

JAK ČÍST VĚDECKÉ STUDIE

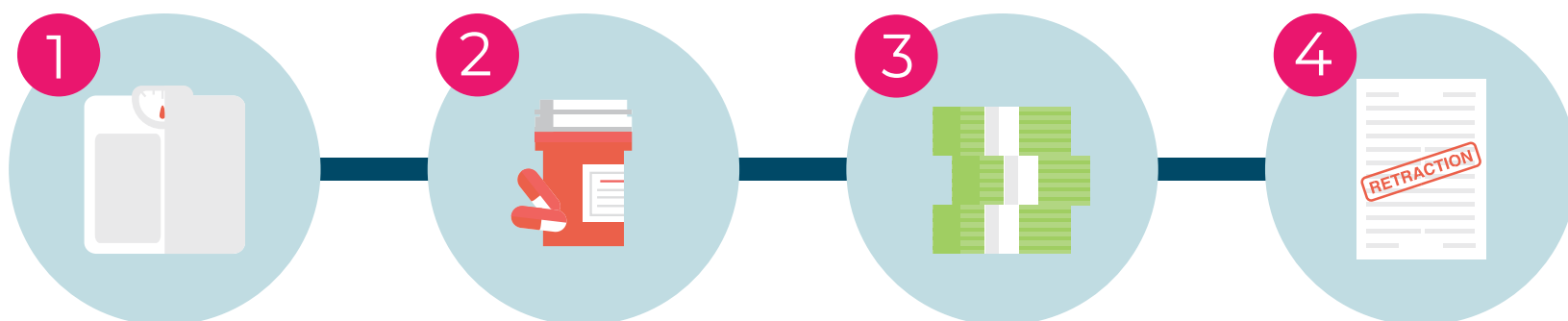
(a proč je to tak obtížné)



Proč se učit číst vědecké studie?

Marketingová tvrzení pronikají do zdravotního a fitness průmyslu zahalená pod rouškou „vědeckých důkazů“. Výrobci doplňků stravy prodávají suplementy jako např. [extrakt ze zelené kávy](#) (GCE, pro který je možné jen stěží najít jakýkoli vědecký výzkum na lidech) s podobnou jistotou, jako by byla jejich účinnost prokázána srovnatelně jako s [kreatinem](#) (u něhož existují stovky studií na lidech). Někdy, pokud se vydáme po stopě marketingového tvrzení, se můžeme dopátrat skutečně publikované studie – ne všechny studie jsou však rovnocenné. Abychom se vyhnuli plýtvání peněz nákupem neefektivních výrobků produktů a doplňků stravy, musíme být schopni posoudit různé aspekty studie, např. její důvěryhodnost, použitelnost a klinickou relevanci předkládaných výsledků.

Schéma č. 1: Extrakt ze zelené kávy – varovný příklad.



Studie v *The Journal of Diabetes, Metabolic Syndrome, and Obesity* tvrdí, že účastníci, kteří užívali GCE ztratili na hmotnosti přibližně 0,7 kg/měsíc.

Dr. Oz začal vychvalovat GCE jako „záračnou pilulku na hubnutí“.

V následujících letech GCE vydělává milióny dolarů.

Originální studie je stažena. Správnost dat studie není možné nezávisle ověřit.

Špatně provedené studie mohou vést k nevšedním výsledkům. Předtím, než vyvodíte nějaký závěr, je obvykle nejlepší cestou vyčkat a zjistit, zda je možné výsledky replikovat, než vyvodíte závěry.

Abychom porozuměli studii a její souvislosti s předchozím výzkumem na dané téma, je nutné číst více než abstrakt. Pokud je prezentován nový výzkum, kriticky důležitý je kontext studie. Samotné abstrakty jsou proto často zavádějící

Typy studií

Studií existuje celá řada. Tato příručka byla navržena tak, aby vám pomohla lépe porozumět každé jedné z nich se zvláštním důrazem na experimentální výzkum.

Tabulka č. 2: Přehled typů studií

TYPY STUDIÍ		DESIGN	PŘEDNOSTI	SLABINY
SOUHRN DŮKAZŮ	Meta-analýza	Shromažďuje data z veškeré dostupné literatury a analyzuje je	Poskytuje větší statistickou sílu dané výzkumné otázky	Časově velmi náročné, vyžaduje pokročilou znalost statistiky
	Systematický přehled	Představuje odborné review (přezkum) dostupných důkazů na dané téma	Poskytuje přehled-důkazy v oblastech s omezeným množstvím výzkumu	Rozdíly v designu mohou ztížit porovnání jednotlivých studií
INTERVENČNÍ STUDIE	Randomizovaná kontrolovaná studie (RCT)	Pacienti jsou náhodně rozdělováni do intervenční a kontrolní skupiny	Randomizace napomáhá eliminovat populační bias ve výzkumném souboru	Může být velmi drahá a nákladná na zdroje
	Nerandomizovaná kontrolovaná studie	Pacienti jsou rozdělováni do intervenční nebo kontrolní skupiny	Může být pro účastníky studie zaslepena	Chybí randomizace
OBSERVAČNÍ STUDIE	Kohortová studie	Sleduje návyky a rizikové faktory výzkumného souboru populace v průběhu času	Je jednodušejší proveditelná než RCT	Její provedení může trvat roky
	Studie případů a kontrol	Porovnává anamnestické údaje skupin pacientů se specifickou nemocí či bez ní nebo zdravotním důsledkem	Pomáhá identifikovat potenciální rizikové faktory	Bývá často ovlivněný biasem odpovědí (paměť účastníků)
	Kazuistika	Poskytuje podrobný popis jednotlivých případů	Pomáhá identifikovat nové trendy	Není zobecnitelná

Randomizované, dvojitě zaslepené, placebem kontrolované studie jsou běžně považovány za zlatý standard biomedicínského výzkumu. V takových studiích jsou pacienti náhodně rozdělováni do intervenční (ta, která intervenci obdrží) nebo kontrolní skupiny (která dostává placebo), přičemž ani pacienti ani samotní výzkumníci nevědí, do které skupiny ten který pacient patří.



Porozumění abstraktu a úvodu

Článek je rozdělován do sekcí (oddílů)-tyto oddíly se mohou u jednotlivých článků lišit, zpravidla však obsahují abstrakt, úvod, metodologickou sekci (která poskytuje demografické údaje, prezentuje design studie a může vysvětlovat použití vybraných cílových parametrů) a závěr (ten je obvykle rozdělen na „výsledky“ a „diskusi“).

