

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Název práce: Etnické/rasové rozdíly ve výskytu kolorektálního karcinomu v USA (Ethnic/race differences in the incidence of colorectal cancer in the USA)

Autor práce: Bc. Jana Slaměňíková

1. Obsah práce

Cíle, výzkumné otázky, hypotézy (*přiměřenost cílů, geografický charakter a relevance cílů, teoretické podložení hypotéz aj.*)

Prvním hlavním cílem práce je analyzovat vliv vybraných socio-demografických faktorů, faktorů zdravotního stavu a faktorů životního stylu na incidenci kolorektálního karcinomu. Druhým hlavním cílem je zjistit etnické rozdíly ve výskytu kolorektálního karcinomu v USA a stanovit příspěvek vlivu jednotlivých vybraných sociodemografických faktorů a faktorů životního stylu.

Cíle v této podobě jsou stanoveny vhodně a mají značný informační potenciál. V úvodní kapitole jsou navíc ještě definovány výzkumné hypotézy, a to na základě prostudované literatury (je snad ke zvážení, zda by hypotézy vč. jejich zdůvodnění a předpokladů neměly být ve struktuře práce zařazeny až za rešerši relevantní literatury) a obsahu převzatých dat. Na tomto místě není bohužel zcela jasné, co je myšleno obsahem převzatých dat a stanovením hypotéz na základě dat. Hypotéza stanovená pouze na základě dostupných dat by pravděpodobně postrádala teoretické podložení.

Autorka k jednotlivým hypotézám dodává stručné vysvětlení či doplnění založené na odborných publikacích. Přesto však u čtenáře mohou vznikat otázky ohledně volby a formulace hypotéz – např. u hypotézy 2 není zcela jasný její účel (potvrzení vyššího rizika vzniku kolorektálního karcinomu u afroamerické rasy) – hypotéza byla již dříve v textu uvedena jako konstatování faktu – str. 14 („Nejvíce zasažená rasa kolorektálním karcinomem je afroamerická, kde je nejvyšší rizikovitost oproti ostatním rasám.“ – toto konstatování však není v textu řádně citováno). Bylo by tedy vhodné stanovené hypotézy zdůvodnit i z hlediska jejich volby a významu.

Práce s literaturou (*teoretické ukotvení práce, diskuse a kritické zhodnocení literatury aj.*)

V teoretické části práce se autorka nejprve věnuje obecným charakteristikám kolorektálního karcinomu a gastrointestinální soustavy. Tato část nepochybně vytváří důležitý kontext, nicméně přehled klíčové odborné literatury týkající se etnických rozdílů v incidenci a dalších ukazatelích jsem očekávala především ve třetí kapitole. Najdeme v ní sice představení vybraných studií věnovaných tématu, autorka však částečně odbíhá k metodickým otázkám analýzy, ale také k popisu screeningových metod či léčby. Celkově dost částí této kapitoly se věnuje něčemu zcela jinému než slibuje její název.

Následující kapitola je věnována faktorům ovlivňujícím vznik kolorektálního karcinomu a je z hlediska práce s literaturou a celkového uspořádání a logické ucelenosti textu výrazně kvalitnější.

Autorka v práci využívá relevantní zdroje českého i zahraničního původu, množství využití literatury je v podstatě nadstandardní a dobře využito, citováno je důsledně (až na pár výjimek zmíněných

níže).

Metodika práce (vhodnost aplikovaných metod, kvalita a vhodnost použitých dat aj.)

Sice stále ještě v teoretické části práce, ale přesto na obr. 5 je uvedena standardizovaná míra, bylo by vhodné převzít nejen graf z citovaného zdroje, ale také např. informaci o tom, co je použito jako standard.

Popis využitých zdrojů dat a metod práce je obsahem šesté kapitoly. Data pochází z databáze Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO), což lze zcela samozřejmě považovat za zdroj naprosto odpovídající záměru práce. Volbu zdroje považuji nejen za vhodnou, ale i velmi užitečnou a ocenění hodnou. Za jeden z přínosů práce lze považovat právě i představení této databáze (slabinou tohoto představení však může být absence citací, která je jen na konci popisu, ačkoli ten zabírá cca 4 stránky a je dělen do více oddílů, resp. jde o takto dlouhý text přebraný z jediného zdroje, což navzdory účelu není z formálního hlediska optimální).

V kontextu popisu vzniku datového souboru pro vlastní analýzu však narážíme na nevysvětlenou „zvláštnost“ nebo spíše nejasný popis. Na str. 42 autorka uvádí, že celá databáze PLCO obsahuje údaje za cca 155 tis. účastníků. Na str. 46 autorka popisuje vlastní úpravy datového souboru, kdy uvádí vyřazení 31 případů s anamnézou kolorektálního karcinomu před vstupem do studie, dále uvádí, že ze souboru nebyla vypuštěna pozorování s chybějícími daty, nicméně navazuje větou říkající, že po této úpravě „...soubor obsahoval pouze jedince diagnostikované kolorektálním karcinomem během studie...“ – což by dávalo smysl tak, že patrně byli vyřazeni ti jedinci, kteří kolorektální karcinom diagnostikovaný neměli, nicméně autorka pokračuje „...[soubor] zahrnoval informace o 154 855 respondentech“. Toto je bohužel zcela matoucí – pokud na vstupu máme 155 tis. respondentů, téměř nikoho nevyřadíme, nemohou v souboru zůstat jen respondenti s diagnostikovaným karcinomem kolorekta, protože by to znamenalo, že tímto onemocněním onemocněli vlastně všichni respondenti. Bohužel to tak naznačují i uváděné počty – skutečně ze 155 tis. onemocnělo zkoumaným onemocněním 154 855 respondentů? Nebo již v původním souboru respondentů studie byli jen ti, kterým byl zkoumaný karcinom diagnostikován? To bohužel z předchozího popisu není zřejmé. Nicméně to opět naráží na fakt, jak na souboru jen respondentů s potvrzenou diagnózou může autorka studovat vliv protektivních a rizikových faktorů na riziko vzniku karcinomu, navíc s definovanou **závislou proměnnou jako „Zda má účastník potvrzen kolorektální karcinom vzniklý během studie?“ (srov. viz str. 46 „...soubor obsahoval pouze jedince diagnostikované kolorektálním karcinomem během studie...“** – s uvedenou citací ze str. 46 je následně v rozporu i konstatování ze str. 52 „Při srovnání osob bez kolorektálního karcinomu a s kolorektálním karcinomem jasně převažují osoby bez kolorektálního karcinomu s celkovým počtem 152 496 osob. Zatímco osob s kolorektálním karcinomem bylo pouze 2 359.“).

Na str. 46 bohužel narážím na další nejasnost „Tato vysvětlovaná (závislá) proměnná v rámci druhého hlavního modelu byla vybrána vždy pouze pro specifickou etnicitu. Pro popisnou statistiku a následující logistické regresní analýzy byly vytvořeny specifické modely pro jednotlivá etnika: ... aby bylo možné porovnat výsledné hodnoty.“ – zcela to opět není možné jednoznačně chápat, právě vytvoření samostatných modelů pro jednotlivá etnika v podstatě znemožňuje jejich vzájemné porovnání. Autorka tím patrně myslela možnost zhodnocení vlivů zkoumaných faktorů odděleně v rámci jednotlivých etnik, nikoli vzájemné porovnání modelů. V souvislosti s touto druhou částí

analýzy se však nabízí ještě další otázka vztahující se k tvrzení na str. 52 „Ve druhé části analýzy hodnotící vliv faktorů životního stylu na vznik kolorektálního karcinomu v rámci etnicit byly sestaveny specifické modely podle etnicity, kde byla každá proměnná testována zvlášť a kontrolována o věk a pohlaví.“ – proč byly modely sestavovány odděleně pro jednotlivé proměnné a nebyl (stejně jako v první části analýzy) sestaven souhrnný model, kde by vliv jednotlivých faktorů mohl být kontrolován na změny a hodnoty jiných proměnných?

Lze tedy konstatovat, že zvolená metoda, tedy logistická regresní analýza, je vhodná a zcela vyhovující stanovenému záměru práce. Bohužel z textu není zcela jasná konstrukce všech dílčích modelů a jejich zdůvodnění. Největší nejasnosti panují v popisu konstrukce a složení analyzovaného souboru. V textu postrádám informaci o tom, jak se zvolená metoda dokáže vypořádat s chybějícími pozorováními.

Analytická část práce (správnost a interpretace výsledků, detailnost, struktura a provázanost analýzy)

Samotná analýza, resp. její výsledky, je předmětem kapitoly 7. Opět bohužel v popisu výsledků narážíme minimálně na nejasná vyjádření – např. str. 57 „Naopak statisticky významná ($p < 0,001$) se prokázala míra konzumace alkoholu, nicméně OR (1,00) nezvyšuje ani nesnižuje riziko vzniku kolorektálního karcinomu.“ Vždyť právě použitý test ověřuje rozdílnost OR od hodnoty 1 (v tomto případě je problém v desetinných místech, kdyby v tabulce byl použit jednotný počet desetinných míst, což by tak také být mělo, byla by nejasnost celkem dost vyjasněna).

