ktoré súvisia s adherenciou pacienta.

Do výskumného súboru zahrnuli 161 pacientov po transplantácii a v rozhovore sa ich pýtali, ako hodnotia svoje zdravie, či majú niekoho, na koho sa môžu obrátiť, kto ich podporí (teda na sociálnu oporu), a tiež, či im dodržiavanie pokynov a liečba spôsobujú nejaké problémy alebo starosti. Zistilo sa, že pacienti, ktorí hodnotia svoje zdravie horšie, sú menej spokojní so svojou sociálnou oporou a ktorých dodržiavanie pokynov či liečby viac obťažuje, boli častejšie neadherentní, teda nedodržiavali pokyny lekára.

Z tejto štúdie však nevieme zaručiť, či sa pacienti najprv cítili horšie, boli menej spokojní so sociálnou oporou, či mali nežiadúce sprievodné prejavy liečby, alebo či to bolo naopak - najprv porušovali liečebný režim, a až potom sa cítili horšie, pohnevali sa s partnerom a objavili sa nežiadúce symptómy. Môžeme to predpokladať, ale nemôžeme to zaručiť, lebo oboje sme sa pacientov pýtali v jednom a tom istom čase.

Medzi **výhody prierezových štúdií** patrí, že sú *relatívne nenáročné, rýchle a lacné*. Sú menej náročné z hľadiska dodržiavania *etických aspektov*, keďže „*expozícia*“ je prirodzená, nemanipulujeme s ňou, nemáme na ňu vplyv. *Ich nevýhodou* však je, že *expozícia* aj *dôsledok* sa

sledujú v tom istom čase, a preto nie je vždy možné odlíšiť, či „expozícia“ predchádzala, alebo nasledovala sledovaný parameter. Nevieme sa teda vyjadriť ku *kauzalite*.

Ďalším problémom je *kvalita údajov*, ktorá môže byť problematická (závisí od respondentovej pamäte, ochoty reportovať odpovede, od miery sociálnej žiadúcnosti a podobne). Prierezová štúdia taktiež *nie je vhodná pri sledovaní zriedkavých dôsledkov expozície*. Ak by sme chceli s prierezovým dizajnom sledovať faktory súvisiace so zlyhaním obličky u darcov, potrebovali by sme obrovský súbor takýchto respondentov. Našťastie sa to nedeje často. Ak sa to ale predsa udeje, tak na sledovanie takýchto prípadov nám lepšie poslúži iný dizajn štúdie.

> [Transpl Int. 2005 Sep;18\(9\):1072-8. doi: 10.1111/j.1432-2277.2005.00183.x.](#)

Prevalence and characteristics of noncompliant behaviour and its risk factors in kidney transplant recipients

Jaroslav Rosenberger¹, Andrea Madarasova Geckova, Jitse P van Dijk, Iveta Nagyova, Robert Roland, Wim J A van den Heuvel, Johan W Groothoff

Affiliations + expand

PMID: 16101729 DOI: [10.1111/j.1432-2277.2005.00183.x](#)

[Free article](#)

Abstract

Noncompliance with therapy is one possible explanation for the observation that long-term graft survival is not sufficiently improved by the development in immunosuppression. The aim of the study was to explore the prevalence, characteristics and risk factors of noncompliance with immunosuppression. A total of 161 adult kidney transplant recipients were interviewed about their self-rated health, social support, education, stress from adverse effects and compliance with the immunosuppression. The prevalence of subclinical noncompliance was 54%. Noncompliant patients declared significantly worse self-rated health, less satisfaction with social support and higher stress from adverse effects. Male gender (OR 7.5, CI 2.4-23.39), high stress from adverse effects (OR 12.27, CI 2.44-61.88), fair self-rated health (OR 4.45, CI 1.04-19.55) and fair satisfaction with social support (OR 4.55, CI 1.08-19.24) were predictors of noncompliance. Standardized detection methods should be developed with the aim of identifying patients who are at risk of noncompliance in order to prevent graft loss.

[PubMed Disclaimer](#)

Obrázok číslo 5: Príklad prierezovej štúdie.

Teraz si uvedieme *príklad kohortovej štúdie* od autorov Vahtera a kol. (2009), znázornený na obrázku číslo 6. Vo francúzskych elektrárňach GAZEL prebiehala v 90. rokoch rozsiahla longitudinálna štúdia, ktorá sledovala viac ako 20 000 zamestnancov. Tento konkrétny príspevok popisuje štúdiu, v ktorej sledovali, či odchod do dôchodku súvisí s kvalitou spánku. Na problémy so spánkom sa opýtali zamestnancov sedem rokov pred odchodom do dôchodku a následne sedem rokov po odchode do dôchodku. Dobrá správa je, že po odchode do dôchodku mali respondenti významne menej problémov so spánkom. Oslovených bolo takmer 15 000 zamestnancov, pričom na výsledky čakali 14 rokov.

Iným **príkladom kohortovej štúdie** (obrázok číslo 7) je štúdia ELSPAC - Európska longitudinálna štúdia tehotenstva a detstva, ktorá bola iniciovaná Svetovou zdravotníckou organizáciou v roku 1985. Jej cieľom bolo sledovať 40 000 detí narodených v šiestich krajinách Európy. V Čechách boli do tejto štúdie zaradené všetky deti narodené v roku 1991 a 1992 v Brne a Znojme, a to už od tehotenstva matky až po dospelosť narodených detí. Rodičia týchto detí boli sledovaní tiež. Cieľom štúdie bolo zistiť, ktoré biologické, psychologické, sociálne, ekonomické a environmentálne faktory ovplyvňujú priebeh tehotenstva, pôrod, popôrodnú adaptáciu, následný vývin dieťaťa a zdravie adolescentov.

Pri takomto dlhom sledovaní respondentov je však veľmi ťažké udržať ich ochotu zúčastňovať sa na ďalších a ďalších vlnách zberu údajov. Pri tejto štúdii bola spolupráca s rodičmi relatívne dobrá a do piateho roku veku dieťaťa sa im podarilo udržať 70% respondentov, response rate bol teda 70%, pričom 30% respondentov sa už z rôznych dôvodov nepodarilo zapojiť do štúdie. Po 11 rokoch už spolupracovala len polovica rodičov a po 15 rokoch už iba tretina. Od väčšiny respondentov sa teda nepodarilo získať ďalšie informácie (Piler a kol., 2017). Strata respondentov je veľký metodologický problém, pretože nevieme, či nám nechýbajú práve tí najdôležitejší, ktorých účasť by mohla úplne zmeniť zistenia.

> [Sleep](#). 2009 Nov;32(11):1459-66. doi: 10.1093/sleep/32.11.1459.

Effect of retirement on sleep disturbances: the GAZEL prospective cohort study

Jussi Vahtera ¹, Hugo Westerlund, Martica Hall, Noora Sjösten, Mika Kivimäki, Paula Salo, Jane E Ferrie, Markus Jokela, Jaana Pentti, Archana Singh-Manoux, Marcel Goldberg, Marie Zins

Affiliations + expand

PMID: 19928385 PMCID: [PMC2768952](#) DOI: [10.1093/sleep/32.11.1459](#)

Abstract

Objectives: Changes in health following retirement are poorly understood. We used serial measurements to assess the effect of retirement on sleep disturbances.

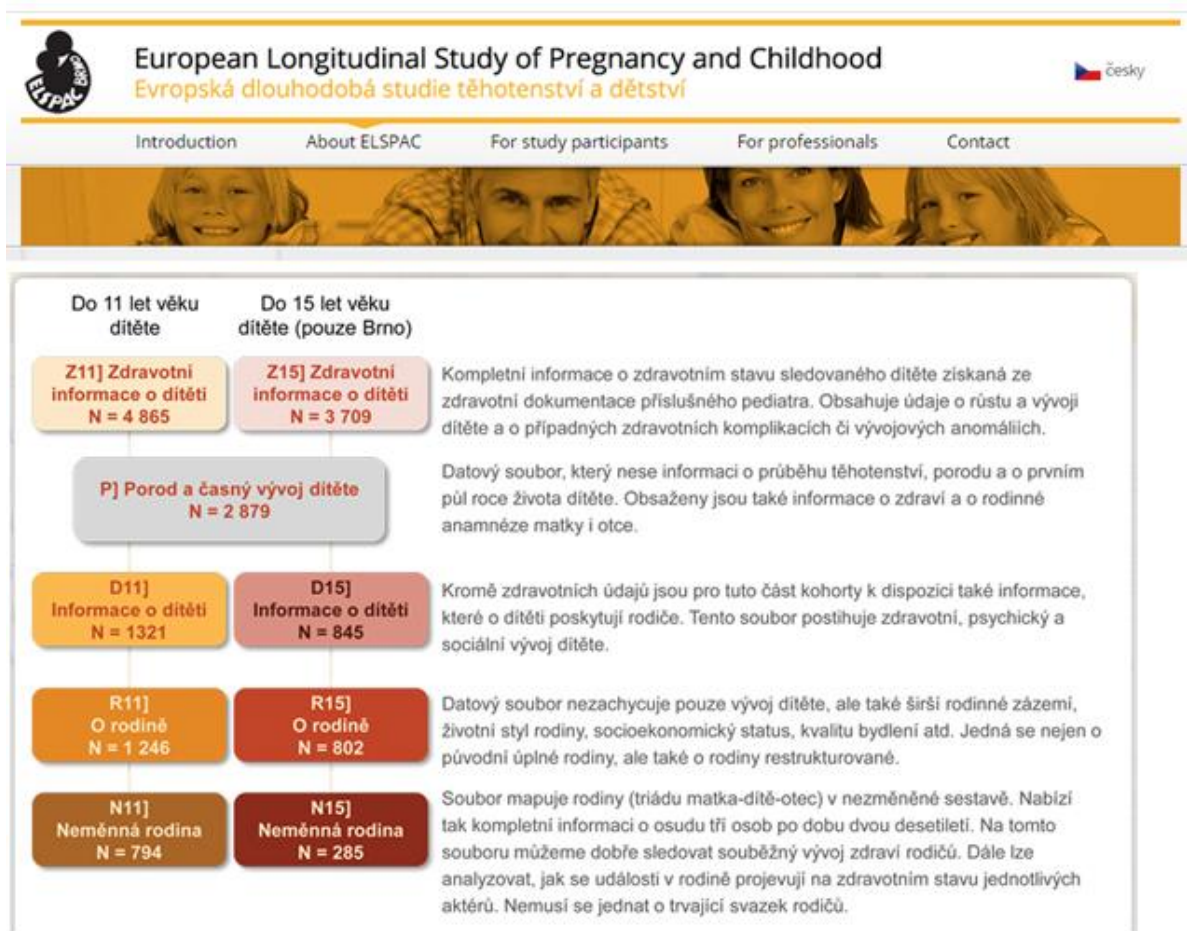
Design: Prospective cohort study.