Abstrakt je krátkou sumarizací pokrývající hlavní body studie. Protože rozsah informací, které studie obsahuje, nelze jednoduše obsáhnout do několika vět, může být abstrakt neúmyslně zavádějící. Abstrakt neposkytuje kontext, nevysvětluje limitace experimentu a ani to, jak jsou publikované výsledky použitelné v klinické praxi (reálném světě). Před samotnou citací studie jako vědeckého důkazu v diskusi si nejprve **přečtěte celý článek**, neboť po přečtení se může ukázat, že studie poskytuje pouze slabý důkaz.

Úvod představuje zkoumanou problematiku. Měl by jasně identifikovat výzkumnou otázku, kterou se autoři studie snaží zodpovědět. Zde autoři obvykle sumarizují výsledky předchozích výzkumů a vysvětlují, z jakého důvodu se téma rozhodli dále zkoumat.

Ku příkladu užívání nekalorického sladidla [stévie](#) se ukázalo jako slibný způsob, jak zlepšit kontrolu [hladiny krevního cukru](#), zejména u diabetiků. Vědci se tedy rozhodli provést rozsáhlejší a přesnější studie, aby určili, zda by se stévie mohla stát účinnou léčbou diabetu. Úvod je často skvělým místem k nalezení dalších materiálů k danému tématu, protože zde se autoři často odkazují na předchozí relevantní výzkum.

Jedna studie je pouze dílkem skládačky

Čtení několika studií na dané téma poskytuje více informací – více dat – a to i přes neznalost vytváření vlastní meta-analýzy. Ku příkladu, pokud si přečtete pouze jedinou studii sledující efekt [kreatinu na hladinu testosteronu](#) s výsledkem jeho zvýšení, pak 100 % vašich dat říká, že kreatin zvyšuje hladinu testosteronu. Pokud však přečtete deset studií sledujících efekt kreatinu a najdete pouze jedinou studii nacházející efekt zvýšení, pak 90 % vašich dat říká, že kreatin hladinu testosteronu nezvyšuje.

Porozumění abstraktu a úvodu

(Toto je zjednodušený příklad, ve kterém jsme použili „sčítání důkazů“: porovnali jsme počet studií, které nachází efekt, a studií, které efekt nenachází. Meta-analýzy jsou však daleko komplikovanější: berou v úvahu různá kritéria, jako design studie, počet účastníků a zkreslení, která mohou ovlivňovat výsledky, a tedy nejen pozitivní či negativní výsledek studie.)

Není překvapivé, že společnosti vyrábějící doplňky stravy běžně studie účelově vybírají („cherry-picking“). Pokud např. chce taková společnost prodávat kreatin jako booster testosteronu, zmíní se o té jedné studii, která zjistila zvýšení hladiny testosteronu po suplementaci kreatinem, nikoli však o dalších devíti, kde efekt prokázán nebyl.

Obdobná situace nastává v případě, kdy se tábory na opačné straně spektra snaží prokázat názor, který zastávají, házením jednotlivých protikladných studií do pléna (proti sobě). Pokud budete hledat studii, která prokazuje, že strava s nízkým množstvím tuků je pro hubnutí efektivnější než dieta nízkosacharidová, důkazy podporující tento fakt zcela jistě najdete. Pokud však budete hledat studii s opačným efektem, najdete ji také. Při hledání pravdy je proto důležité (a nejen za účelem zajištění munice pro „rvačku“ na Twitteru) pečlivě projít celé spektrum důkazů a spravedlivě zvážit i studie, které nesouhlasí s vaším původním přesvědčením (pokud nějaké máte – většina z nás však má).

K této poznámce mějte na paměti, že společnosti nejsou v účelovém vybírání studií samy. Někdy mají tuto tendenci výzkumníci. Pokud víte, že spektrum důkazů v dané oblasti je sporné, a studie, kterou čtete, zmiňuje pouze výzkumy podporující závěry autorů, zvažte [vlastní vyhledání dalších publikací na toto téma](#) (což je dobrý nápad vždy).

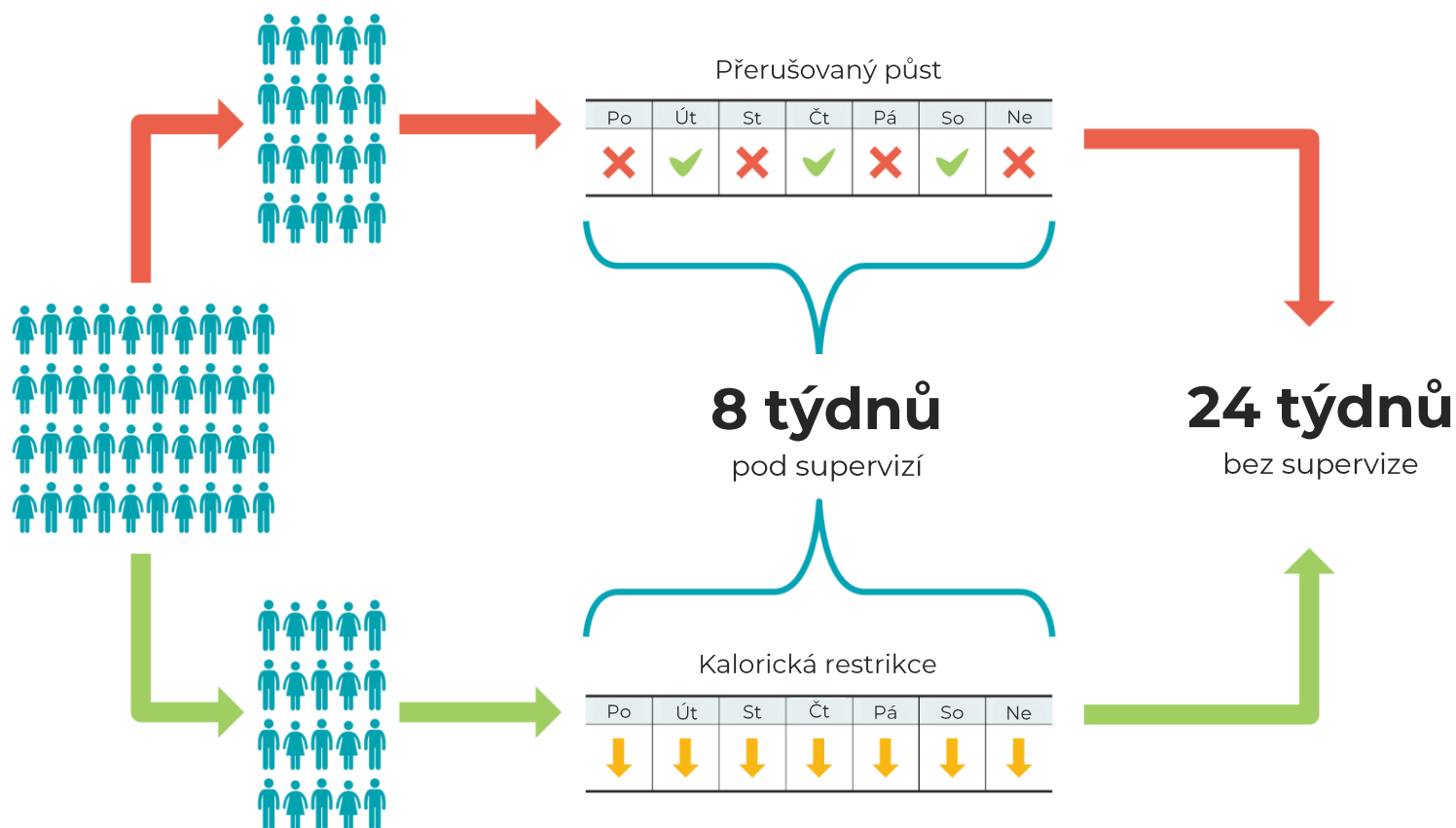
Metodologie: nejdůležitější část studie

Sekce „Metody“ (nebo také „Materiál a metody“) přináší informace o designu studie a jejích účastnících. V ideálním případě by měla být natolik jasná, přehledná a detailně popsaná, aby ostatní výzkumníci mohli studii zopakovat bez nutnosti kontaktovat autory. V této sekci bude třeba prozkoumat silné a slabé stránky výzkumu, které ovlivňují interpretaci výsledků studie.

Demografie

Sekce „Metody“ obvykle začíná popisem informací o účastnících, jako je věk, pohlaví, životní styl, zdravotní stav a způsob náboru. Tyto informace vám pomohou s rozhodnutím, jak je studie relevantní pro vás, pro vaše blízké nebo vaše klienty.

Schéma č. 3: Příklad protokolu studie porovnávající dvě různé diety.



Informace o demografických údajích studie mohou být zdlouhavé a vy budete pravděpodobně v pokusu přeskochit, je však třeba mít na paměti, že tyto údaje ovlivňují jak **spolehlivost**, tak i **použitelnost** studie.

Spolehlivost. Čím větší je studie (tzn. čím více má účastníků), tím spolehlivější jsou její výsledky. Povšimněte si, že studie často začíná s větším počtem účastníků, než končí. Zejména ve výživových studiích se běžně setkáváme se značným počtem „odpadlíků“ („dropouts“).

Použitelnost. Ve světě zdraví a fitness se použitelností rozumí, že látka nebo intervence (např. cvičení, dieta, suplement), které jsou použitelné pro jednoho, mohou být plýtváním peněz (nebo hůře, nebezpečím) pro druhé. Např. zatímco [kreatin](#) je široce uznáván jako bezpečný a účinný, existuje určitá část populace, která na efekt kreatinu neodpovídá („nonresponders“) a tento suplement u nich nezlepšuje výkon při fyzické aktivitě.

Jak je možné vidět na příkladu s kreatinem, i když použitelnost a rozsah výzkumu může být variabilní, demografické údaje napomáhají k posouzení jeho použitelnosti. Např. pokud jsou do studie [nabíráni pouze muži](#), ženy, které studii čtou, by měly mít na paměti, že výsledky mohou být pro ně daleko méně použitelné. Stejně tak výsledky intervence testované u vysokoškolských studentů se mohou lišit, pokud se provádějí na seniorech v domově důchodců.

Schéma č. 4: Některé studie jsou specifické pro pohlaví



Různé metody náboru účastníků budou navíc přitahovat různorodé demografické údaje a budou tedy ovlivňovat použitelnost studie. Ve většině případů budou výzkumníci používat nějakou formu „příhodného testování“. Např. univerzity provádějící studie budou velmi často nabírat vlastní studenty. Některé studie však budou využívat „nahodilé testování“, aby jejich výsledky studií byly více aplikovatelné na obecnou populaci. Takové studie jsou obecně nazývány „rozšířené randomizované kontrolované studie“.

Zavádějící (matoucí) faktory

Konečně, v demografických informacích budou běžně zmiňovaní účastníci vyřazení ze studie a pokud se tak stalo, pak z jakého důvodu. Nejčastějším důvodem je přítomnost zavádějících faktorů – proměnných, které by „zmátly“ (tzn. ovlivnily) výsledky.

Ku příkladu, pokud zkoumáte efekt odporového tréninkového programu na množství svalové hmoty, nebudete chtít, aby část participantů studie brala anabolické suplementy, zatímco ostatní ne. Bud' si budete přát, aby všichni užívali stejné doplňky, nebo spíše aby nikdo neužíval žádný.

Stejně tak, pokud zkoumáte efekt anabolických suplementů na množství svalové hmoty, nebudete chtít, aby jen někteří participanté studie cvičili, zatímco ostatní necvičí. Budete chtít, aby všichni dodržovali stejný tréninkový program, či méně pravděpodobně aby necvičil nikdo.

Je samozřejmě možné mít ve studii více než dvě skupiny. Můžete mít ku příkladu jednu studii sledující efekt odporového tréninkového programu s následujícími čtyřmi skupinami:

- Odporový tréninkový program + žádný suplement
- Odporový tréninkový program + kreatin
- Bez odporového cvičení + žádný suplement
- Bez odporového cvičení + kreatin

Avšak pokud vaše studie bude mít čtyři skupiny místo dvou, pro každou skupinu je nutné zachovat stejný počet participantů, je tedy nutné nabrat jich dvakrát více – což studii ztěžuje a prodražuje.

Pokud se podíváme hlouběji, všechny rozdíly mezi participanty jsou variabilní a jsou tedy potenciaálními matoucími faktory. Proto studie na myších vždy využívají druhy, které jsou si geneticky velmi podobné. A proto se také studie na lidech zřídka kdy provádějí na různorodém vzorku osob. Studie jsou např. omezené na starší ženy, aby se eliminovaly potenciální matoucí faktory pohlaví a věku.

Jak jsme mohli vidět výše, s rozsáhlejší velikostí studie je možné mít více skupin. Více skupin můžeme vytvořit dokonce po ukončení studie. Ku příkladu, pokud provádíte observační studii s tisíci osobami, která zkoumá efekt [červeného masa](#), můžete separovat data pouze pro „muže“ nebo „ženy“ a provést selektivní samostatnou analýzu pro tyto podskupiny. Analýzy podskupin takového charakteru jsou nicméně považovány spíše za průzkumné, nikoli průkazné, a mohou vést k falešné pozitivě výsledků. (Pokud např. krevní testy **mylně detekují** onemocnění, nazýváme takový výsledek jako **falešně pozitivní**.)

Design a cíle studie

V sekci „Metody“ bude také popsáno, jakým způsobem byla studie prováděna. Design studie zahrnuje např. jednostranně zaslepenou studii, kdy pouze pacienti nevědí, zda dostávají placebo; observační studie, ve kterých se pouze sledují demografické údaje a provádějí měření; a mnoho dalších (viz [tabulka č. 2](#) pro více příkladů).

Přesněji řečeno, dozvíte se zde informace o délce studie, podávaných dávkách léčiv, tréninku, testovaných metodách atd. V ideálním případě, jak již bylo zmíněno, informace zde popsané by měly být jasné a detailní do té míry, aby studie mohla být opakována bez potřeby kontaktování autorů.

Konečně, sekce „Metody“ může objasnit výzkumné cíle, na které se autoři zaměřovali. Např. studie zkoumající efekt odporového cvičení může využívat jako primární kritérium množství svalové hmoty (hlavní hodnotící kritérium studie)

Jedním z triků ve studiích, kde je žádoucí prokázat efekt (někdy proto, aby takový výsledek sloužil jako marketingový materiál pro nějaký produkt, ale často jednoduše proto, že studie, které prokazují efekt bude snadnější publikovat), je shromáždit větší množství výsledků (stanovit si více cílů) a poté sepsat článek pouze o takových výsledcích, které měly efekt, buď snížením významu ostatních cílových charakteristik, anebo tím, že jsou zcela vynechány. Za účelem prevence takového „datového dolování/lovení“ – „data dredging/fishing“ (metody, která byla prokázána úsměvným [čokoládovým hoaxem](#)) – mnoho vědců prosazuje [předběžnou registraci studií](#).

Mít čich na odhalování triků, které používají méně svědomití autoři, je bohužel dovedností, kterou si budete muset osvojit k objektivnímu posuzování publikovaných studií.

“
**Mít čich
na odhalování triků,
které používají
méně svědomití
autoři, je bohužel
dovedností, kterou
si budete muset
osvojit k objektivnímu
posuzování
publikovaných
studií.**”

a množství tukové hmoty, svalové síly a hladinu testosteronu jako kritérium sekundární.

Interpretace statistiky

Sekce „Metody“ obvykle končí obsáhlou diskusí o statistických metodách. Zjišťování, zda byla ve studii použita vhodná statistická analýza, je obsáhlé téma samo o sobě, proto doporučujeme nezabývat se detaily, ale soustředit se na celkový obraz.

Pojďme si primárně ujasnit dvě častá nedorozumění. Možná jste četli, že efekt byl **signifikantní**, a až později zjistili, že byl ve skutečnosti velmi malý. Stejně tak jste možná četli, že nebyl nalezen **žádný efekt**, ale při čtení článku jste zjistili, že intervenovaná skupina ztratila na hmotnosti více než skupina s placebem. Co to znamená?

Problém je jednoduchý: tihle podivní vědci nemluví stejným jazykem jako normální lidé.

Pro vědce **signifikantní** neznámá *významný* – ale *statisticky významný*. Výsledek je signifikantní, pokud by bylo nepravděpodobné, že data nasbíraná v průběhu studie skutečně neprokazují žádný efekt.

Proto může být i [velmi malý efekt](#) signifikantní – např. 0.2 kg (0.5 lb) ztráta hmotnosti za rok. Mimo to, efekt může být signifikantní, ale nikoli klinicky relevantní (to znamená, že nemusí mít žádné znatelné účinky na zdraví).

Obdobně, pro vědce, **žádný efekt** obvykle znamená *žádný statisticky signifikantní efekt*. Proto si můžete projít měření nasbíraná v průběhu studie a zaznamenat zvýšení nebo snížení, a přesto se v závěru dočíst, že nebyly nalezeny žádné změny. *Byly zde změny, které však nebyly signifikantní*. Jinými slovy, byly zde změny, ale tak malé, že *mohly* být způsobené nahodilými výkyvy (*mohou* být také důsledkem skutečného efektu; to však nevíme jistě).

[Jak jsme viděli výše](#) v sekci „Demografie“, větší rozsah studie poskytuje spolehlivější výsledky. Obdobně čím větší je rozsah studie, tím je větší možnost zjistit, zda jsou malé účinky signifikantní. Malá změna je s menší pravděpodobností způsobená nahodilými výkyvy, pokud se vyskytuje u studie s tisíci pacienty než u studie s pacienty deseti.

To vysvětluje, proč meta-analýzy mohou najít signifikantní změny sloučením několika studií, které nezávisle na sobě žádný signifikantní efekt nenašly.

P-hodnoty 101

Nejčastěji je efekt označen za signifikantní, pokud statistická analýza (prováděná výzkumníky po skončení studie) přináší hodnotu p , která není vyšší než stanovený práh (stanovený výzkumníky před začátkem studie). Takovému prahu říkáme *hladina spolehlivosti* (průkaznosti).

Pochopení, jak interpretovat p -hodnoty správně, může být složité, a to dokonce i pro experty, zde však předkládáme intuitivní způsob, jak o nich přemýšlet:

Představte si hod mincí. Hodíte mincí 100krát a dostanete výsledek zhruba 50/50 panna vs. orel. To by nebylo nijak překvapivé. Co když ovšem hodíte mincí 100krát a pokaždé padne panna? Teď je to překvapivé! Jen tak pro zajímavost, šance, že se něco takového skutečně stane, je přibližně 0,000000000000000000000000000008 %.

Můžete přemýšlet o p -hodnotách jako o hodu mincí vždy se stejným výsledkem – pana.

- P -hodnota 5 % ($p = 0,05$) není nic jiného než hodit v každém ze 4 pokusů pokaždé pannu.
- P -hodnota 0,5 % ($p = 0,005$) není nic jiného než hodit v každém ze 8 pokusů pokaždé pannu.
- P -hodnota 0,05 % ($p = 0,0005$) není nic jiného než hodit v každém ze 11 pokusů pokaždé pannu.

Na rozdíl od všeobecného přesvědčení, „ p “ v „ p -hodnotě“ neznamena „pravděpodobnost“. Pravděpodobnost výsledků hodu mincí o 4 pannách je 6,25 % nikoli 5 %. Pokud chcete převést hodnotu p na hody mincí (technicky nazývané [s-hodnoty](#)) a procentuální pravděpodobnost, podívejte se [sem](#).

[Jak bylo řečeno dříve](#), efekt je signifikantní, pokud by bylo nepravděpodobné, že data nashbíraná v průběhu studie skutečně nemají žádný význam. K tomuto faktu je možné doplnit, že čím je p -hodnota nižší, tím více si můžeme být jistí významností účinku.

P -hodnoty 201

Tak dobře. Férové varování: takhle sekce je pro šprty. No, lehce pro šprty. V klidu ji přeskočte a čtěte dále [zde](#).

Pořád jste tady? Dobře, tak se do toho pustíme. Jak jsme se již dozvěděli, výzkumníci provádějí statistickou analýzu výsledků jejich studií (obvykle se provádí jedna analýza k jednomu cílovému kritériu) za účelem průkazu, zda daná intervence měla či neměla efekt. Běžně jsou taková rozhodnutí prováděná na základě p -hodnot výsledku, které říkají, jak pravděpodobný by byl výsledek, alespoň tak velký jako ten, který byl pozorován ve studii, kdyby **nulová hypotéza**, mimo jiné předpoklady, byla pravdivá.

Ach, ten žargon! Nepanikařte, koncept vám vysvětlíme a ilustrujeme.

V každém experimentu jsou obecně dvě protichůdná stanoviska: **nulová hypotéza** a **alternativní hypotéza**. Pojdme si představit fiktivní studii sledující vliv doplňku stravy „Better Weight“ na hubnutí ve srovnání s placebem. Dvě protichůdné hypotézy by pak vypadaly takto:

- **Nulová hypotéza:** ve srovnání s placebem Better Weight **nezvyšuje** nebo **nesnižuje** tělesnou hmotnost (= hypotéza říká, že účinek doplňku stravy je **nulový**).
- **Alternativní hypotéza:** ve srovnání s placebem Better Weight **zvyšuje** nebo **snižuje** tělesnou hmotnost (= hypotéza říká, že doplněk stravy má pozitivní nebo negativní účinek na tělesnou hmotnost).

Účelem je vědět, zda účinek (zde na tělesnou hmotnost) intervence (zde doplňku stravy nazvaného „Better Weight“) je lepší, horší či stejný ve srovnání s kontrolní skupinou (zde placebo, někdy však může být kontrola jiná, např. dobře prostudovaná intervence, nové léčivo je zkoumáno ve srovnání s referenčním léčivem).

K tomuto účelu výzkumníci obvykle používají hladinu spolehlivosti (α) stanovenou před začátkem studie. Pokud je na konci studie p -hodnota (p) nižší nebo rovná hladině spolehlivosti ($p \leq \alpha$), můžeme říci, že zde existuje signifikantní rozdíl v účinku mezi léčebnými metodami. (Vzpomeňte si v tomto kontextu na [„signifikantní znamená statisticky signifikantní“](#).)

Schéma č. 5: Hladina spolehlivosti studie



Nejčastěji používanou hladinou spolehlivosti je 5 % ($\alpha = 0,05$). To znamená, že pokud **nulová hypotéza** (tzn. předpoklad, že **neexistuje** rozdíl mezi léčebnými metodami) je pravdivá, dostali by vědci po nekonečném počtu opakování pokusu falešně pozitivní výsledek (tzn. detekovali by signifikantní efekt, aniž by skutečně existoval) v nejvýše 5 % případů ($p \leq 0,05$).

Obecně je p -hodnota měřítkem konzistence mezi výsledky studie a předpokladem, že dvě léčebné metody mají stejný efekt. Pojdme se nyní podívat, jak se tyto poznatky promítnou ve studii Better Weight, ve které je jednou z léčebných metod doplněk stravy a druhou placebo:

- **Scénář 1:** p -hodnota je 0,80 ($p = 0,80$). Výsledek je tedy více konzistentní s **nulovou hypotézou** (tzn. předpokladem, že zde neexistuje rozdíl mezi léčebnými metodami). Shledáváme, že Better Weight **nemá žádný efekt** na hubnutí ve srovnání s placebem.
- **Scénář 2:** p -hodnota je 0,01 ($p = 0,01$). Výsledek je tedy více konzistentní s **alternativní hypotézou** (tzn. předpokladem, že zde existuje rozdíl mezi léčebnými metodami). Shledáváme, že Better Weight **má signifikantní efekt** na hubnutí ve srovnání s placebem.

Stejně jako je signifikantní výsledek $p = 0.01$, je signifikantní i $p = 0.000001$. Takže jaké informace nám vlastně nabízejí nižší hodnoty p ? Dávají nám větší důvěru ve výsledek. V našem případě by nám hodnota $p = 0.000001$ dala větší jistotu říci, že Better Weight má významný vliv na změnu tělesné hmotnosti.

Pamatujte si, že [signifikantní efekt nemusí být klinicky relevantní](#). Řekněme, že jsme našli signifikantní výsledek při hodnotě $p = 0,01$ prokazující, že Better Weight podporuje redukci tělesné hmotnosti. Chyták: Better Weight zapříčinil o 0,2 kg větší úbytek hmotnosti než placebo po jednom roce užívání – rozdíl je příliš malý na to, aby měl smysluplný efekt na zdravotní stav. V takovém případě, i když je výsledek signifikantní statisticky, v reálném světě (klinické praxi) je příliš malý na to, aby ospravedlnil doporučení pro suplementaci. (Takový typ scénáře bude pravděpodobnější u studií většího rozsahu, protože, [jak jsme již viděli](#), větší rozsah studie má vyšší potenciál najít malé, ale signifikantní účinky).

Konečně bychom měli zdůraznit, že přestože je nejčastěji využívána hladina významnosti je 5 %, některé studie vyžadují spolehlivost vyšší. Např. v případě genetických epidemiologů je pro průkaz statisticky signifikantní genetické asociace (řekněme průkazu, že určitý gen je asociován s nárůstem hmotnosti) běžně stanovena hladina významnosti na 0,0000005 % ($p \leq 0,0000005$), což je srovnatelné s 28násobným hodem mincí, kdy pokaždé padne panna. Pravděpodobnost, že se něco takového stane, je 0,00000003 %.

P-values: Neuctívejte je!

A konečně mějte na mysli, že i přes svou důležitost, **p -hodnoty nemají poslední slovo o tom, zda jsou závěry studie správné a adekvátní.**

[Viděli jsme](#), že vědci příliš dychtiví najít efekt ve své studii mohou sklouznout k „datovému dolování“ – „data dredgingu“. Mohou se také snažit snížit p -hodnoty různými cestami: např. mohou provádět různé statistické analýzy na stejných datech a dokumentovat pouze signifikantní p -hodnoty nebo mohou rekrutovat další participanty do té doby, než výsledek vyjde statisticky signifikantní. Tyto špatné vědecké praktiky jsou známé jako „ p -hacking“ nebo „selektivní reportování“. (O příkladu takové studie v reálném světě se můžete dočíst [zde](#)).

Metodologie: nejdůležitější část studie

Zatímco výběr statistické analýzy studie obvykle odpovídá proměnným, které se vědci snaží kontrolovat, p -hodnoty mohou být ovlivněny (úmyslně či neúmyslně) rovněž designem studie, skrytými matoucími faktory, typem použitých statistických testů, a mnoha, mnoha dalšími. Pokud hodnotíte sílu designu studie představte si sami sebe jako výzkumníka a zhodnoťte, jak byste dokázali překroutit studii do té míry, aby vám poskytla takové výsledky, které si přejete, a umožnila další kariérní postup.



Čtení výsledků

Závěrem studie diskutují vědci v sekci „Výsledky“ nebo také „Výsledky a diskuze“ primárně sledovaná data nebo data, která je nejvíce zajímala. Přeskočení k této sekci rovnou po přečtení abstraktu může být lákavé, často však vede k nesprávné interpretaci výsledků a šíření dezinformací. Nikdy nečtěte výsledky bez předchozího prostudování sekce „Metody“; **vědět, jak vědci dospěli k předkládaným výsledkům, je stejně důležité jako samotný závěr.**

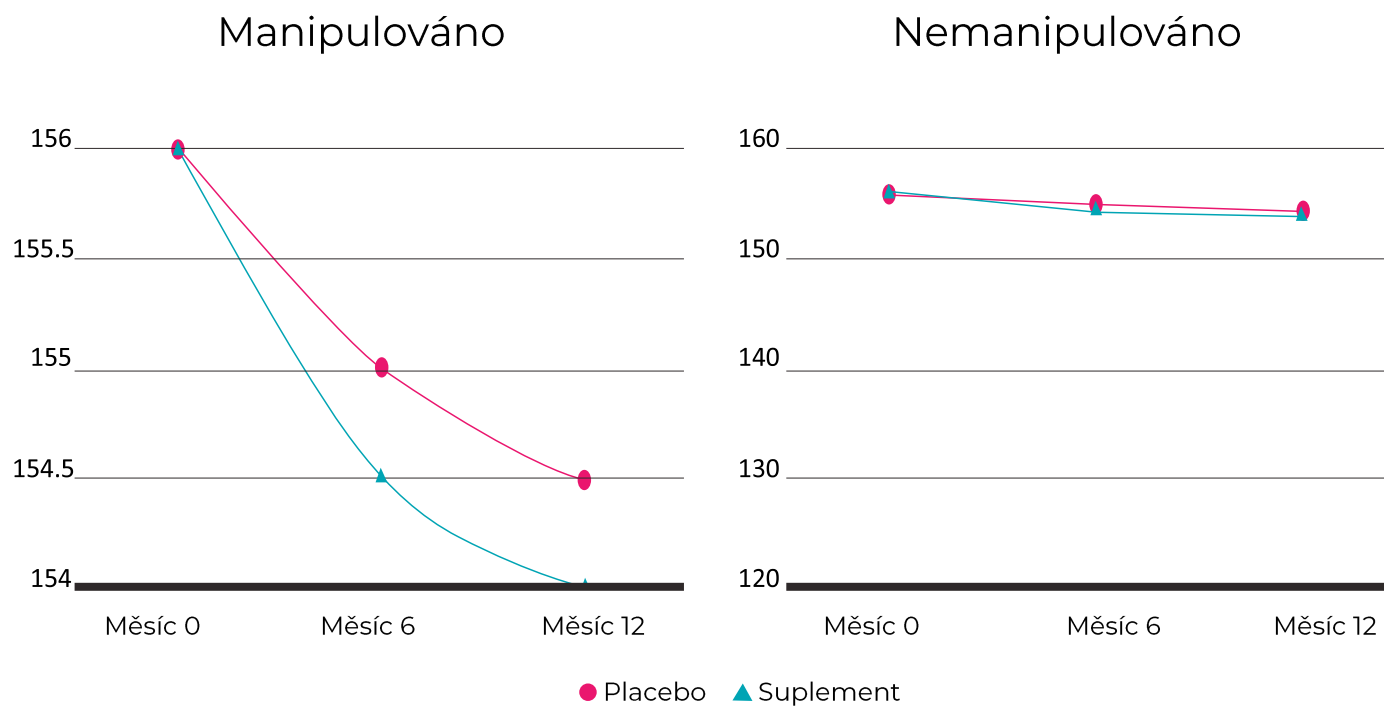
Jedna z prvních věcí, na kterou se v sekci „Výsledky“ zaměřit, je srovnání vlastností testovaných skupin. Velké rozdíly v počátečních („baseline“) charakteristikách po randomizaci mohou znamenat, že dané skupiny ve skutečnosti nejsou porovnatelné. Tyto rozdíly mohou být dílem náhody nebo způsobeny nesprávnou aplikací randomizační metody.

Výzkumníci taktéž musí informovat o počtu odpadlíků („dropout“) a hodnotit adherenci účastníků k terapii. Téměř každá studie má určitou část participantů, kteří studii nedokončí nebo se nejsou schopni řídit danými instrukcemi. To je případ především dlouhotrvajících nebo omezujících/náročných studií (např. experimenty s různými dietními pokusy). Při nápadně velkém počtu odpadlíků a vysoké nonadherenci účastníků k terapii bychom měli zpozornět, především pokud je míra předčasného ukončení v jedné skupině mnohem vyšší než ve zbývajících.

Vědci využívají dotazníky, krevní testy a další metody sběru dat. Všechna získaná data lze vyobrazit pomocí grafů a schémat. U grafů vždy sledujte, v jakých hodnotách (měřítku) na vertikální ose (osa Y) jsou výsledky zobrazeny. Co může na první pohled vypadat jako velká změna, může být ve skutečnosti změna pouze minimální.

V naší „Better Weight“ studii, zkoumající redukci hmotnosti, vedla intervence k rozdílu pouhých 0,2 kg (0,5 lb) mezi testovanou a kontrolní skupinou. Manipulací s osou Y však můžeme i z tak malého rozdílu udělat na první pohled rozdíl velký.

Tabulka 6: Manipulace osy Y (lb)



Sekce „Výsledky“ může zahrnovat také sekundární analýzy, jako **analýzu podskupin** nebo **analýzu citlivosti**.

Analýza podskupin. [Jak jsme mohli vidět](#) na konci sekce „Zavádějící faktory“, analýza podskupin spočívá v opětovném provedení analýzy, avšak pouze na omezené podmnožině participantů. Ku příkladu, pokud vaše studie zahrnuje muže i ženy všech věkových skupin, je možné analyzovat pouze data „žen“ nebo pouze data populace „starší 65 let“ a zjistit, jak a zda se výsledky liší.

Analýza citlivosti. Slouží ke kontrole, zda výsledky zůstanou stejné, i když je provedena odlišná analýza nebo se vynechají některá data (v meta-analýze lze např. vynechat jednu studii a provést analýzu znovu, bez ní).

[Jak jsme viděli](#) v sekci „Demografie“, spolehlivost studie záleží na velikosti testovaného vzorku. Pokud dojde k vyřazení některých účastníků studie z analýzy, zmenší se velikost testovaného vzorku a zvýší se riziko vzniku falešně pozitivních výsledků. To znamená, že pokud si s daty více "pohrajeme", je možné, že nakonec dosáhneme pozitivních výsledků.

Ukažme si to na extrémním příkladu: řekněme, že autor je placený, aby prokázal, že „Better Weight“ funguje. Testoval Better Weight na 20 participantech obou pohlaví ve věkovém rozmezí 21–87 let. Z 20 účastníků bohužel zredukoval hmotnost pouze 1. Je to 65letá žena. Autor se může rozhodnout provést analýzu podskupiny, kdy vyloučí všechny muže a všechny účastníky, kteří nejsou ve věku 65 let. Následně může tvrdit, že „Better Weight“ je účinný u žen ve věku 65 let.



Objasnění závěru

Někdy bývá závěr nahrazen a rozdělen do sekcí „Výsledky“ a „Diskuze“.

V sekci „Diskuze“ autoři vysvětlují hodnotu své práce. Mohou taktéž objasňovat svou interpretaci výsledků nebo polemizovat nad mechanismy účinku (tj. nad biochemií v pozadí). Daná studie je zde porovnávána s předchozími již provedenými studiemi a jsou zde uvedena doporučení pro další směr výzkumu, kterým by se měly následné studie na základě aktuálních výsledků ubírat. Je kriticky důležité zapamatovat si, že [jedna studie je pouze dílkem celkové skládky](#). Kam lze zařadit jednu konkrétní studii v celém spektru vědeckých důkazů na dané téma?

Autoři by měli uvést silné a slabé stránky své studie. Kriticky je zhodnotte. Popsali je autoři podrobně? Vynechali některé limitace studie? Nemusíte vycházet pouze z autorova hodnocení studie – analyzujte ji sami.

Stejně jako úvod poskytuje závěr cenný kontext a celkový vhled do tématu. Pokud v závěru autoři výsledky nadhodnocují nebo extrapolují výsledky nad daný demografický celek, nebojte se studii pročíst znovu (především sekci „Metody“).



Střet zájmů

Střet zájmů (SZ), pokud existuje, bývá běžně v textu uváděn na konci studie, po závěru. SZ se může objevit u osob, které se podílejí na návrhu, tvorbě, provedení a analýze výzkumu a mají zároveň motiv vyzkoumat určité konkrétní výsledky. Nejčastější důvod SZ tvoří finance – objevuje se v případě, kdy je výzkum sponzorován firmou nebo institucí, nebo pokud sám autor pro financující firmu pracuje, a ta tak může z výsledku studie sekundárně profitovat.

Smutným faktem zůstává (ověřeno [studii](#)), že neuvedení SZ je mezi autory běžné. Zároveň, co je považováno u jednoho časopisu za SZ, nemusí být stejně posuzováno i v časopise druhém. Někdy může být ve SZ dokonce celý časopis, přesto jej nemusí zveřejňovat. Například časopis pocházející ze země, která těží z exportu určitých bylin, může být skrytě motivován k publikaci studií, které prokazují pozitivní účinky těchto specifických bylin. Proto i když je studie zaměřená na určitou bylinu pouze obecně, tedy ne na konkrétní značku či produkt z této byliny, neznamená to, že zde není třeba předpokládat SZ.

SZ je nutné hodnotit opatrně – pokud ve studii není uveden žádný SZ, neznamená to, že zde SZ neexistuje. Zároveň, pokud SZ uveden je, není nutné automaticky předpokládat, že jsou výsledky výzkumu ovlivněny.

Hledání pravdy

Jak bylo možné vidět v sekci „Demografie“, výsledky studií jsou zřídka kdy paušálně aplikovatelné na každého. Například první studie zabývající se [glutaminem](#) byly prováděny na pacientech s popáleninami, kteří jsou na tuto aminokyselinu deficientní vlivem jejich zranění (popálenin). Následné studie se suplementací glutaminu na pacientech bez této deficience však stejný účinek neprokázaly.

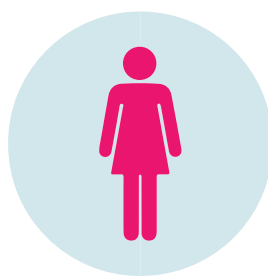
Schéma č. 7: Řada faktorů ovlivňujících použitelnost překládaných výsledků

DEMOGRAFICKÉ

VĚK POD 50



ŽENSKÉ POHLAVÍ

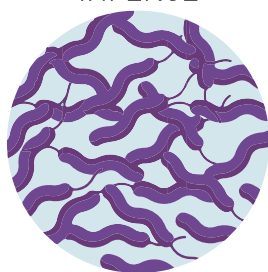


RODINNÁ ANAMNÉZA & GENETIKA



VLIVY PROSTŘEDÍ / BIOLOGICKÉ

GASTROINTESTINÁLNÍ
INFEKCE



DYSBIÓZA
STŘEVNÍ MIKROBIOTY



STRES, ÚZKOST, DEPRESE



Záměrná selekce určité (demografické) populace má své uplatnění pro výzkumníky, kteří hledají způsob, jak pomoci specifickému typu pacientů, na druhé straně může být takové jednání strategií k propagaci konkrétního výsledku – není např. s podivem a neobvyklé, že do studií se „spalovači tuku“ jsou běžně přijímány pouze postmenopauzální ženy s nadváhou. Pokud je tato informace vynechána z abstraktu studie a novináři pak přeskočí metodologickou sekci (nebo dokonce celou studii), dochází k publikaci zavádějících informací.

Nikdy nepředpokládejte, že média četla a zkoumala celé studie. Průzkum hodnotící [kvalitu důkazů v oblasti výživového poradenství publikovaný v britských národních novinách](#) poukázal na to, že 69 %–72 % tvrzení týkajících se zdraví založeno na nedostatečných nebo nevyhovujících vědeckých poznatcích. Novináři jsou často limitováni časem (uzávěrkou) a spoléhají pouze na zkrácené verze studií, které jen málokdy adekvátně sumarizují jejich výsledky.

Závěrem lze říct, že není náhrady za vlastní hodnocení studie, proto je v případě pochybností pro lepší posouzení silných a slabých stránek daného výzkumu nezbytné opakovaně prostudovat sekci „Metodologie“.



Proč je důležité mít tým

Posouzení, byť i jedné studie může být pracné. Zabere hodiny. Jsme si vědomi, že znát základy hodnocení studií je důležité, rozumíme však i lidem, kteří chtějí mít i vlastní život. Neexistuje člověk, který by měl čas přečíst všechny vycházející studie, a některé studie mohou těžit z toho, že jsou čteny různými profesionály z odlišných polí působnosti.

Se vzděláním v oblasti veřejného zdraví, sportovní medicíny, kinesiologie, výživy, farmakologie, toxikologie, mikrobiologie, molekulární biofyziky, biomedicíny, neurověd, chemie a dalších, jsou členové našeho týmu akreditovaní experti s různorodým pozadím a mohou si tak v průběhu revize výzkumu udělat komplexní obraz. Kromě toho má každý člen svou vlastní síť, na kterou se může obrátit, kdykoli potřebuje kontaktovat špičkové experty v daném tématu.

[Profesionálové](#), jejichž živobytí závisí na tom, aby měli aktuální a spolehlivé informace, důvěřují Examine.com, aby je informoval o nejnovějších výzkumech v oblasti výživy; důvěřují, že specialisté pečlivě prověří každou studii s maximální péčí a informace z ní poté sdělí jasně, stručně a přehledně. I pokud nejste zdravotník, můžete těžit z odbornosti Examine.com navštívením [webové stránky](#) (se stovkami volně dostupných článků) nebo přečtením jednoho u našich praktických průvodců z oblasti [fitness](#), [keto](#), [syrovátkového proteinu](#) a [dalších](#).



Základní checklist

Jelikož je celý průvodce značně obsáhlý, připravili jsme na konec zjednodušený seznam („checklist“), který bude po ruce, kdykoli se budete chtít ponořit do čtení další studie o výživě.

- ✓ **Jaká je hlavní hypotéza? (Na jakou otázku se autoři snaží odpovědět?)**
- ✓ **Je ve studii jasně a precizně popsán její design?**
 - O jaký [typ studie](#) se jedná?
 - Jak dlouho studie trvala?
 - Jaké cíle („endpoints“) byly sledovány primárně a jaké sekundárně?
- ✓ **Pokud se jednalo o klinickou studii, lze na základě informací v ní obsažených, tuto studii reprodukovat?**
 - Jednalo se o randomizovanou studii? Pokud ano, jak?
 - Byla studie zaslepená? Pokud ano, byla zaslepená jednou? Dvojitě? Trojitě?
 - Jaká léčba byla podávána? (Jsou dostatečně detailně popsány intervence v kontrolní a intervenované skupině? Co která skupina „dostala“ a co „nedostala“?)
- ✓ **Jaká populace byla studována?**
 - Jaký byl rozsah studie? (Kolik participantů se studie účastnilo?)
 - Byla jasně dána kritéria pro přijetí nebo vyřazení z účasti na studii?
 - Jakým způsobem byli participanté do studie nabíráni?
- ✓ **Jaké byly výsledky analýzy?**
 - Kolik lidí v každé skupině studii nedokončilo? („dropouts“)
 - Byly výsledky statisticky významné?

✓ **Jsou výsledky studie aplikovatelné v reálném světě?**

- Byly výsledky klinicky relevantní?
- Na jaké populaci byla studie prováděna; pro kterou je tedy pravděpodobně aplikovatelná?
- Byly by užívané dávky aplikovatelné v reálném životě (realistické)?

✓ **Co bylo zdrojem potenciálního biasu?**

- Byly nerovné hodnoty „dropouts“ mezi skupinami? Pokud ano, proč?
- Dodržovala intervenovaná skupina danou intervenci (léčbu)?
- Byla studie preregistrována k prevenci tzv. „data dredging“ (= hledání jakéhokoli možného vztahu nebo vlivu dané intervence)
- Byl deklarován nějaký konflikt zájmu?

Překlad zpracovala Česká asociace nutričních terapeutů, z.s.