Bohužel text je v podstatě jen zcela popisné povahy a rekapituluje hodnoty z tabulek. Velkou část zabírají jednoduchá deskriptivní třídění – např. struktury respondentů z hlediska vybraných charakteristik. I u popisu těchto základních třídění můžeme však narazit na nejasnosti nebo příliš zkratkovitá a nepřesná vyjádření - někdy vedou až k tak bizarním tvrzením jako na str. 71 „Z popisné statistiky a grafu č. 4 lze vyčíst, že aktuální konzumenti alkoholu jsou nejčastěji běloši (55,9 %).“ – v grafu ani analýze nejde o to, jaké etnikum je nejčastější konzument alkoholu, ale uvedený graf porovnává podíl aktuálních konzumentů alkoholu mezi respondenty s diagnostikovaným karcinomem. Obdobně na str. 73 „Graf č. 7 znázorňuje zastoupení respondentů v jednotlivých kategoriích BMI v rámci etnicit.“ – v grafu není zastoupení kategorií BMI v rámci etnicit, ale pouze mezi respondenty s diagnostikovaným karcinomem dle etnicit.

Další z nepřesností je na str. 79 „Obecně u bělochů lze vidět signifikantní hodnoty u stejných proměnných jako u celkového zhodnocení, pravděpodobně z důvodu převahy bělochů ve studii, kdy vliv jednotlivých proměnných u všech etnicit celkem není kontrolován o proměnnou etnickou příslušnost.“ – není to „pravděpodobně“ z tohoto důvodu, je to téměř jednoznačně. Především ale, je zde na místě dotaz, o jakém celkovém modelu autorka mluví, když konstatuje, že není kontrolován na etnickou příslušnost – celkový model (tab. 4) etnickou příslušnost obsahuje.

Vzhledem k popisnosti 7. kapitoly lze uvítat kapitolu osmou, která slibuje diskusi výsledků a stanovených hypotéz. Zde se autorka vrací k vymezeným hypotézám, poskytuje jejich vyhodnocení na základě dosažených výsledků a zároveň je konfrontuje se závěry jiných studií. I zde však nacházíme ne zcela jasná vyjádření nebo podložené úvahy, např. str. 84 „Jako logické vysvětlení vyššího rizika u bývalých kuřáků může být, že během mnoha let kouření se vliv škodlivých látek v těle kumuluje a rakovina vzniká až s určitým časovým odstupem...“ (jak by to vysvětlovalo větší vliv kategorie bývalý kuřák než aktuální kuřák?) „... Zároveň do kategorie aktuálního kuřáka jsou

zahrnutý osoby, které sice aktuálně kouří, avšak i ve velmi malém množství...“ (i bývalí kuřáci mohli kouřit jen malé množství cigaret) „... Rovněž kategorie bývalých kuřáků nám nepodává žádnou bližší informaci o množství vykouřených cigaret nebo časovém období, kdy dotyčný přestal kouřit (před 10 lety nebo třeba před týdnem).“ – oceňuji, že autorka se snaží nalézt možné zdůvodnění dosažených výsledků, nicméně při zcela korektním postupu analýzy by předpoklady výsledků a jejich význam měly být známy před samotným výpočtem. V tomto konkrétním případě, ale dalo by se to připomenout u více částí analýzy, je třeba vnímat také to, že logistická regrese potvrzuje souvislost mezi proměnnými, nikoli kauzální vztah – nenabízí se tedy např. v tomto případě vysvětlení, že výsledek nemusí znamenat, že zanechání kouření vede k diagnostice nemoci, ale že vztah může být opačný – v důsledku diagnostiky nemoci nebo zdravotních problémů před samotnou diagnostikou je možné, že část respondentů zanechala kouření? Nebo šlo o status bývalého kuřáka na počátku studie, tedy před diagnostikou?

Celkově lze konstatovat, že zvolený analytický postup je zvolen logicky a je v pořádku proveden. Interpretace je však někdy formulována ne zcela jasně, což je pravděpodobně dáno buď slabší stylistickou úrovní práce nebo nejistotou autorky v provedené analýze.

Závěry práce (odpovědi na výzkumné otázky, zasazení závěrů do širšího kontextu výzkumu aj.)

Autorka analýzou potvrdila vliv věku, pohlaví, etnicity, vzdělání, anamnézy kolorektálního karcinomu v rodině nebo diabetu na výskyt kolorektálního karcinomu mezi respondenty zpracovávané studie. Podařilo se prokázat spíše protektivní efekt vitamínu D a léků obsahujících léčebnou látku kyselinu acetylsalicylovou (aspirin). V některých případech nebyly však výsledky zcela průkazné a jednoznačné. Zastoupení etnických skupin nebylo vždy dostatečné pro provedení zamýšlené analýzy, případně bylo provedení analýzy zatíženo chybějícími pozorováními. Autorka dosažené výsledky zasadila do kontextu jiných prací, ne vždy je však jasné zdůvodnění případných rozdílů.

2. Formální náležitosti práce (zatrhněte Vaši volbu)

Formální úprava: **výborná** / velmi dobrá / dobrá / nevyhovující

Stylistická úroveň: **výborná** / velmi dobrá / **dobrá** / nevyhovující

Citační úroveň: **výborná** / velmi dobrá / dobrá / nevyhovující

Komentář (vč. kvality grafiky, (ne)přiměřenosti délky práce aj.)

Autorka má někde nešťastně formulované věty, kde např. ne vždy je dodržena shoda podmětu s přísudkem apod. V textu se vyskytují kromě toho ještě gramatické a stylistické nedostatky – str. 31 („vyšetření, které mohou vézt k časnému záchytu“).

V textu jsou zaměňovány individuální charakteristiky a charakteristiky populace – např. již v abstraktu „byl zjištěn vliv věkové struktury a pohlaví“ – není jasné, zda autorka hovoří o populační rovině (věková struktura) či individuální rovině (pohlaví).

Také některá tvrzení nemusí být pro čtenáře zcela jasná a pochopitelná – např.:

- „v posledních letech došlo k poklesu celkové míry incidence kolorektálního karcinomu ve Spojených státech... Tento pokles je zapříčiněn z velké části z důvodu screeningu.“ (str. 14) – z věty není jasné, zda se snížila účast na screeningu, tím i vykázaná incidence, nebo v čem

vlastně „důvod screeningu“ leží a co je tím myšleno.

- Na str. 36 se píše o škodlivém vlivu konzumace červeného masa na vznik karcinomu, autorka uvádí „Na rizikovitost vzniku kolorektálního karcinomu při konzumaci zpracovaného masa poukazuje mnoho studií zabývajících se specificky tímto problémem ... Riziko se odhaduje mezi 20 až 50 % ve srovnání s ne-konzumenty (Santarelli et al. 2008).“ – uvedené tvrzení by ale znamenalo výrazně protektivní efekt konzumace masa – snížení rizika na 20 až 50 % proti ne-konzumentům masa.

Některá uváděná tvrzení nejsou řádně odcitována příslušným zdrojem (např. str. 14 „Nejvíce zasažená rasa kolorektálním karcinomem je afroamerická, kde je nejvyšší rizikovitost oproti ostatním rasám.“).

3. Souhrn

Celkové zhodnocení práce (vč. přínosů práce, provázanosti teoretické a analytické části aj.)

Téma práce lze považovat za nepochybně aktuální a zásadní, může mít značný přínos k výzkumu tohoto tématu ve světě. Autorka si stanovila reálné a zajímavě pojaté cíle. Naformulované hypotézy však již někdy postrádají jednoznačné zdůvodnění – není např. jasné, zda je cílem ověřit dříve potvrzené závěry na analyzovaném datovém souboru.

Teoretická část práce představuje důležitý kontext pro vlastní analýzu. Poměrně hodně nejasností panuje především v popisu konstrukce datového souboru a konstrukce konkrétních modelů v rámci zvolené metody. Zcela přesvědčivě nepůsobí ani interpretace výsledků, samotné závěry se však zdají být logické a v mnoha případech přináší jednoznačně zajímavé poznání.

Předloženou práci **doporučuji / nedoporučuji** k obhajobě. (zatrhněte Vaši volbu)

Otázky k obhajobě

1. Proč specifické modely formulované pro jednotlivá etnika nebyla standardizovaná na zdravotní stav zkoumaných osob (viz str. 47: „Nezávisle proměnné lze rozdělit na socio-demografické ukazatele, ukazatele zdravotního stavu a ukazatele životního stylu. Tyto ukazatele vstupují do prvního modelu. Do specifických modelů druhého hlavního modelu vstupují pouze socio-demografické ukazatele a ukazatele životního stylu.“)?
2. Předložený text vyznívá z hlediska interpretace jako poměrně dost zobecňující na populační úroveň, nicméně stále jde o analýzu souboru, kde je jen několik tisíc potvrzených výskytů zkoumaného onemocnění. Otázka tedy míří na to, jak dalece lze dle autorky získané výsledky skutečně vnímat jako obecně platné pro úroveň populace?

Datum: 16. 1. 2021

Autor posudku: RNDr. Klára Hulíková, Ph.D.

Podpis: