

Posudek oponenta na diplomovou práci

oponentský posudek

Jméno posuzovatele: Šárka Kaňková

Datum: 3.9. 2018

Autor: Daniel Pokorný

Název práce:

Udržování RhD polymorfismu v lidské populaci selekcí ve prospěch heterozygotů

Cíle práce

Cílem této diplomové práce bylo najít doklady pro selekční výhodu RhD pozitivních heterozygotních jedinců.

Struktura (členění) práce, odpovídá požadovanému? ANO

Rozsah práce (počet stran): 138

Je uveden anglický abstrakt a klíčová slova ANO

Je uveden seznam zkratek? NE, nicméně předložená práce nevyžadovala samostatný seznam zkratek, případné zkratky jsou v textu vždy správně vysvětleny

Literární přehled:

Odpovídá tématu? ANO

Je napsán srozumitelně? většinou ANO, nicméně místy působí chaoticky

Použil(a) autor(ka) v rešerši relevantní údaje z literárních zdrojů? ANO

Jsou použité literární zdroje dostatečné a jsou v práci správně citovány? většinou ANO

Pouze citace „Kaňková et al. 2007; Kaňková and Flegr 2007“ použité hned v úvodu na str. 9 nejsou v seznamu použité literatury.

Dále bych si dovolila upozornit i na drobné odchylky v jednotnosti formátu citací, autor užívá zkratku „et al.“ a zároveň “et al“ bez tečky, v citacích o dvou autorech užívá spojení jednou „a“ a jindy „and“, nejednotné je i užívání čárek a středníků mezi jednotlivými citacemi, případně mezi autorem a rokem vydání.

Citace knih např. „Flegr J. Evoluční biologie. 2. vyd. *Academia*. 2009; Praha“ a “Potter, E. L. Rh... Its relation to congenital hemolytic disease and to intragroup transfusion reactions. *Year Book Publishers*; Chicago: 1947.“ postrádají čísla stránek, z nichž byl čerpán text v literárním přehledu.

Citace „Perry G. H. et al. Evolutionary genetics of the human Rh blood group system. *Hum Genet*. 2012; 131:1205–16.“ má odlišný formát než ostatní citace.

Materiál a metody:

Odpovídají použité metody experimentální kapitole? ANO

Kolik metod bylo použito? 3

Jsou metody srozumitelně popsány? NE

V metodice není popsáno, jak byly vypočteny skóry jednotlivých proměnných a jakých hodnot tyto skóry nabývaly. Sledované proměnné jsou nedostatečně popsány a není zdůvodněna jejich volba.

Experimentální část:

Je vysvětlen cíl experimentů? Částečně ANO

Je dokumentace výsledků dostačující? ANO

Postačuje množství experimentů k získání odpovědí na zadané otázky? ANO i NE

Cílem práce bylo najít doklady pro selekční výhodu RhD pozitivních heterozygotních jedinců.

Experimentů a statistických hodnocení bylo provedeno dostatek, nicméně autorem stanovený cíl je velmi obecný a navrhované hypotézy ani literární přehled čtenáře nijak nepřesvědčí o cílové volbě hodnocených proměnných použitých v této práci.

Diskuze:

Je opravdu diskuzí, nejde jen o konstatování vlastních výsledků? ANO

Jsou výsledky porovnávány s literaturou? ANO

Jsou uvedeny nějaké hypotézy či návrhy na další řešení problematiky? ANO

Závěry (Souhrn):

Jsou výstižné? ANO

Formální úroveň práce (obrazová dokumentace, grafika, text, jazyková úroveň):

Členění práce je logické, jednotlivé kapitoly na sebe jasně navazují.

Číslování kapitol je zbytečně podrobné a z mého osobního pohledu dělá práci více nepřehlednou. Např. stránky 35. a 36., zde je samostatná podkapitola pro každou proměnnou a pod nadpisem je pouze odkaz na přílohu.