Setting: The French national gas and electricity company.

Participants: Fourteen thousand seven hundred fourteen retired employees (79% men).

Measurements and results: Annual survey measurements of sleep disturbances ranging from 7 years before to 7 years after retirement (a mean of 12 measurements). Before retirement 22.2% to 24.6% of participants reported having disturbed sleep. According to repeated-measures logistic-regression analysis with generalized estimating equations estimation, the odds ratio (OR) for having a sleep disturbance in the postretirement period was 0.74 (95% confidence interval 0.71-0.77), compared with having a sleep disturbance in the preretirement period. The postretirement improvement in sleep was more pronounced in men (OR 0.66 [0.63-0.69]) than in women (OR 0.89 [0.84-0.95]) and in higher-grade workers than lower-grade workers. Postretirement sleep improvement was explained by the combination of preretirement risk factors suggesting removal of work-related exposures as a mechanism. The only exception to the general improvement in sleep after retirement was related to retirement on health grounds. In this group of participants, there was an increase in sleep disturbances following retirement.

Obrázok číslo 6: Príklad kohortovej štúdie.



Obrázok číslo 7: Príklad kohortovej štúdie.

Problémom môžu byť aj **mätúce premenné**. Predstavte si, že robíte výskum, ako spoločné budovanie bylinkovej záhrady v škole zlepšuje duševné zdravie školákov a ich vzájomné vzťahy. V školách, ktoré sa plánujú a neplánujú zapojiť do tohto projektu, zmeriate duševné zdravie školákov a čakáte, kým školáci zapojených škôl upravujú záhony, vyberajú rastlinky, sadia, plejú. Lenže v meste, kde prebieha vaša expozícia, dôjde k výbuchu plynu, pri ktorom zomrie niekoľko ľudí. Celé mesto žije touto udalosťou niekoľko týždňov. Nech by bola expozícia projektu bylinkovej záhrady akokoľvek silná, nemôže sa porovnávať so silou traumatizujúceho zážitku, ktorého sú plné noviny, televízia alebo rozhovory ľudí. Môžete teda zabaliť svoj projekt bylinkovej záhrady a overovať vplyv hromadného nešťastia na nepriamych svedkov. Pri čakaní na dôsledky v kohortových štúdiách sa vždy deje aj niečo iné, vplývajú tam aj iné mätúce premenné. Čím je to čakanie dlhšie, tým môže byť ich vplyv výraznejší.

Ak vieme o premenných, ktoré by nám pri sledovaní súvislosti expozície s očakávaným dôsledkom mohli robiť neporiadok, mali by sme sa snažiť ich zaznamenať. Napríklad školákov, ktorí neboli v projekte s bylinkovou záhradou, sa budeme pýtať, či nie sú v podobnom projekte vo svojej komunite. Naopak, u školákov, ktorí boli v bylinkovom projekte, by sme mohli sledovať, koľkokrát sa aktívne zúčastnili prác na záhrade a podobne. Môžeme to pridať ako ďalšiu premennú do modelu, ktorú budeme kontrolovať. Alebo ich môžeme eliminovať. Napríklad vyradíme z analýz deti, ktoré sa aktívne nezapájali do prác na bylinkovej záhrade, a tiež tie, ktoré mali byť neexponované, ale zapájali sa do podobného projektu mimo školy.

Medzi *výhody kohortových štúdií* patrí, že je možné *overiť si, či expozícia predchádzala dôsledku*, teda vyjadrovať sa ku kauzalite. Sú výhodné v prípade, že máme *málo poznatkov* o zdravotných dôsledkoch expozície alebo ak je *expozícia zriedkavá*, lebo si vieme expozíciu vyhľadať a na dôsledok počkať. Kohortové štúdie sú menej problematické z *hľadiska etiky*, keďže expozícia je prirodzená, stalo by sa to tak či tak. Nevýhodou kohortových štúdií je ich nákladnosť, *vyššie investície* (časové, personálne, finančné). Môže trvať dlho, kým bude možné merať dôsledky expozície, napríklad kým sa u fajčiara rozvinie chronická obštrukčná choroba pľúc. Ďalšou nevýhodou je, že je veľmi náročné sledovať respondentov, udržať ich spoluprácu v dlhšom období a je teda *možná strata respondentov*. Problémom môže byť aj pôsobenie *mätúcich premenných*.

Teraz si uvedieme **príklad prípadovej kontrolovanej štúdie**, tzv. case-control studies, od autorov Kido a kol. (2009), znázornený na obrázku číslo 8. Autori chceli zistiť, čím sa líšia darcovia obličiek, u ktorých došlo k zlyhaniu zostávajúcej obličky, od darcov, ktorým oblička naďalej dobre slúži. Zlyhanie obličky u človeka, ktorý daroval jednu svoju zdravú obličku, je situácia, ktorej sa treba za každú cenu vyhnúť. Takýto človek nemal byť darcom, lebo bude potrebovať obe svoje obličky.

Takýchto prípadov našťastie nie je veľa a dr. Kido našiel iba 8 darcov, u ktorých došlo k zlyhaniu zostávajúcej obličky v priebehu 10 rokov po darovaní. Kontrolnú skupinu vytvoril z darcov, u ktorých k zlyhaniu nedošlo a snažil sa ich vyberať tak, aby mali podobný vek, rod a čas po darovaní ako skupina so zlyhanou obličkou. Z lekárskeho záznamov vybrali údaje a tieto skupiny následne porovnali. Jeden z darcov prišiel o obličku pri autohavárii, ale u nikoho z darcov nedošlo k zlyhaniu bezprostredne po darovaní. Funkcia obličiek sa zhoršovala postupne, s objavovaním sa pridružených ochorení. V tejto štúdií sa im podarilo potvrdiť aj rozdiely medzi prípadmi a kontrolami vo vybraných faktoroch.

[> Am J Transplant. 2009 Nov;9\(11\):2514-9. doi: 10.1111/j.1600-6143.2009.02795.x. Epub 2009 Aug 14.](#)

How do living kidney donors develop end-stage renal disease?

R Kido ¹, Y Shibagaki, K Iwadoh, I Nakajima, S Fuchinoue, T Fujita, S Teraoka

Affiliations + expand

PMID: 19681812 DOI: [10.1111/j.1600-6143.2009.02795.x](https://doi.org/10.1111/j.1600-6143.2009.02795.x)

[Free article](#)

Abstract

The clinical course and risk factors for developing end-stage renal disease (ESRD) after heminephrectomy in living kidney donors have scarcely been investigated. We reviewed medical records and identified eight case donors who developed chronic kidney disease (CKD) stage 5 or ESRD, and subsequently investigated the association between postoperative clinical courses and changes in renal function. To conduct a case-control study, we also selected a control group comprising 24 donors who had maintained stable renal function and were matched for age, sex and follow-up time since donation. Except for one donor who developed ESRD caused by a traffic accident, none of the donors developed progressive renal dysfunction immediately after donation. Their renal functions remained stable for a long period of time, but started to decline after developing new comorbidities, especially risk factors known as progression factors (proteinuria or hypertension) or accelerating factors (cardiovascular [CV] event or infection) of CKD. As compared with the control donors, incidence of postoperative persistent proteinuria, acute CV event, severe infection and hospitalization due to accelerating factors of CKD were significantly higher in the case donors. These results suggest the importance of long-term (more than 10 years) follow-up of donors with special attention on the risk factors of CKD.

[PubMed Disclaimer](#)

Obrázok číslo 8: Príklad prípadovej kontrolovanej štúdie.

Medzi *výhody prípadových kontrolovaných štúdií* patrí možnosť zozbierať *dostatočný počet „prípadov“*, najmä ak je výskyt zriedkavý v populácii, alebo ak by bolo potrebné na „dôsledok“ dlho čakať. Sú výhodné, ak je málo poznatkov o etiológii ochorenia, pretože umožňujú zozbierať mnoho informácií o možných rizikových faktoroch. Zaberú *menej času, sú menej nákladné a efektívnejšie*, ako napríklad kohortové štúdie. Ich nevýhodou je, že sú náročné na kvalitné zostavenie *kontrolnej skupiny a kvalita údajov závisí od kvality reportovania respondenta*. Nie sú vhodné na *detekovanie slabých súvislostí*, lebo umožňujú vyššie riziko chybného úsudku o súvislostiach.

Ešte tu máme jeden druh observačných štúdií, ktorý pracuje prevažne s rutinne zbieranými údajmi, ako sú údaje zo sčítania obyvateľstva, povinných hlásení zdravotných ukazovateľov, povinných hlásení o nezamestnanosti a mnohých ďalších rutinne, teda pravidelne zhromažďovaných údajov. Takéto štúdie analyzujú agregované údaje, teda parametre za celú jednotku, okres, kraj, školu, nemocnicu a nie údaje za individuálneho respondenta. Nazývajú sa aj **ekologické štúdie**. Na rôznych úrovniach, ako sú kraje, okresy či mestské časti, vieme zozbierať údaje o rôznych parametroch a vložiť ich do analýzy. Napríklad údaje o úmrtnosti v jednotlivých okresoch Slovenska alebo údaje o miere nezamestnanosti v mestských častiach Košíc.

Príklad ekologickej štúdie od autorov Rosicova a kol. (2011) je uvedený na obrázku číslo 9. S využitím voľne dostupných databáz rutinne zbieraných údajov zo sčítania obyvateľstva a z vykazovania úmrtnosti v jednotlivých okresoch Slovenska sa autori pokúsili identifikovať expozíciu a premenné, ktoré by mohli súvisieť s úmrtiami, ktoré je možné pripísať nadmernej konzumácii alkoholu. Použili na to stromovú analýzu, ktorá spomedzi premenných vyhľadáva a vyberá tie, ktoré prispievajú k vysvetleniu variácie, teda rozdielov v sledovanom parametri - v tomto prípade miere úmrtnosti pripísateľnej nadmernej konzumácii alkoholu. Takmer štvrtina úmrtí mužov v sledovanom období (2001-2003) bola zapríčinená nadmernou konzumáciou alkoholu. Miera úmrtnosti podľa príčin smrti súvisiacich s konzumáciou alkoholu mužov (ale nie žien) súvisela so socioekonomickými charakteristikami okresu. Najvýznamnejšie prispievala miera nezamestnanosti a nízka vzdelanostná úroveň, zatiaľ čo podiel rómskej populácie a priemerný príjem neboli štatisticky významnými faktormi.

Medzi *výhody ekologických štúdií* patrí to, že sú *rýchle, finančne nenáročné* a vystačia si často s *rutinne zbieranými údajmi*, ktoré bývajú voľne dostupné. Ak sú opatrne interpretované, môžu významne prispieť k poznaniu. Ich nevýhodou je *ekologická chyba*, teda chybné vzťahovanie zistení na úrovni skupiny alebo jednotlivcov, ktorí nemusia byť aktérmi týchto zistení. Ekologické štúdie majú tendenciu *nadhodnotiť mieru vzťahu* medzi skúmanými premennými.

V rámci analytických dizajnov štúdií si ako posledný dizajn predstavíme **experimentálne štúdie**. V experimentálnych štúdiách výskumník manipuluje s expozíciou v kontrolovaných podmienkach. To má obrovský dosah na etické aspekty, pretože tým výskumník preberá zodpovednosť za dôsledky expozície. Čo ak odloženie liečby v experimente, alebo práve naopak podanie lieku spôsobí zhoršenie stavu alebo nebudaj úmrtie pacienta? Experimentálne štúdie teda majú obrovskú silu z hľadiska možnosti vyjadrovať sa ku kauzalite, lebo kontrolujú expozíciu, kontrolujú podmienky experimentu, kontrolujú dôsledky a ich časovú následnosť po expozícii, ale sú limitované možnými dôsledkami na pacienta a etické aspekty musia byť dôsledne zvažované. Najznámejším typom experimentálnych štúdií sú **randomizované kontrolované štúdie**, v ktorých sa znáhodňuje, kto bude v kontrolnej a kto v experimentálnej skupine, pričom môžu byť rôznym spôsobom zaslepené. Buď iba respondent alebo respondent aj výskumník nevedia, kto do ktorej skupiny patrí - napríklad komu bola podaná účinná látka a komu placebo. Experimentálnymi štúdiami sú ale aj **intervenčné štúdie**, v ktorých sa porovnáva, či intervencia implementovaná v danej komunite je účinná.

Regional socioeconomic indicators and ethnicity as predictors of regional infant mortality rate in Slovakia

Katarina Rosicova · Andrea Madarasova Geckova ·
Jitse P. van Dijk · Jana Kollarova · Martin Rosic ·
Johan W. Groothoff

Received: 29 October 2009 / Revised: 29 April 2010 / Accepted: 12 September 2010 / Published online: 26 October 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract

Objective Exploring the associations of regional differences in infant mortality with selected socioeconomic indicators and ethnicity could offer important clues for designing public health policy measures.

Methods Data included perinatal and infant mortality in the 79 districts of the Slovak population in 2004. Linear regression was used to analyse the contribution of education, unemployment, income and proportion of Roma

population on regional differences in perinatal and infant mortality rates.

Results All the explored socioeconomic indicators and ethnicity individually contributed significantly to both perinatal and infant mortality, with the exception of income. In the model exploring the influence of all these variables together on perinatal and infant mortality, only the effect of the proportion of Roma population remained significant. This model explained 34.9% of the variance for perinatal and 36.4% of the variance for infant mortality.

Conclusions Living in Roma settlements indicates an accumulation of socioeconomic disadvantage. Health literacy, health-related behaviour and many other factors might contribute to the explanation of the differences in infant mortality, and a better understanding of these processes might help us to design tailored interventions.

This paper belongs to the special issue "Migrants and ethnic minorities in Europe: new challenges for public health research and practice".

K. Rosicova (✉)
Kosice Self-Governing Region, Department of Regional
Development and Land-Use Planning, Kosice, Slovakia
e-mail: katarina.rosicova@vucke.sk

K. Rosicova · A. Madarasova Geckova · J. P. van Dijk

Keywords Infant mortality · Regional differences ·
Roma population · Socioeconomic indicators

Obrázok číslo 9: Príklad ekologickej štúdie.

Každý druh výskumného dizajnu reprezentuje odlišný spôsob získavania informácií. V princípe s každým typom dizajnu sa dá vedecky poznávať, líšia sa však v sile dôkazu, ktorá je do veľkej miery závislá od toho, nakoľko sa nám podarilo vyhnúť sa chybám pri realizácii výskumu. Pri overovaní účinnosti lieku, od ktorého závisí prežitie pacienta, asi budete hľadať dizajn s čo najväčšou silou dôkazu a siahnete po randomizovanej dvojito zaslepenej štúdií. Na identifikovanie faktorov, ktoré by mohli mať vplyv na rozdiely v úmrtnosti na vybrané ochorenie medzi okresmi, však vieme získať údaje aj z voľne dostupných databáz, bez nutnosti terénneho výskumu.

Ako si teda vybrať vhodný výskumný dizajn? Dobrým spôsobom je odpovedať si na nasledujúce štyri otázky:

- (1) Umožní mi vybraný dizajn odpovedať na moju **výskumnú otázku**?
- (2) Je tento dizajn v poriadku z hľadiska **etiky**?
- (3) Aké **zdroje** mám k dispozícii (časové, finančné, personálne)?
- (4) Existuje **iný spôsob**, ktorý by bol lacnejší, rýchlejší a zároveň by odpovedal na výskumnú otázku?

Cieľom je popísať vzťah medzi expozíciou a dôsledkom **validne** (s vylúčením skreslenia, chýb), **presne** a s **využitím minimálnych zdrojov**. Výber vhodného dizajnu štúdie je kľúčovým faktorom kvality údajov. Predpoklad, že údaje sú zbierané dôveryhodným spôsobom, by nemal byť prijatý automaticky.

Rozličné vonkajšie faktory môžu interferovať s hodnoteným vzťahom medzi expozíciou a dôsledkom a máriť presné zhodnotenie tohto vzťahu. Dizajn štúdie je len taký dobrý, ako je dobrá logika a použité nástroje na identifikovanie potencionálnych mäťúcich premenných. Chyba pri výbere dizajnu alebo realizácii štúdie vedie k identifikovaniu falošnej súvislosti medzi expozíciou a dôsledkom.

3.2 Iniciatívy a odporúčania

V rámci tejto časti si predstavíme iniciatívy a odporúčania, ktoré nám môžu pomôcť skontrolovať, či sme pri výbere konkrétneho dizajnu na nič nezabudli, či sme všetko správne popísali v protokole. S cieľom zlepšiť a skvalitniť reportovanie výskumníkov o použitých metódach, teda zlepšiť transparentnosť vedy a uľahčiť verifikovateľnosť zistení, vzniklo viacero iniciatív a odporúčaní, ktorých výstupom sú rôzne kontrolné zoznamy (anglicky checklist). Jedným z nich je napr. STROBE (skratka z anglického STrengthening the Reporting of OBservational studies in Epidemiology) na obrázku číslo 10, voľne dostupný na www.strobe-statement.org.

Je to akýsi kontrolný zoznam, či ste v recepte popísali všetky ingrediencie a postupy takým spôsobom, že to zvládne aj váš partner, keď ostane na víkend sám doma. Na obrázku číslo 10 je len časť kontrolného zoznamu iniciatívy STROBE pre prierezové štúdie a v ľavom stĺpci môžete vidieť zoznam ingrediencií, ako sú názov a abstrakt, úvod, zdôvodnenie, ciele, typ dizajnu, kontext, populácia a vzorka, premenné, meracie nástroje atď. V pravom stĺpci sú pripomienky, na čo pri popise v protokole či vedeckom článku netreba zabudnúť. Napríklad pri vzorke je potrebné popísať inkluzívne a exkluzívne kritéria a spôsob výberu respondentov.

STROBE Statement—Checklist of items that should be included in reports of *cross-sectional studies*

	Item No	Recommendation
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why

Obrázok číslo 10: Dokument STROBE obsahujúci kontrolný zoznam položiek, ktoré by mali byť zahrnuté v prierezových štúdiách.

Takýchto iniciatív je viac a kontrolné zoznamy či iné pomôcky sú voľne dostupné na stiahnutie. STROBE ponúka kontrolné zoznamy pre prierezové, kohortové a kontrolované prípadové štúdie (voľne dostupné na www.strobe-statement.org). MOOSSE (anglicky Meta-analysis of Observational Studies in Epidemiology) a QUOROM (anglicky Quality Of Reporting Of Meta-analyses Statement) sa venujú metaanalýzám, CONSORT (anglicky The Consolidated Standards of Reporting Trials) sa venuje randomizovaným kontrolovaným štúdiám (voľne dostupné na <https://www.consort-spirit.org/>).

OTÁZKY

1. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Identifikovať príčinu zriedkavého ochorenia.
 - A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia

2. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Popísať nové ochorenie, alebo zriedkavý priebeh ochorenia.
 - A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia

3. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Popísať výskyt nejakého ochorenia v rôznych krajinách alebo v rôznom čase (napr. v roku 1990 a 2010).
 - A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia

4. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Analyzovať výskyt chrípky v zimných a jarných mesiacoch.
 - A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia

- E) Prípadová štúdia
- F) Experimentálna štúdia (RCT)
- G) Prevalenčná štúdia
- H) Metaanalýza
- I) Incidenčná štúdia

5. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Skúmať súvislosť medzi expozíciou a dôsledkom ak potrebujeme informácie relatívne rýchlo a bez veľkých nákladov. Napr. overiť či rod súvisí s fyzickou aktivitou, teda či sú vo veku 15 rokov chlapci fyzicky aktívnejší ako dievčatá.

- A) Ekologická štúdia
- B) Prierezová štúdia
- C) Prípadová kontrolovaná štúdia
- D) Kohortová štúdia
- E) Prípadová štúdia
- F) Experimentálna štúdia (RCT)
- G) Prevalenčná štúdia
- H) Metaanalýza
- I) Incidenčná štúdia

6. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Identifikovať dlhotrvajúci vplyv zriedkavej expozície, napr. posttraumatický stres alebo posttraumatický rast účastníkov teroristického útoku.

- A) Ekologická štúdia
- B) Prierezová štúdia
- C) Prípadová kontrolovaná štúdia
- D) Kohortová štúdia
- E) Prípadová štúdia
- F) Experimentálna štúdia (RCT)
- G) Prevalenčná štúdia
- H) Metaanalýza
- I) Incidenčná štúdia

7. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Ak potrebujete overiť či expozícia predchádzala dôsledku. Napr. potrebujete overiť či problémy so spánkom súvisia s nočnými zmenami alebo si nočnú prácu vyberajú ľudia, ktorí aj tak v noci nemôžu spať.

- A) Ekologická štúdia
- B) Prierezová štúdia
- C) Prípadová kontrolovaná štúdia
- D) Kohortová štúdia
- E) Prípadová štúdia
- F) Experimentálna štúdia (RCT)
- G) Prevalenčná štúdia
- H) Metaanalýza
- I) Incidenčná štúdia

8. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Ak potrebujete overiť súvislosť expozície s dôsledkom, ale z etických dôvodov nie je možné manipulovať s expozíciou. Napr. overiť či vypitie litra vína súvisí s nauzeou alebo fajčenie s kašľaním.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
9. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Ak máme k dispozícii rutinne zbierané údaje (štatistické hlásenia) a vieme dohľadať regionálne rozdiely v niektorej expozícii a sledovanom dôsledku. Napr. rozličný prístup k zdravotníckej starostlivosti a mieru dojčenskej úmrtnosti.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
10. Aký typ dizajnu štúdie sa podľa Vás hodí na nasledovný výskumný problém: Máme niekoľko štúdií na menších súboroch, ale potrebujeme sa oprieť o zistenie s väčšou silou. Napr. v rôznych krajinách sú publikované štúdie o prežívaní po transplantácii. Takýchto pacientov však nie je veľa, súbory týchto štúdií sú malé, ale my potrebujeme silnejší dôkaz integrujúci tieto zistenia na menších súboroch.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia

11. Aký typ štúdie je popísaný v nasledujúcom scenári? Štúdia skúma úmrtnosť z dôvodu rakoviny krčka maternice v krajinách EÚ v závislosti od priemerného zastúpenia žien daného štátu, ktoré podstupujú každoročne skrining (Pap test).
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
12. Aký typ štúdie je popísaný v nasledujúcom scenári? Štúdia porovnáva výskyt bolesti chrbta v skupine zamestnancov údržby, v skupine pekárov a v skupine cukrárov.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
13. Aký typ štúdie je popísaný v nasledujúcom scenári? Štúdia hodnotí vzťah medzi výskytom rakoviny prsníka a kojením v anamnéze žien. Výskumník vybral ženy, u ktorých bola diagnostikovaná rakovina prsníka a podľa veku k nim vybral ženy z tých istých oblastí. Respondentiek sa pýtal, či v minulosti kojili niektoré zo svojich detí.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
14. Aký typ štúdie je popísaný v nasledujúcom scenári? Štúdia skúma 2 spôsoby liečby rakoviny prsníka. Ženy v prvom stupni ochorenia boli náhodne rozdelené do dvoch skupín, pričom v prvej skupine bola aplikovaná iba lumpektómia, kým v druhej skupine bola aplikovaná lumpektómia a radiačná terapia. Ženy boli sledované 5 rokov a skúmali sa rozdiely v recidíve a prežívaní.
- A) Ekologická štúdia

- B) Prierezová štúdia
- C) Prípadová kontrolovaná štúdia
- D) Kohortová štúdia
- E) Prípadová štúdia
- F) Experimentálna štúdia (RCT)
- G) Prevalenčná štúdia
- H) Metaanalýza
- I) Incidenčná štúdia

15. Aký typ štúdie je popísaný v nasledujúcom scenári? Skúmaný bol vzťah medzi ožiarením pľúc a následným rizikom pre rakovinu prsníka. Štúdia začala v roku 2005. V tejto štúdii ženy, ktoré podstúpili radiačnú terapiu v súvislosti s popôrodnou mastitídou v roku 1940 boli porovnávané so ženami, ktoré podstúpili neradiačnú terapiu pre popôrodnú mastitídu v roku 1940. Výskyt (incidencia) rakoviny prsníka v oboch skupinách bola sledovaná v období 50 rokov po podstúpení tej či onej terapie.
- A) Ekologická štúdia
 - B) Prierezová štúdia
 - C) Prípadová kontrolovaná štúdia
 - D) Kohortová štúdia
 - E) Prípadová štúdia
 - F) Experimentálna štúdia (RCT)
 - G) Prevalenčná štúdia
 - H) Metaanalýza
 - I) Incidenčná štúdia
16. Observačný dizajn štúdie skúmajúci prevenciu alebo liečbu je obvykle zvolený, ak experimentálny dizajn štúdie by bol neetický alebo neprípustný.
- A) Je to pravda.
 - B) Nie je to pravda.
17. Hlavným obmedzením observačného dizajnu štúdie je, že výskumník nemôže mať úplnú kontrolu nad externými faktormi, ktoré by mohli skresliť štúdiu.
- A) Je to pravda.
 - B) Nie je to pravda.
18. V štúdii, ktorá mala prierezový dizajn bola skúmaná súvislosť medzi typom krvnej skupiny a výskytom šedého zákalu u respondentov. Potvrdená súvislosť (častejší výskyt šedého zákalu v skupine ľudí s určitým typom krvnej skupiny) môže viesť k nesprávnym záverom o kauzálnom vzťahu medzi príčinou (typ krvnej skupiny) a následkom (výskyt šedého zákalu), pretože tento typ dizajnu nám neumožňuje posúdiť časovú následnosť, teda či príčina predchádzala následku.
- A) Je to pravda.
 - B) Nie je to pravda.
19. V štúdii, ktorá má ekologický dizajn bola skúmaná súvislosť medzi percentuálnym zastúpením obyvateľstva žijúceho v Rómskych osadách v okrese a detskou úmrtnosťou v danom okrese.

Táto súvislosť bola štatisticky významná na úrovni okresov. Na základe tejto štúdie môžeme tvrdiť, že novorodenci matiek z Rómskych osadách zomierajú častejšie ako novorodenci matiek, ktoré v Rómskych osadách nežijú.

- A) Je to pravda.
- B) Nie je to pravda.

20. V štúdiu, ktorá mala prierezový dizajn bola skúmaná súvislosť medzi aktuálnou konzumáciou alkoholu (za posledný týždeň) a výskytom šedého zákalu u respondentov. Potvrdená súvislosť (častejší výskyt šedého zákalu v skupine ľudí, ktorí v uplynulom týždni konzumovali alkohol) môže viesť k nesprávnym záverom o kauzálnom vzťahu medzi príčinou (aktuálna konzumácia alkoholu) a následkom (výskyt šedého zákalu), pretože tento typ dizajnu nám neumožňuje posúdiť časovú následnosť, teda či príčina predchádzala následku.

- A) Je to pravda.
- B) Nie je to pravda.

21. Strata respondentov je problémom v prípade kohortových štúdií, ale nie je problémom v prípade experimentálneho dizajnu štúdie.

- A) Je to pravda.
- B) Nie je to pravda.

4 VALIDITA ŠTÚDIE

V tejto kapitole si predstavíme, ako rozlíšiť a popísať vnútornú a vonkajšiu validitu štúdie. Ukážeme si, ako identifikovať alternatívne vysvetlenia pozorovaných súvislostí, ako je náhodná chyba, systematická chyba a vplyv mäťúcich premenných. Taktiež si ukážeme, ako rozpoznať najvýznamnejšie zdroje systematickej chyby.

4.1 Vnútoraná a vonkajšia validita štúdie

Učiteľka v jednej škole chcela povzbudiť žiakov, aby si poriadne umývali ruky a navrhla experiment. Vybrali troch dobrovoľníkov: jeden mal neumyté ruky, druhý si ruky natrel antibakteriálnym gélom a tretí si ruky vyumýval mydlom a vodou. Potom ich pani učiteľka požiadala, aby čo najviac ochytali pridelený krajec chleba. Krajce chleba následne vložili do igelitových vreciek na chladné a suché miesto a každý deň pozorovali čo sa deje. Výsledok experimentu sa dostavil po niekoľkých dňoch a bol to krásny spôsob, ako ukázať žiakom, prečo je dôležité si ruky umývať, aj keď nevyzerajú špinavé.

Môžeme sa na takéto zistenie spoľahnúť? Je tento školský experiment dôveryhodný? Je vzťah medzi čistotou rúk a rastom plesní na krajci chleba skutočná, náhodná alebo zapríčinená niečím iným? Je pleseň na krajci chleba skutočne dôkazom patogénov na rukách? Môžeme tieto zistenia zovšeobecniť? Ak vyrástla na krajci chleba veľká pleseň, dá sa z toho usudzovať, že špinavými rukami prenášame choroby? Alebo, že umývanie rúk mydlom pod tečúcou vodou je efektívnejšie ako natieranie antibakteriálnym gélom? Inými slovami je táto štúdia validná?

Štúdia je validná, ak sme merali skutočne to, čo sme merať chceli a pozorované súvislosti sú skutočne súvislosti sledovaných premenných a nie iných. Taktiež pri opakovanom sledovaní na inej vzorke z danej populácie by malo byť možné zopakovať pozorovanie týchto súvislostí.

Validita štúdie je podmienkou objektivity a verifikovateľnosti poznávania, čo je jeden zo základných princípov vedeckého poznávania. K chybám, zmäteniu, teda ohrozeniu validity môže pritom dôjsť v každom kroku výskumného procesu: už pri formulovaní výskumnej otázky, pri zvolení dizajnu štúdie, pri výbere vzorky, výbere premenných, ich úprave, zbere, analýze, interpretovaní. Počas celého výskumného procesu musíme uvažovať o validite štúdie a zbierať údaje, ktoré by nám ju pomohli priebežne hodnotiť a v závere pri reportovaní výsledkov popísať.

Rozlišujeme **vnútornú a vonkajšiu validitu**.

Pri **vnútornej validite** hodnotíme zdroje chýb smerov dovnútra, pýtame sa, či je vzťah medzi expozíciou a následkom pravdivý.

Vonkajšia validita vypovedá o uplatniteľnosti zistení vo vybranej skupine na populáciu, teda smerom von. Hodnotí, nakoľko tieto zistenia budú platiť v inej nemocnici, v inej krajine, v inom mesiaci, v inom období, v inej kultúre.

Pri hodnotení vonkajšej validity sa zamýšľame, či sú naše zistenia zovšeobecniteľné, teda či použité metódy, výber vzorky alebo poznanie o danom výskumnom probléme nám umožňuje usudzovať, že by to malo platiť aj v inom výskumnom súbore pri opakovaní merania.

V školskom experimente použili bežne zašpinené ruky bežného školáka, bežný antibakteriálny gél a bežné mydlo a tečúcu vodu a bežné krajce chleba. Takáto špina na rukách, gél, mydlo, voda či chlieb by sa našlo asi aj vo vedľajšej triede, asi v každej triede tejto krajiny a asi aj v susednej krajine a môžeme uvažovať, že ruky školákov v Afrike alebo na Sibíri sú osídlené aj inými baktériami. Naše metódy dobre zastupujú bežné podmienky školákov na celom svete. Naša vzorka nemá dôvod líšiť sa od populácie

školákov na celom svete. A naše poznanie o danom probléme ilustruje napríklad výskum v jednom laboratóriu, kde urobili stery z rúk školákov a v agarovom géli im vyrástli kolónie baktérie rodu Bacillus, baktéria E.coli zo stolice, kvasinky, nejaké chlpaté plesne a dokonca aj baktérie druhu Salmonela. U týchto všetkých sa dá predpokladať, že budú na rukách školákov od severného pólu po južný, a k nim sa možno pridajú aj nejaké miestne verzie. Krajec chleba a igelitové vrecúško však nie je petriho miska a agarový géľ. V tomto bode máme trochu problém, lebo ak by tam boli baktérie, ktorým chlieb vyslovene nechutí, tak na ňom nenarastú. Výskumný problém, použité metódy, výber vzorky školákov a poznanie o danom jave nám dáva nádej, že tieto zistenia sú zovšeobecniteľné – špinavé ruky školáka dokážu vyčarovať plesne aj napriek antibakteriálnemu gélu. Máme však trochu problém so zovšeobecnením, lebo nevieme, ako by reagovali baktérie, ktoré nemajú rady chlieb.

4.2 Rozlíšenie náhodnej chyby, systematickej chyby a vplyvu mäťúcich premenných

Vnútorňú validitu štúdie môžu ohrozovať najmä tri zdroje. Súvislosť medzi expozíciou a dôsledkom je pravdivá iba ak sa podarilo eliminovať tri alternatívne vysvetlenia pozorovanej súvislosti:

1. náhodná chyba,
2. systematická chyba,
3. vplyv mäťúcich premenných.

O **náhodnej chybe** hovoríme, ak je pozorovaná súvislosť medzi expozíciou a sledovaným parametrom náhodná. To, či je alebo nie je náhodná, overujeme použitím štatistických metód a odhadom intervalov spoľahlivosti, ktoré sme si vysvetlili v 2. kapitole s názvom Meranie vplyvu. Hovoria o tom, či by sme s 95% alebo 99% pravdepodobnosťou pri opakovaní výskumu na inej vzorke danej populácie dospeli k rovnakým zisteniam.

Systematická chyba je chyba v dizajne alebo realizácii štúdie, ktorá vedie k nepravdivému pozorovaniu.

Vzťah medzi expozíciou a sledovaným parametrom môže byť skreslený, teda maskovaný, zvýraznený, sprostredkovaný ich súvislosťou s inou premennou, tzv. mäťúcou premennou, kvôli **vplyvu mäťúcich premenných**. Ak dochádza k modifikácii súvislosti, teda maskovaniu alebo zvýrazneniu, hovoríme o moderátoroch a o moderácii. Ak dochádza k sprostredkovaniu vplyvu, hovoríme o mediátoroch a o mediácii.

Podrobnejšie sa budeme venovať systematickej chybe, ktorá sa môže vyskytnúť v ktorejkoľvek fáze výskumného procesu. Môže sa udiť počas dizajnovania štúdie, zberu údajov, analýzy údajov a ich interpretácie a aj pri reportovaní zistení. Najčastejšou skupinou systematických chýb sú chyby pri výbere vzorky, tzv. selection bias a chyby pri pozorovaní, tzv. observational bias.

Na vysvetlenie systematickej chyby pri výbere vzorky použijeme príklad zo života. Počas druhej svetovej vojny v snahe znížiť straty v letectve sa rozhodli analyzovať všetky poškodené lietadlá, ktoré doleteli späť na základňu. Zakreslili, kde všade boli zasiahnuté strelami a obrázky lietadiel preložili cez seba, aby identifikovali najčastejšie zasiahnuté časti lietadla. Nemohli opancierovať celé lietadlo, ale mohli pridať pancier na niektoré miesta. Zdalo sa im logické, že najčastejšie zasiahnuté miesta sú tou správnou voľbou. Pre istotu si ale zavolali experta, aby posúdil či uvažujú správne. A on sa opýtal, kde sú tie lietadlá, ktoré na letisko nedoleteli. Lietadlá, ktoré doleteli sú tie, ktoré to zvládli napriek zásahu. Tie, ktoré nedoleteli boli pravdepodobne zasiahnuté v strede trupu alebo na chvoste, teda na úplne

iné časti lietadla. Tento príklad ilustruje systematickú chybu pri výbere vzorky, ktorá viedla k nepravdivej súvislosti medzi skúmanými premennými. Pri zaradení lietadiel, ktoré nedoleteli, by sme dospeli k úplne iným záverom. Preto je dôležité pri výbere vzorky rozmýšľať, kde sú tie lietadlá (alebo pacienti, respondenti), ktoré nedoleteli (v prípade pacientov, respondentov, ktorí neboli zastihnutí, oslovení, odmietli účasť) a či to nie sú práve oni, ktorí sú pre náš výskum dôležití?

Napríklad, výskumu o vedľajších účinkoch lieku sa môžu zúčastniť dobrovoľne iba pacienti, ktorí majú tendenciu dobre spolupracovať, byť organizovaní a dodržiavajú odporúčania. U týchto pacientov sa ale nemusia žiadne vedľajšie účinky vyskytnúť. V bežnej praxi môže byť skupina pacientov, ktorí nie sú veľmi naklonení spolupracovať, nezúčastňujú sa teda výskumov a tiež nie sú organizovaní, nedodržia odporúčania. Práve táto skupina môže byť riziková z hľadiska výskytu nežiadúcich účinkov lieku. To sú tie lietadlá, ktoré nám dobrovoľne a ochotne do výskumnej vzorky nepristanú.

Systematická chyba pri výbere vzorky je chyba v dôsledku systematických odlišností v charakteristikách tých, ktorí sa zúčastnili a ktorí sa nezúčastnili výskumu. Vyplýva teda zo spôsobu výberu respondentov, ich oslovovania a sledovania.

V kontrolovanej prípadovej štúdiu môžeme použiť odlišné kritéria pre výber prípadov a pre výber kontrolnej skupiny, pričom tieto kritéria môžu súvisieť so sledovanými parametrami. Hovoríme vtedy o **systematickej chybe pri výbere kontrol**. Môžeme skúmať, či preventívne opatrenie súvisí s výskytom ochorenia. Ak by sme chceli zistiť, či absolvovanie preventívnej gynekologickej prehliadky vrátane sterov z krčka maternice súvisí s výskytom rakoviny krčka maternice, mohli by sme využiť kontrolovanú prípadovú štúdiu.

Novo diagnostikované prípady by sme vyhľadali v zdravotných záznamoch nemocnice a kontrolnú skupinu by sme vybrali zo spádovej oblasti nemocnice, teda z lokalít, odkiaľ pochádzajú aj prípady. Počas pracovných dní a v pracovnom čase by sme prechádzali touto oblasťou a podľa istého kľúča klopali na dvere bytov. Žien, ktoré by boli doma a otvorili dvere, by sme sa pýtali, či boli v uplynulom roku na preventívnej gynekologickej prehliadke a absolvovali ster z krčka maternice.

Predpokladajme, že máme 250 prípadov a 250 kontrol. Z toho v oboch skupinách absolvovalo preventívnu prehliadku 100 žien, čo znamená, že 40% žien v skupine prípadov aj kontrol absolvovalo preventívnu prehliadku. Vypočítame si pomer šancí, v našom prípade sa to rovná 1, čo znamená, že exponované i neexponované ženy majú rovnakú šancu byť prípadom, teda mať rakovinu krčka maternice. Ale čo ak došlo k systematickej chybe pri výbere kontrol? Čo ak sme týmto spôsobom oslovili iba nezamestnané ženy, ktoré majú aj nižšiu pravdepodobnosť absolvovať zdravotné prehliadky hrazené zo zdravotného poistenia, prípadne vyžadované zamestnávateľom u zamestnaných žien?

Ak by sme zahrnuli aj zamestnané ženy, počet exponovaných a neexponovaných by ostal v skupine prípadov rovnaký, ale v skupine kontrol by sme mali viac exponovaných a menej neexponovaných. V skupine prípadov by teda bolo 40% exponovaných a v skupine kontrol 60% exponovaných, teda tých, ktoré absolvovali preventívnu prehliadku. Po výpočte pomeru šancí dostaneme hodnotu 2,25. To znamená, že skupina neexponovaných, teda tých čo neabsolvovali preventívnu prehliadku, má viac ako dvakrát vyššiu pravdepodobnosť výskytu rakoviny krčka maternice. V tomto prípade by sme sa tomuto problému mohli vyhnúť, ak by sme použili rovnaké kritéria pri výbere respondentov do skupiny prípadov a do skupiny kontrol, napríklad, ak by sme na dvere klopali aj počas víkendov alebo mimo pracovného času.

Systematická chyba pri oslovovaní respondentov je ďalší typ systematickej chyby pri výbere vzorky, ku ktorej môže dôjsť. Pri oslovovaní respondentov sa môže stať, že ochota zúčastniť sa alebo odmietnuť účasť bude súvisieť s dôsledkom. Predstavte si, že robíte výskum, ako fajčenie súvisí s kašľaním u

dospievajúcich. Oslovíte v škole žiakov, či by neboli ochotní zúčastniť sa takéhoto výskumu. Nájde sa 50 dobrovoľníkov, ktorých sa opýtate, či fajčia a či kašľú. V skupine kašľajúcich by bolo 10 fajčiarov a 10 nefajčiarov, teda polovica z nich by bola exponovaná fajčeniu. V skupine nekašľajúcich by bolo 10 fajčiarov a 20 nefajčiarov, teda približne tretina z nich by boli fajčiari. Z pomeru šancí zistíme, že fajčiari majú dvakrát vyššiu šancu kašľať.

Čo však v prípade, že fajčiarov tento výskum veľmi neoslovil. Nie je príjemné si priznať, že fajčenie by mohlo mať zdravotné dôsledky. Nie je to niečo, s čím by sa potrebovali chváliť. Práve tí, ktorí skutočne majú problém s kašľaním sa mohli rozhodnúť nezúčastniť sa výskumu a mohol to byť až každý druhý z kašľajúcich fajčiarov. Ak by to bol každý druhý a nám by sa ich podarilo presvedčiť, aby sa zapojili do výskumu, zmenil by sa počet kašľajúcich fajčiarov. Namiesto 10 by sme ich mali vo vzorke 20. V skupine kašľajúcich by bolo exponovaných fajčeniu 66% respondentov a v prípade nekašľajúcich 33% respondentov. Po výpočte pomeru šancí by tak fajčiari mali 4-krát vyššiu pravdepodobnosť kašľať oproti nefajčiarom, čo je skoro dvojnásobne viac ako v prípade, keď došlo k systematickej chybe vo výbere vzorky. Predísť tejto chybe je možné stimulovaním vysokej účasti, motivovaním respondentov zúčastniť sa výskumu, ale tiež overovaním parametrov u tých, ktorí odmietnu účasť. V tomto prípade by sme napríklad mohli pri oslovovaní požiadať tých respondentov, čo odmietnu, aby nám aspoň odpovedali na otázku či fajčia, a podľa toho by sme mohli zistiť, že účasť na štúdiu odmietlo viac fajčiarov ako nefajčiarov.

V kohortových alebo experimentálnych štúdiách strácame respondentov počas sledovania. Ak táto strata nie je náhodná, ale systematická a súvisí so sledovanými parametrami, môže to viesť k systematickej chybe pri výbere vzorky súvisiacej so stratou respondentov počas sledovania. Príkladom takejto **systematickej chyby pri sledovaní respondentov**, konkrétne strata respondentov pri sledovaní, môže byť štúdia ELSPAC, o ktorej sme hovorili v predchádzajúcej kapitole 3 o dizajnoch štúdií. Pri štúdiu autorov Piler a kol. (2017) bola spolupráca s rodičmi zo začiatku relatívne dobrá. Do piateho roku veku dieťaťa sa im podarilo udržať 70% respondentov, ale po 11 rokoch už spolupracovala len polovica rodičov a po 15 rokoch už iba tretina. Od väčšiny respondentov sa nepodarilo získať informácie a s veľkou pravdepodobnosťou im chýbajú práve tí najdôležitejší. Iný príklad na ukážku systematickej chyby pri sledovaní respondentov je štúdia Clerencia-Sierra a kol. (2020) znázornená na obrázku číslo 1. Respondentov môžeme stratiť aj v dôsledku vyššej mortality a morbidoty, ktorá môže zredukovať vzorku a ostanú len tí „najzdravší“. Španielska štúdia porovnávala tých, ktorí zomreli vo veku 80-89 rokov, 90-99 rokov a vo veku 100 a viac rokov a zistila, že skupina storočných mala vo všeobecnosti nižšiu mieru morbidoty, vyžadovala menej zdravotnej starostlivosti a boli aj menej často hospitalizovaní oproti tým, ktorí boli vlastne o 10 či 20 rokov mladší. Na prvý pohľad by sa mohlo zdať, že vyšší vek súvisí s lepším zdravím, no pravda je skôr, že vyššieho veku sa dožijú len tí najzdravší. Tento prirodzený výber však predstavuje systematickú chybu vo výbere respondentov.

K systematickej chybe môže viesť aj odlišné sledovanie, diagnostika a hlásenie prípadov v skupine exponovaných a neexponovaných. Ako môže dôjsť k **systematickej chybe pri reportovaní prípadov a exponovaných** si ukážeme na nasledujúcom príklade. Predstavte si, že realizujete kontrolovanú prípadovú štúdiu a skúmate riziko venózneho tromboembolizmu u užívateľiek orálnej antikoncepcie. V nemocnici vyberiete ženy hospitalizované kvôli tejto diagnóze a ženy hospitalizované kvôli inému akútnemu ochoreniu alebo plánovanej operácii. Oboch skupín sa opýtate, či užívali hormonálnu antikoncepciu v uplynulom roku. Podarí sa Vám zozbierať 92 prípadov žien hospitalizovaných kvôli venóznemu tromboembolizmu a k nim 98 žien do skupiny kontrol. V skupine prípadov 72 žien prizná, že užívalo hormonálnu antikoncepciu a v prípade kontrol 28 žien prizná, že užívalo hormonálnu antikoncepciu. Medzi prípadmi teda máte 78% užívateľiek hormonálnej antikoncepcie, kým v prípade kontrol iba 26% užívateľiek. Po výpočte pomeru šancí mali užívateľky hormonálnej antikoncepcie v

tomto prípade 10 násobne vyššiu pravdepodobnosť trpieť venóznym tromboembolizmom v porovnaní so skupinou neužívateľiek. Je to ale skutočne tak? K systematickej chybe mohlo dôjsť v kritériách na prijatie pacientiek do nemocnice, čo viedlo k chybe v diagnostike a hlásení prípadov. Keďže existovalo podozrenie na súvislosť užívania antikoncepcie s výskytom tohto ochorenia, pravdepodobnosť hospitalizácie ženy, ktorá v anamnéze uviedla užívanie antikoncepcie a mala podozrivé symptómy, bola oveľa vyššia v porovnaní so ženou, ktorá mala rovnaké symptómy, ale neužívala antikoncepciu. V skupine prípadov sme pravdepodobne zachytili všetky užívatelky, lebo ich automaticky hospitalizovali, ale neužívatelky mohli pri rovnakých symptómoch ochorenia ostať nehospitalizované. Ak by sme ten istý výskum zopakovali v populácii, teda nie u hospitalizovaných žien, zachytili by sme pravdepodobne aj neužívatelky antikoncepcie v skupine prípadov a došli by sme k iným výsledkom.



Journal of
Clinical Medicine



► J Clin Med. 2020 May 21;9(5):1563. doi: [10.3390/jcm9051563](https://doi.org/10.3390/jcm9051563)

Do Centenarians Die Healthier than Younger Elders? A Comparative Epidemiological Study in Spain

[Mercedes Clerencia-Sierra](#)^{1,2,*}, [Ignatios Ioakeim-Skoufa](#)^{3,*}, [Beatriz Poblador-Plou](#)^{2,4}, [Francisca González-Rubio](#)^{1,2}, [Mercedes Aza-Pascual-Salcedo](#)^{1,2}, [Mónica Machón](#)^{2,5,6}, [Antonio Gimeno-Miguel](#)^{2,4,*}, [Alexandra Prados-Torres](#)^{2,4,*}

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#)

PMCID: PMC7291259 PMID: [32455809](https://pubmed.ncbi.nlm.nih.gov/32455809/)

Abstract

This study aims to describe the clinical course, drug use, and health services use characteristics during the last year of life of elders who die being centenarians and to identify key aspects differentiating them from elders who die at an earlier age, with a particular focus on sex differences. We conducted an observational, population-based study in the EpiChron Cohort (Aragón, Spain). The population was stratified by sex and into three age sub-populations (80–89, 90–99, and ≥ 100 years), and their characteristics were described and compared. Multimorbidity was the rule in our elders, affecting up to 3 in 4 centenarians and 9 in 10 octogenarians and nonagenarians. Polypharmacy was also observed in half of the centenarian population and in most of the younger elders. Risk factors for cardiovascular disease (i.e., hypertension, dyslipidaemia, diabetes), cerebrovascular disease and dementia were amongst the most common chronic conditions in all age groups, whereas the gastroprotective drugs and antithrombotic agents were the

Obrázok číslo 1: Príklad epidemiologickej štúdie zo Španielska.

Keď si to zosumarizujeme, k systematickej chybe pri výbere vzorky môže dôjsť pri *výbere kontrol* v kontrolovanej prípadovej štúdií, pri *oslovovaní respondentov*, ktorí z nejakého dôvodu súhlasia alebo nesúhlasia s účasťou vo výskume, pri *sledovaní respondentov*, ktorých z nejakého dôvodu nevieme dohľadať alebo sa kvôli chorobe či úmrtiu už štúdie nemôžu zúčastniť a pri odlišnom sledovaní, diagnostike a *hlásení prípadov v skupine exponovaných a neexponovaných*.

Týmto systematickým chybám *môžeme predísť* alebo *ich môžeme lepšie zhodnotiť*, ak používame rovnaké kritéria pri výbere prípadov a kontrol. Taktiež, ak sa snažíme o vysokú mieru účasti pri oslovovaní respondentov a zároveň sa snažíme minimalizovať stratu respondentov v priebehu štúdie. V neposlednom rade pomáha, keď sa snažíme používať rovnaké spôsoby diagnostikovania a reportovania prípadov v skupine exponovaných a neexponovaných a zbierame a analyzujeme údaje o respondentoch, ktorí odmietli účasť, a ktorých sme nemohli dohľadať alebo nemohli pokračovať v štúdií.

Iným typom systematickej chyby je **systematická chyba pri pozorovaní**. Je to chyba v meraní expozície alebo dôsledku, ktorej výsledkom je systematický rozdiel v kvalite, presnosti merania v porovnávaných skupinách alebo medzi administrátormi. Ak by ste na lúku poslali dvoch administrátorov a požiadali ich, aby spočítali čmeliaky, včely a iné opeľovače, pričom jeden by bol špecialista na hmyz a druhý laik, s veľkou pravdepodobnosťou by sa ich zistenia výrazne líšili. Opeľovačov je veľmi veľa druhov. Niektoré muchy imitujú včely, osy, niektoré čmeliaky sa podobajú na včely a naša najväčšia včela, drevárik fialový, pripomína pancierovú stíhačku a nie včelu. Botanik by vedel kde hľadať a zaznamenal by ich viac vo väčšej druhovej rôznorodosti, kým laik by možno prišiel s krásnymi fotkami lietajúceho hmyzu a v jeho tabuľke by boli včely medonosné, čmeliaky a iný hmyz, pričom v kategórii iný hmyz by sa ocitol aj drevárik fialový, či včela chlpačka.

K **systematickej chybe pri pozorovaní** môže dôjsť z mnohých dôvodov a môže k nej dôjsť na **strane meracieho nástroja, administrátora alebo respondenta**.

Na úrovni merania môže k **systematickej chybe pri pozorovaní** dôjsť spôsobom dopytovania či zisťovania, čo môže viesť ku skresleniu. Je rozdiel či sa dopytujeme uzavretými alebo otvorenými otázkami, či otázku formulujeme citlivejšie a nepriamo alebo priamo. Ak sa spôsob dopytovania systematicky líši v skupinách, môže to tiež viesť ku skresleniu. Ďalším zdrojom systematickej chyby môže byť spôsob, ako sa zbierajú údaje. V jednej skupine môžete nechať respondentov vyplňať dotazník, ale v druhej zvoliť čítanie otázok. Môžete chcieť porovnávať majoritnú populáciu, ktorá zvládne vyplniť dotazník samostatne a populáciu z vylúčených rómskych komunit, ktorá nemá skúsenosť s vyplňaním dotazníkov, a tak im otázky budete čítať. Môžete jednej skupine umožniť vyplniť dotazník online, ale druhej pre výpadky internetu doniesť papierové dotazníky. Výpovede respondentov sa budú líšiť v závislosti od toho, akým spôsobom ste v danej skupine zbierali údaje.

Ďalším príkladom zdroja systematickej chyby na úrovni merania je odlišná kalibrácia meracieho nástroja. V štúdií s názvom HepaMeta realizovanej na Slovensku, ktorá sa venovala rizikovým faktorom vírusovej hepatitídy B/C a metabolického syndrómu v skupine obyvateľstva žijúceho v rómskych osadách, sme počas preventívnych prehliadok potrebovali zmerať výšku, váhu a krvný tlak respondentov. Zistili sme však, že používané výškomery, váhy i tlakomery v ambulanciách sa dosť rôznia. Ak sme chceli mať všetkých respondentov zvážených porovnateľným spôsobom, museli sme na preventívne prehliadky doniesť vlastnú váhu, výškomer i tlakomer. Inak by sa nám mohlo stať, že niektoré skupiny respondentov by mali vyšší tlak alebo nižšiu váhu v dôsledku inej kalibrácie meracieho nástroja.

Tabuľka číslo 1 ilustruje rozdiely vo výpovediach respondentov pri rôznych typoch dopytovania. V poslednom stĺpci je skutočný počet užívateľov vybraných liekov, teda 100%. Ak sa respondentov dopytovali na užívanie konkrétnych liekov, iba 22-39% respondentov si na užívanie tohto lieku spomenulo a reportovalo ho. Ak sa v otázke spomínal dôvod užívania liekov, napr. brali ste v poslednom čase lieky na upokojenie, proti bolesti či na zníženie horúčky, tak si spomenulo 40-77% respondentov. Pri otvorenej otázke, napríklad „užívali ste nejaké lieky?“, si spomenulo najmenej respondentov – len 1-25% z nich.

N (%)	Otvorené otázky	Dôvod užívania	Názov lieku	Skutoční užívatelia
Valium	4 (13)	18 (58)	9 (29)	31 (100)
Darvon	5 (25)	8 (40)	7 (35)	20 (100)
Aspirin	2 (1)	146 (77)	41 (22)	189 (100)
Acetaminophen	-	58 (61)	37 (39)	95 (100)

Tabuľka 1: Rozdiely vo výpovediach respondentov pri rôznych typoch dopytovania.

K systematickej chybe pri pozorovaní môže dôjsť aj kvôli administrátorovi. Niekedy je to o schopnosti administrátora získať potrebné údaje. Spomeňte si na príklad špecialistu na hmyz a laika pri počítaní hmyzu na lúke. V jednej štúdií skúmali kvalitu života seniorov, kde časť respondentov spovedal lekár a časť psychologička. Pri porovnaní údajov sa ukázali rozdiely. Lekár, trénuvaný na skrining zdravotných problémov si ich viac všimol, lepšie sa na nich pýtal a zaznamenal ich viac. Psychologička, trénuvaná na mapovanie psychickej pohody či nepohody a sociálnych vzťahov, bola zase úspešnejšia v tejto oblasti. V tomto prípade by pomohlo mať viac administrátorov a lepšie ich zaškoliť, aby sa čo najmenej líšili v spôsobe dopytovania a vypracovať čo najpodrobnejší protokol administrovania, teda spôsobu ako sa pýtať a ako reagovať na otázky respondentov. Iným zdrojom chyby môže byť, ak administrátor vie, kto je prípad a kto neprípad, exponovaný alebo neexponovaný. Alebo ak administrátor vie, aké hypotézy sú skúmané, či je to dokonca samotný výskumník, ktorý vedome či nevedome túži byť vo svojom výskume úspešný. V takom prípade sa môže do administrácie vniesť tendencia vyhľadávať potvrdzujúce údaje a ignorovať tie ostatné. Napríklad, kašľajúcich respondentov sa sugestívnym spôsobom pýtať na fajčenie a u nekašľajúcich respondentov si nevšimáť, bagatelizovať či nerozvádzať ich výpovede o fajčení.

K systematickej chybe pri pozorovaní môže dôjsť aj na strane respondenta. Ak máte hodinky, ktoré sledujú kroky, možno ste na začiatku ich používania prešli viac, než by ste prešli inokedy, len preto, že vás sledovanie motivovalo. Výskumy, ktoré nemôžu skryť, že merajú výkon, sa musia vysporiadať so zmenou spávania respondenta, ktorý si uvedomuje, že je pozorovaný, meraný. V prípade krokomerov alebo akcelometrov sa vie, že tento výkyv – nadvýkon po niekoľkých dňoch vyprchá, takže sa s tým zvyknú vysporiadať tak, že do analýz vložia údaje až po stabilizácii výkonu, teda po niekoľkých dňoch. Respondenti môžu tiež odpovedať sociálne žiaducim spôsobom, teda tak, ako sa domnievajú, že sa od nich očakáva. To sa deje najmä pri otázkach, ktoré sú v konflikte s ich potrebou udržať si dobrú mienku v očiach ostatných. Čím si je istejší sociálnym statusom a čím menej je otázka v konflikte s žiaducim obrazom o ňom, tým otvorenejší bude vo výpovedi. Inú odpoveď na neveru dostanete medzi chlapmi v krčme, inú v kostole alebo pred manželkou. Okrem iného môže byť výpoveď respondenta skreslená z toho dôvodu, že si nevie spomenúť, nie je to pre neho dôležité, cíti sa vinný alebo z rôznych ďalších dôvodov.

PRAKTICKÁ ÚLOHA 1:

V kontrolovanej prípadovej štúdií skúmali vplyv užívania vybraných liekov na výskyt vrodenej chyby srdca u novorodenciakov. Prípadmi boli matky detí, ktoré sa narodili s vrodenuou chybou srdca v období rokov 1973-1975. Ako kontroly vybrali matky náhodne zvolených zdravých detí, narodených v rovnakom období a v rovnakej oblasti. Matiek sa pýtali na užívanie liekov počas tehotenstva, pričom údaje o užívaní liekov získali otvorenou otázkou: „Užívali ste nejaké predpísané alebo nepredpísané lieky počas tehotenstva?“

Ženy užívajúce liek proti nevoľnosti mali 1,8-násobne vyššiu pravdepodobnosť, že porodili dieťa s vrodenuou srdcovou chybou. Otázky boli položené matkám detí s vrodenuou chybou aj matkám zdravých detí. Položená bola otvorená otázka, aké lieky brali počas tehotenstva.

Mohlo dôjsť k skresleniu výpovede matiek? Ak áno, ktorá skupina matiek mohla mať tendenciu skresliť výpoveď o užívaní liekov a akým spôsobom?

RIEŠENIE PRAKTICKEJ ÚLOHY 1:

Je veľmi pravdepodobné, že výpoveď matiek (prípadoov) bude podrobnejšia a presnejšia, lebo počas tehotenstva boli podrobnejšie sledované a v ich prežívaní sú tieto informácie dôležité, teda hodné zapamätania. Ak viete, že je s dieťaťom problém alebo sa narodí s problémom, hľadáte dôvod, pýtate sa prečo, a teda informácie o užívaní liekov budú patriť k tým, ktoré ostanú zapamätané. Môžete mať aj vyššiu ochotu vypovedať. Alebo naopak, môžete si ich pamätať, ale byť menej ochotná sa priznať-napríklad môžete zatajovať užívanie liekov alebo konzumáciu alkoholu z obavy, aby ste neboli obvinená z následkov zo strany rodiny, zdravotníckeho personálu alebo samých seba. Zatiaľ čo výpoveď matiek z kontrolnej skupiny, ktoré prežili tehotenstvo bezstarostne, bez problémov a tešia sa zo zdravého dieťaťa, môže byť menej podrobná a nemajú dôvod loviť v pamäti a trápiť sa nad dotazníkom. V tomto konkrétном prípade bola informácia o užívaní liekov získaná z otvorenej otázky a tá je veľmi citlivá na skreslenie výpovede. Pacienta sa opýtate: „Brali ste nejaké lieky?“ Odpoveď je nie. A tak sa dopytujete ďalej, či ani na bolesť hlavy, na tlak alebo nejaké vitamíny? A odpoveď je zrazu áno a vymenujú zoznam rôznych liekov, ktoré si dávajú k rannému čaju.

Štúdiu, ktorá skúmala vplyv Benedectínu, lieku proti nevoľnosti, neskôr zopakovali. Lepšie však monitorovali jeho užívanie, kontrolovali to v zdravotnej dokumentácii a vyvrátili akúkoľvek súvislosť medzi užívaním lieku a výskytom vrodenej srdcovej chyby u dieťaťa. Iným spôsobom, ktorý by sa v tomto prípade dal použiť, je vybrať kontrolnú skupinu žien zo skupiny žien, ktoré porodili dieťa s inou vrodenuou chybou alebo zdravotným postihnutím, teda také, ktoré by mali podobné prežívanie a postoj k skúmaným informáciám ako matky, ktoré porodili dieťa s vrodenuou chorobou srdca.

Ukázali sme si rôzne druhy systematických chýb pri pozorovaní a spôsobov ako im predísť. Systematickej chybe pri pozorovaní vo svojom výskumnom projekte môžeme predísť tak, že kontrolujeme kalibráciu meracích nástrojov, kontrolujeme charakteristiky administrátorov a analyzujeme ich vplyv na výpovede. Ďalej to môže byť tak, že používame viac administrátorov a prideliujeme im respondentov náhodne, dobre zaškoliujeme administrátorov, podrobne vypracujeme protokol administrovania, neinformujeme administrátorov o skúmaných hypotézach, lepšie formulujeme otázky a mnohé iné.

Pri hodnotení systematickej chyby v štúdií hodnotíme zdroj systematickej chyby, jej rozsah a silu, spôsob a smer skreslenia.

Skúsime si to ukázať na konkrétnom príklade. V zdravotníckom zariadení sa pýtali pracovníkov, ako často si umývajú ruky a sledovali výskyt nozokomiálnych nákaz. Analýza údajov odhalila, že čím nižšia bola frekvencia umývania rúk, tým vyšší bol výskyt nozokomiálnych nákaz. Máme však podozrenie, že tí, ktorí si ruky umývajú menej často sa skôr vyhli odpovedi alebo skreslili svoju odpoveď (uviedli viac umývaní rúk ako skutočne). Aké skreslenie ktorej premennej, ktorého výsledku a v akom smere môžeme očakávať?

Ak by sa nám podarilo zapojiť do štúdie aj tých zdravotníckych pracovníkov, ktorí skutočne zanedbávali hygienu rúk, a ak by sme získali úprimnejšie odpovede, počet nozokomiálnych nákaz by sa nezmenil. Tento údaj bol totiž získaný zo zdravotných záznamov. Došlo by však k zmene v rozdelení respondentov: niektorí pracovníci, pôvodne zaradení do skupiny „čistých rúk“, by boli presunutí do skupiny „špinavých rúk“. Zároveň by k skupine „špinavých rúk“ pribudli ďalší respondenti, ktorí sa predtým vyhli odpovedi.

V prípade, že pôvodné skreslenie maskovalo skutočný vzťah medzi zanedbaním hygieny rúk a výskytom nozokomiálnych nákaz, môžeme očakávať, že pri presnejšom zbere údajov by sa tento vzťah ukázal ako ešte silnejší. To znamená, že skreslenie pravdepodobne viedlo k podhodnoteniu výsledkov.

Opačný scenár – nadhodnotenie vzťahu – by mohol nastať iba v prípade, že zdravotnícki pracovníci so špinavými rukami, ktorí sa nezapojili do štúdie alebo zatajili svoje správanie, by zároveň neboli zodpovední za šírenie nákaz. Mohlo by to byť napríklad preto, že sa z nejakého dôvodu nedotýkajú pacientov ani povrchov. Tento scenár je však málo pravdepodobný.

OTÁZKY

1. Týka sa nasledujúce tvrdenie vonkajšej alebo vnútornej validity?

	vonkajšia validita	vnútorná validita
Skúma, či je vzťah medzi expozíciou a následkom pravdivý.		
Skúma, či sa zistenia z danej štúdie dajú zovšeobecniť aj na respondentov, ktorí sa štúdie nezúčastnili, iný kontext, alebo inú dobu.		
Hodnotí vplyv náhodnej chyby, systematickej chyby a mätúcich premenných.		

2. Priradte správne, čoho sa týka nasledujúce tvrdenie: Pozorovaná súvislosť medzi expozíciou a dôsledkom je náhodná, teda rozdiely medzi skupinou exponovaných a neexponovaných nie sú štatisticky významné.
- A) Náhodná chyba
 - B) Systematická chyba
 - C) Vplyv mätúcich premenných
3. Priradte správne, čoho sa týka nasledujúce tvrdenie: Chyba v dizajne alebo realizácii štúdie, ktorá vedie k nepravdivému pozorovaniu súvislosti medzi expozíciou a dôsledkom.
- A) Náhodná chyba
 - B) Systematická chyba
 - C) Vplyv mätúcich premenných
4. Priradte správne, čoho sa týka nasledujúce tvrdenie: Vzťah medzi expozíciou a sledovaným parametrom je skreslený ich súvislosťou s treťou premennou (moderátorom alebo mediátorom).
- A) Náhodná chyba
 - B) Systematická chyba
 - C) Vplyv mätúcich premenných
5. Priradte správne, čoho sa týka nasledujúce tvrdenie: Pozorujeme vzťah medzi vzdialenosťou od školy a častotou respiračných ochorení. Potvrdíme aj súvislosť medzi vzdialenosťou a používaním MHD a tiež medzi častotou používania MHD a výskytom respiračných ochorení. Vzťah medzi vzdialenosťou od školy a výskytom respiračných ochorení je teda nepravdivý.
- A) Náhodná chyba
 - B) Systematická chyba
 - C) Vplyv mätúcich premenných
6. Priradte správne čoho sa týka nasledujúce tvrdenie: Skúmame vzťah medzi vysielaním futbalového šampionátu a spotrebou piva. V skupine máme mužov i ženy. Muži počas sledovania vypijú 2 krát viac piva ako obvykle, ženy 2 krát menej. Analýza údajov ale nepotvrdí súvislosť medzi futbalom a pitím piva. O akú chybu môže ísť v tomto prípade?
- A) Náhodná chyba
 - B) Systematická chyba
 - C) Vplyv mätúcich premenných

7. Priradte správne, čoho sa týka nasledujúce tvrdenie: Skúmame rozdiely v počte krokov u ľudí, ktorí sa vozia do práce autom alebo MHD. Rozdáme im krokometry a nasledujúci deň si odpíšeme počet krokov, ktoré urobili. V ten deň si ale šoféri áut všimli, že majú málo krokov a pridali prechádzku. O akú chybu ide?
- A) Náhodná chyba
B) Systematická chyba
C) Vplyv mäťúcich premenných
8. Priradte správne, čoho sa týka nasledujúce tvrdenie: V skupine respondentov, ktorí konzumovali ovocie z ekologických fariem, sa chrípka v poslednej sezóne vyskytla o 5% menej často ako v kontrolnej skupine respondentov, ktorí konzumovali akékoľvek ovocie. Rozdiel medzi skupinami však nie je štatisticky významný.
- A) Náhodná chyba
B) Systematická chyba
C) Vplyv mäťúcich premenných
9. Priradte správne, čoho sa týka nasledujúce tvrdenie: Sledujeme súvislosť medzi veľkosťou sídla a výskytom úmrtí na podchladenie. Vo väčších sídlach je však vďaka neziskovej organizácii vyššia miera informovanosti, takže sa táto príčina uvádza do úmrtných listov, kým v menších sídlach sa v takých istých prípadoch uvádzajú rôzne iné dôvody úmrtia (napr. kardiovaskulárne zlyhanie). O akú chybu ide?
- A) Systematická chyba pri výbere kontrolnej skupiny
B) Systematická chyba pri oslovovaní respondentov
C) Systematická chyba pri sledovaní respondentov
D) Systematická chyba pri reportovaní prípadov a exponovaných
10. Priradte správne, čoho sa týka nasledujúce tvrdenie: Skúmame súvislosť striktného dodržiavania pokynov fyzioterapeuta na zmiernenie bolesti chrbta. Je však potrebné aspoň polročné cvičenie a sledovanie. Tak dlhú dobu sa nám podarí sledovať iba menej ako polovicu respondentov, a navyše tých, ktorí mali na začiatku len mierne bolesti. Respondenti s vážnejšími problémami to vzdali a skúšajú niečo iné. Respondenti, ktorí nedokázali striktno dodržiavať pokyny, to tiež vzdali. V štúdiu sa nám nepodarilo potvrdiť súvislosť. O akú chybu ide?
- A) Systematická chyba pri výbere kontrolnej skupiny
B) Systematická chyba pri oslovovaní respondentov
C) Systematická chyba pri sledovaní respondentov
D) Systematická chyba pri reportovaní prípadov a exponovaných

11. Môžeme takýmto spôsobom predísť systematickej chybe pri pozorovaní?

	áno	nie
Skontrolovať kalibráciu meracích nástrojov (či každú skupinu meriame rovnakým spôsobom).		
Používať otvorené otázky.		
Dobre zaškoliť administrátorov.		
Vypracovať protokol administrovania.		
Informovať administrátorov o skúmaných hypotézach.		
Neinformovať administrátorov o skúmaných hypotézach.		
Prideľovať respondentov administrátorom náhodne.		

Prideľovať jednému administrátorovi len exponovaných a druhému len neexponovaných.		
--	--	--

12. Skúmaný bol vzťah medzi problémami v učení (mal si niekedy problém s písaním, čítaním, matematikou) a postojom k škole u žiakov. Žiaci, ktorí reportovali problémy, menej často reportovali, že sa im v škole páči. Avšak 30% žiakov nevyplnilo dotazník. S veľkou pravdepodobnosťou práve tí, ktorí majú problém s učením, najmä s čítaním. Ak by sa ich podarilo zapojiť do štúdie, akým smerom a nakoľko by to zmenilo zistenia? Teda, aký je smer skreslenia? Označte správnu odpoveď.

- A) Zistenia by sa nezmenili.
- B) Zistenia sú pravdepodobne podhodnotené a ak by sa zapojili všetci, vzťah medzi problémami v učení a negatívnym postojom k škole by bol ešte výraznejší.
- C) Zistenia sú pravdepodobne nadhodnotené a ak by sa zapojili všetci, vzťah medzi problémami v učení a negatívnym postojom k škole by bol slabší alebo žiadny.

13. Môžeme takýmto spôsobom predísť systematickej chybe pri výbere vzorky?

	áno	nie
Zvoliť rovnaké kritéria pri výbere prípadov aj kontrol.		
Stimulovať vysokú účasť oslovených respondentov.		
Snažiť sa o čo najmenšiu stratu respondentov počas sledovania.		
Zvoliť rovnaké spôsoby diagnostikovania a reportovania.		
Priebežne zbierať a analyzovať údaje o respondentoch, ktorí odmietnu účasť alebo ich nevieme ďalej sledovať.		
Zozbierať si všetky údaje sám, vlastnoručne.		
Vyberať si iba respondentov vyznačujúcich sa dobrou spoluprácou.		
Vybrať do súboru iba aktívnych dobrovoľníkov.		

Zoznam použitej literatúry:

- Aschengrau, A., & Seage, G. R. (2008). *Essentials of epidemiology in public health*. Jones & Bartlett Publishers. *Sudbury, MA*.
- Clerencia-Sierra, M., Ioakeim-Skoufa, I., Poblador-Plou, B., González-Rubio, F., Aza-Pascual-Salcedo, M., Machón, M., ... & Prados-Torres, A. (2020). Do centenarians die healthier than younger elders? A comparative epidemiological study in Spain. *Journal of clinical medicine, 9*(5), 1563.
- Ferjenčík, J. (2010). *Úvod do metodologie psychologického výzkumu*. 2. vyd. *Praha: Portál*. ISBN978-80-7367-815-9.
- Kido, R., Shibagaki, Y., Iwadoh, K., Nakajima, I., Fuchinoue, S., Fujita, T., & Teraoka, S. (2009). How do living kidney donors develop end-stage renal disease?. *American Journal of Transplantation, 9*(11), 2514-2519.
- Lötvall, J., Ekerljung, L., Rönmark, E. P., Wennergren, G., Lindén, A., Rönmark, E., ... & Lundbäck, B. (2009). West Sweden Asthma Study: prevalence trends over the last 18 years argues no recent increase in asthma. *Respiratory research, 10*, 1-11.
- Parikh, C. R., & Concato, J. (2006). Study designs in patient-oriented research. *American journal of kidney diseases, 47*(2), 356-364.
- Piler, P., Kandrnal, V., Kukla, L., Andrášková, L., Švancara, J., Jarkovský, J., ... & Klánová, J. (2017). Cohort profile: the european longitudinal study of pregnancy and childhood (ELSPAC) in the Czech Republic. *International Journal of Epidemiology, 46*(5), 1379-1379f.
- Rosenberger, J., Veseliny, E., Bena, L., & Roland, R. (2005 a). A renal transplant patient with intractable hiccups and review of the literature. *Transplant infectious disease, 7*(2), 86-88.
- Rosenberger, J., Geckova, A. M., van Dijk, J. P., Nagyova, I., Roland, R., van den Heuvel, W. J., & Groothoff, J. W. (2005 b). Prevalence and characteristics of noncompliant behaviour and its risk factors in kidney transplant recipients. *Transplant international, 18*(9), 1072-1078.
- Rosicova, K., Madarasova Geckova, A., van Dijk, J. P., Kollarova, J., Rosic, M., & Groothoff, J. W. (2011). Regional socioeconomic indicators and ethnicity as predictors of regional infant mortality rate in Slovakia. *International journal of public health, 56*, 523-531.
- van Nieuwenhuijzen, M., Junger, M., Velderman, M. K., Wiefferink, K. H., Paulussen, T. W., Hox, J., & Reijneveld, S. A. (2009). Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Preventive medicine, 48*(6), 572-578.
- Vahtera, J., Westerlund, H., Hall, M., Sjösten, N., Kivimäki, M., Salo, P., ... & Zins, M. (2009). Effect of retirement on sleep disturbances: the GAZEL prospective cohort study. *Sleep, 32*(11), 1459-1466.

Úvod do metodológie výskumu

Vysokoškolský učebný text

Autorka:

Mgr. Jaroslava Kopčáková, PhD

Vydavateľ: Univerzita Pavla Jozefa Šafárika v Košiciach

Vydavateľstvo ŠafárikPress

Počet strán: 70

Rozsah: 4,1 AH

Vydanie: prvé

ISBN 978-80-574-0371-5 (e-publikácia)



ISBN 978-80-574-0371-5



9 788057 403715