Po jazykové stránce na mě práce působí trochu rozpačitě. Místy jsou přehledně a jazykově srozumitelně zpracované úseky textu. Zároveň se však v práci často vyskytují opakující se texty, příliš dlouhá souvětí a velmi těžce čitelné části, které dle mého názoru mohly být částečně způsobeny napojováním textu v různých časových úsecích a možná i nedostatkem času k závěrečné kontrole celého textu. Jako příklady bych uvedla vybrané výňatky z práce:

Str. 9 Úvod - opakované tvrzení

„RhD negativita a RhD heterozygotita často ovlivňují zdraví jedince v opačném směru. To silně naznačuje, že polymorfismus v RhD se může udržet v populacích negativně frekvenčně závislou selekcí, speciálně selekcí ve prospěch heterozygotů.

Předchozí studie ukázaly, že RhD negativita a heterozygotita často ovlivňuje zdraví v opačném směru. Polymorfismus v RhD genu proto může být udržován selekcí ve prospěch heterozygotů.“

Str. 10

„V této práci se provedeme analýzu souvislostí a RhD genotypu a fenotypu a infekce *T. gondii* s psychomotorickými schopnostmi, pracovní a krátkodobou pamětí, sebevědomím, intuicí a psychomotorickými výkony.“

Str. 13 - překlep

„Rh systém je komplexním a jedním z nejdůležitějších krvním systémem“

Str. 16 – dlouhé nesrozumitelné souvětí

„Nicméně existují i názory, které nedávají selekci ve prospěch heterozygotů význam pro udržování polymorfismu, opírající se o výsledků modelů ukazující, že biologická zdatnost inbredních jedinců oproti heterozygotům by musela být v podstatě nulová.“

Str. 28 – překlep

„.....nicméně pro RhD negativní jedinci stále představují podstatný problém“

Str. 35 – jazyková neobratnost

„Kognitivní výkon byl v této studii zjišťován pomocí cognitive reflection testu (CRT), což je standardizovaného test, jenž měří ...“

Str. 37

„Statistické analýzy dat psychomotorického výkony ovšem byly provedeny“

Str.117

„Výsledky CRT v případě velmi připomínaly obecný trend výsledků získaný i pro krátkodobou paměť.“

V práci se vyskytuje nejednotnost v pravopise termínů: RhD gen x RHD gen, heterozygosita x heterozygizita, toxoplazmóza x toxoplasmóza

Dokument obsahuje minimálně pravopisných chyb. Pouze na str. 35 „K tomuto byli využity uvedené extrémní hodnoty“

Grafy a tabulky jsou řazeny v textu plynule a obsahují číslování i správné popisky. V tomto případě bych ovšem raději volila číslování grafů a tabulek chronologické, tj. od 1 počínaje prvním grafem resp. tabulkou v práci (jedná se spíše o subjektivní názor). Autor užívá číslování vztahující se k číslům kapitol.

Splnění cílů práce a celkové hodnocení:

Předložená práce testuje hypotézu, že RhD polymorfismus je udržován selekcí ve prospěch heterozygotů související s opačným efektem genotypů na „výkon“, mimo jiné také v souvislosti s toxoplazmózou. Studium RhD polymorfismu u lidské populace je nepochybně jednou z rozvíjejících se oblastí studia evoluce člověka a má velký potenciál obohatit toto doposud málo prozkoumané vědecké pole o mnoho zajímavých výsledků.

V literární rešerši autor prokázal orientaci v současném stavu poznání a podal vyčerpávající informace úzce související se zadaným tématem. Vzhledem k tomu, že relevantních vědeckých prací není mnoho, autor často opakuje jejich výsledky (mnohdy velmi dopodrobna) a značná část literárního přehledu se tak následně překrývá s diskusí. Zde by bylo možná vhodnější literární přehled více promyslet a některé informace ponechat pouze v diskuzi. Naopak bych do literárního přehledu zařadila navíc podkapitulu týkající se studovaných proměnných a jejich vztahu k stanoveným cílům práce.

Značně nepovedenou částí hodnotím samotný Úvod, text je nejednotný co do formátu citací a formátu psaní odborných slov, obsahuje opakující se části. Citace v textu jsou vkládány minimálně a v místech, kde by měl být text podložen citací, citace chybí. Celý úvod na mě působí dojmem, že byl sepsán na poslední chvíli.

Jako zásadní nedostatek práce vidím velmi obecné a nejasné stanovení hypotéz a s tím související i volba pozorovaných proměnných (dále podrobněji zmíním v části otázky a připomínky oponenta).

Co se samotných výsledků týká, všechna data byla získána pouze z dotazníkových studií.

Určitě by bylo přínosné, kdyby student obohatil svou práci o vlastní, třeba jen malý, pokus, který by nebyl založen pouze na dotaznicích. Student by se tak mohl seznámit například s různými laboratorními technikami nebo metodikou organizace pokusů.

K statistickému vyhodnocování dat byly opakovaně použity obecné lineární modely.

Výsledky jsou přehledně zpracovány do tabulek a grafů a rozděleny zvlášť pro muže a ženy. Bylo by ovšem dobré si uvědomit, že obrazová demonstrace dat v podobě grafů by měla být použita, pokud přehledně dokreslí získané výsledky. V tomto případě jsou v práci některé grafy nadbytečné (např. všechny grafy týkající se vlivu věku na jednotlivé proměnné). Zde je dokonce nevhodně použit krabicový graf a veličina věk zde byla transformována do věkových kategorií, přestože ve všech statistických modelech je používána jako spojitá kovariáta. Místy si nejsem jistá, zda autor dobře rozumí všem statistickým modelům, které používá a následné interpretaci výsledků (dále podrobněji v otázkách oponenta).

Konstatování samotných výsledků v kapitole diskuze je z mého pohledu nadbytečné, velmi podrobné a opakující se, vlastní zamyšlení nad získanými výsledky by mohlo být obsáhlejší. Jak již bylo zmíněno, kapitola 6.6 Shrnutí diskuze se hodně překrývá s kapitolou 2. Literární

přehled.

Pozitivně hodnotím kapitolu 7. Omezení studie. Autor si je vědom nedostatků provedené studie. Přestože v celé práci užívá např. slovní spojení vliv interakce RhD fenotypu a toxoplazmózy na „XXX“, je si vědom, že typ připravené studie neumožňuje jednoznačně určit kauzalitu získaných výsledků. Otevřeně dále diskutuje pozitiva a negativa použitého vzorku osob z komunity „Pokusných králíků“ a volbu dotazníkových studií. V závěru navrhuje budoucí možnosti ověřování stanovených hypotéz.

I přes mnohé výtky, autor předložil práci, která bude jistě přínosným zdrojem mnoha dalších výzkumů. Prokázal schopnost pracovat se zahraniční literaturou a k testování získaných dat používat vhodné statistické analýzy. Prokázal schopnost získaná data interpretovat a zamyslet se nad přínosy a limity práce. Autor tak splnil požadavky kladené na diplomovou práci na PřF UK.

Otázky a připomínky oponenta:

Připomínky:

Str. 11 – pojem syndrom nevhodně užíván pro RhD systém

„Dějiny Rh systému se píší od roku 1937, ovšem jeho význam nebyl rozpoznán okamžitě, ale až ve 40. letech, kdy byl spojen s příčinou vážné hemolytické nemoci novorozenců (Levine a Stetson 1939). Po mnoho let bylo pozorováno, že tento syndrom komplikuje těhotenství.“

Str. 12 – BOX 2.1 – chybí alelická forma RHCE genu RHCE

„**Alelické formy** genu **RHCE** se označují **RHce**, **RHCE** a **RHcE** podle toho, jaký antigen kódují. „

Str. 13: chybí kombinace cE - „RhD protein nese D antigen a protein RhCE nese CE antigeny v různých kombinacích – ce, CE nebo Ce“

Str. 34 – místo FCT má být CFT, „.....pomocí komplement-fixační metody (FCT),..“

V osnově v bodě 6.2. jsou slova sebevědomí a intuice navíc.

Otázky:

1) Autor píše:

„Předkládaná práce testovala dvě hlavní hypotézy:

1. výkonnost RhD negativních osob je horší, než činí průměr v české populaci
2. výkonnost RhD pozitivních heterozygotních osob je lepší, než je průměr v české populaci“

Můžete mi prosím blíže definovat pojem „**výkonnost**“ použitý v hypotézách.

Termín „**průměr v české populaci**“ již dále v práci není použit ani statisticky hodnocen, co tímto pojmem bylo myšleno a jak měla být získána relevantní data o zmíněných průměrech české populace. Jakým způsobem byly tyto hypotézy testovány? Jaké konkrétní analýzy byly provedeny k ověření těchto hypotéz?

Dále autor v metodice uvádí:

„K testování formulovaných hypotéz bylo vybráno z použitého dotazníku pět proměnných, jejichž skóre bylo složeno z několika otázek či úkolů: **krátkodobá paměť**, **pracovní paměť**, **kognitivní výkon**, **sebevědomí a intuice**. K ověření daných hypotéz byla též použita sada úkolů, která testovala **psychomotorické výkony** respondentů.“

Na základě jakých kritérií byly zmíněné proměnné vybrány? Jak sebevědomí a intuice souvisí s pojmem výkonnost uvedeném v hypotézách?

„V této práci se provedeme analýzu souvislostí a RhD genotypu a fenotypu a infekce *T. gondii* s **psychomotorickými schopnostmi**, pracovní a krátkodobou pamětí, sebevědomím, intuicí a **psychomotorickými výkony**.“ Co konkrétně znamenají pojmy psychomotorické schopnosti a psychomotorické výkony v předložené práci? Psychomotorické schopnosti již dále uvedeny

nejdou.

2) V kapitolách 4.4.5. Sebevědomí a 4.4.6. Psychomotorická výkonnost na str. 36 jsou pouze odkazy na kapitolu 10. Příloha, nicméně tato příloha v předložené práci chybí a výše zmíněné proměnné tudíž vůbec nejsou popsány. Ani v žádné z dalších částí práce nelze vyčíst jak a co bylo měřeno v případě proměnných „sebevědomí“ a „psychomotorická výkonnost“.

Můžete mi prosím říci, jak byly tyto proměnné měřeny a hodnoceny?

3) Jak byly vypočteny skóry jednotlivých proměnných a jakých hodnot dosahovaly?

4) Celkově vyplnilo dotazník 11 411 respondentů. Z tohoto datového souboru byli vyřazeni cizinci (2 344 jedinců) z důvodu nejistoty přesnosti informací ohledně jejich RhD. Cizinci byli v datovém souboru rozpoznáváni jednak na základě lokalizování IP adresy a dále na základě jazyka použitého v odpovědích na otevřené otázky. **Z jakých důvodů u cizinců nebyla jistota RhD? Jakým způsobem toto čištění dat probíhalo? Jak jste kontroloval všech 11 411 respondentů v několika otevřených otázkách?**

Dále byla data očištěna o podezřelé hodnoty. K tomuto byli využity uvedené extrémní hodnoty v proměnných jako výška (méně než 140 cm a více než 220 cm bylo vyřazeno), váha (méně než 30 kg a více než 220 kg bylo vyřazeno, počet sourozenců (ti s odpovědí více než 15 byli vyřazeni) a rok narození (narození dříve než v roce 1930 a později než v roce 2005). **Na základě jakých kritérií byly stanoveny výše zmíněné extrémní hodnoty?**

5) Dotazníky byly vytvořeny v prostředí Qualtrics. Pracoval jste v tomto prostředí vy sám? Vytvářel jste dotazníky a následně spravoval experty dat?

6) V kapitole 10. Příloha jsou uvedeny otázky používané v dotazníku k získání proměnné krátkodobá paměť. Jak s touto proměnnou souvisí otázka 7: „MemSh7: Kolik \$ stála akcie General Motors 31. 12. 2012? Možnosti odpovědi: 18.32, 24.68, 28.83, 32,16, 36.64, 37.64“?

7) V kapitole Výsledky na str. 38 jsou 2 tabulky znázorňující rozložení probandů v příslušných kategoriích. Podle těchto tabulek je prevalence toxoplazmózy u žen přibližně 27 % a u mužů 17 %. Jak si tyto rozdíly vysvětlujete?

8) Grafy 5.1 Závislost skóre pracovní paměti žen na interakci RhD fenotypu s nákazou *T. gondii*. a 5.2 Závislost skóre pracovní paměti žen na RhD fenotypu (a podobně u všech dalších testovaných proměnných) zobrazují výsledky na odlišně velikých souborech, jak tuto skutečnost vysvětlíte? Tyto grafy pravděpodobně dokreslují výsledky z tabulky 5.3 Vliv interakce RhD fenotypu a infekce *T. gondii* na pracovní paměť studovaný pomocí GLM. Na jak velkém souboru byla statistika provedena? Jak jsou statisticky významná (hodnota p-value) data zobrazená v grafu 5.2?

S tím dále souvisí i další otázka. Jak vysvětlíte odlišné hodnoty p-value u RhD genotypu v tabulce 5.15. Vliv interakce RhD genotypu a infekce *T. gondii* na pracovní paměť studovaný pomocí GLM a tabulce 5.16 Vliv interakce RhD genotypu a věku na pracovní paměť studovaný pomocí mnohorozměrné analýzy GLM (podobně u všech dalších testovaných proměnných)?

9) Str. 115 – autor píše

„Úlohy zaměřené na pracovní paměť odhalily jak u žen, tak u mužů statisticky významný efekt interakce RhD pozitivního heterozygotního genotypu s přítomností nákazy *T. gondii*.“

Byla tato interakce testována??

10) V kapitole diskuze postrádám zamyšlení nad uměle vytvořenými skupiny v proměnné RhD genotyp. Smíšená skupina, tj. pozit. homozygot + heterozygot, mohla výrazně zkreslit získané výsledky. Proč jste tuto smíšenou skupinu do analýzy zahrnul a proč jste neprovedl pouze porovnání RhD negativních homozygotů se samotnými heterozygoty? Dokázal byste odhadnout (podle statistik pro ČR) v jakém podílu budou zastoupeni ve smíšené skupině pozitivní homozygoti a heterozygoti?

11) Prosím o vysvětlení následujících sdělení:

Str. 14 „Incidence HDN se významně změnila, když se zjistilo, že ABO inkompatibilita mezi matkou a plodem má částečný protektivní efekt proti imunizaci k D antigenu.“

Str. 36 „Po přípravě dat a očištění od testovacích či podezřelých dat (viz kapitola Vzorek respondentů) byla data – v případě psychomotorických výkonů – kontrolována na odlehlá pozorování.“

Str. 25 „Výsledky prokázaly, že RhD negativní toxoplasma-pozitivní jedinci vykazují horší výsledky ve výkonnostních testech, což ale neplatí vždy pro RhD pozitivní jedince.“ – než kdo?

Str. 23 „Probandi byli požádáni o vyvinutí maximálního úsilí stisku rukojeti dominantní rukou v pozici s fixovaným trupem a rameny. Experimentátor stál mimo zorný uhel probanda, který měl za úkol držet podhmatem 5 kg závaží v napjatých předpažených rukách na úrovni ramen. Experimentátor měřil čas, po jaký proband dokázal v této pozici závaží udržet. Po zaznamenání výkonu proband opakoval pokus ještě dvakrát tou samou rukou a následně třikrát svou slabší rukou (Flegr et al 2018).“ – jak tento pokus probíhal, z popisu je to velmi nejasné.

Str. 17 „Například dosud nejsilnější odhalený efekt latentní toxoplasmózy – extrémní zvýšení váhy *T. gondii* infikovaných matek v 16. týdnu těhotenství – byl pozorován jen mezi RhD negativními matkami (Kaňková et al. 2010).“ – co je extrémní zvýšení?

Návrh hodnocení oponenta (známka nebude součástí zveřejněných informací)

výborně velmi dobře dobře nevyhověl(a)

Podpis oponenta: