

Vliv rodinné anamnézy, fyzické aktivity a digitálních technologií na stupně obezity

1. Úvod

Tato práce se zaměřuje na analýzu datového souboru mapujícího úroveň obezity u obyvatel Kolumbie, Peru a Mexika (Palechor & Manotas, 2019). Hlavním cílem práce je identifikovat a kvantifikovat klíčové determinanty, které nejvýznamněji přispívají k rozvoji nadváhy a obezity. K dosažení tohoto cíle využíváme multinomiální logistickou regresi, která nám umožní sledovat simultánní působení více prediktorů najednou. Pro účely této zprávy jsme zvolili 3 proměnné, které budeme v analýze využívat. Patří mezi ně fyzická aktivita, čas strávený využíváním technologií a rodinná anamnéza, respektive, zda se v rodinné anamnéze participanta objevuje nadváha či obezita.

2. Teoretické ukotvení

Obezita je v současné medicíně a sociologii definována jako multifaktoriální chronické onemocnění charakterizované nadměrným hromaděním tukové tkáně, které představuje významné riziko pro rozvoj kardiovaskulárních chorob, diabetes mellitus 2. typu a určitých typů neoplazií (WHO, 2021). Teoretické rámce, jako je socio-ekologický model, zdůrazňují, že rozvoj obezity není pouze výsledkem individuálního selhání energetické bilance, ale je hluboce ovlivněn tzv. „obezogenním prostředím“ (Swinburn a kol., 2011), které zahrnuje dostupnost vysoce energetických potravin, urbanistické řešení podporující sedavý způsob života a socioekonomické determinanty. Současný výzkum se proto odklání od jednoduchých vysvětlení a přiklání se k multikauzálním modelům, které zkoumají interakci mezi energetickou homeostázou těla a vnějšími faktory (Müller & Geisler, 2017). Právě vícerozměrná analýza dat o stravovacích návycích a fyzické kondici umožňuje identifikovat klíčové vzorce, které vedou k progresi různých úrovní obezity u populace ve vybraných zemích jižní Ameriky (Palechor & Manotas, 2019).

3. Charakteristika dat

Dataset obsahuje celkem 2 111 individuálních záznamů, což poskytuje dostatečně robustní základ pro vícerozměrné statistické modelování. Každý záznam je popsán 17 atributy. Ty zahrnují jak demografické údaje (věk, pohlaví), tak specifické metriky stravovacích návyků (počet jídel, konzumace vody, příjem zeleniny) a fyzické kondice (frekvence pohybu, čas u technologií). Klíčovým prvkem je klasifikační proměnná úroveň obezity, která je rozdělena do 7 kategorií: nedostatečná hmotnost, normální hmotnost, nadváha 1. a 2. stupně a obezita typu 1, 2 a 3. Data byla shromážděna za účelem identifikace rizikových faktorů a predikce stupně obezity na základě modifikovatelného chování jedinců v Latinské Americe.

4. Seznam proměnných zahrnutých do statistického modelu

- Závislá proměnná:
 - Obezita (BMI kategorie) – kategorická proměnná (nabývá 7 úrovní: 0-6),
- Nezávislé proměnné:
 - Rodinná anamnéza – kategorická proměnná (nominální)
 - Fyzická aktivita – PAF (= physical activity frequency) – spojitá proměnná (numerická)
 - Technologie – TUE (= time using technology) – spojitá (numerická)

5. Statistická analýza – multinomická logistická regrese

Nulovou hypotézu této analýzy jsme stanovili následovně:

- H0: Rodinná anamnéza obezity, míra fyzické aktivity a čas strávený u digitálních technologií nemají statisticky významný vliv na zařazení jedince do jednotlivých kategorií hmotnosti.

Před samotnou analýzou jsme si zvolili jako referenční kategorii proměnnou „normal weight“. Ačkoliv zdrojový dataset obsahuje 17 atributů, pro účely této vícerozměrné analýzy byly vybrány tři klíčové prediktory: fyzická aktivita, využívání technologií a přítomnost nadváhy či obezity v rodinné anamnéze probanda. Toto metodologické zúžení bylo provedeno za účelem zvýšení interpretační hodnoty modelu a eliminace potenciálního statistického šumu.

6. Interpretace výsledků

Následující analýza interpretuje výsledky multinomiální logistické regrese, která zkoumá vztah mezi vybranými faktory životního stylu a jednotlivými stupni tělesné hmotnosti. Hodnocení probíhá od celkového ověření statistické významnosti modelu až po detailní interpretaci vlivu konkrétních prediktorů na pravděpodobnost výskytu jednotlivých typů obezity.

6.1. Globální test modelu

Model byl nejprve testován jako celek (viz Tabulka 1). Významnost zvolených prediktorů byla ověřena pomocí Likelihood Ratio testu typu 1. Z výsledků vyplývá, že všechny sledované proměnné statisticky významně přispívají k predikční schopnosti modelu ($p < 0,001$). Dominantním faktorem je rodinná anamnéza ($\chi^2 = 673,24$), jejíž vliv na výslednou kategorii hmotnosti je v rámci modelu nejsilnější. Fyzická aktivita i čas u technologií rovněž představují vysoce signifikantní prediktory, ačkoliv jejich relativní přínos k vysvětlení variability je ve srovnání s rodinnou zátěží nižší.

Tabulka 1: Výsledky globálního testu vlivu prediktorů (Type 1 L-R Test)

Proměnná	Stupně volnosti (Df)	Log-Likelihood	Chí-kvadrát (χ^2)	p-hodnota
Rodinná anamnéza	6	-3764,20	673,24	<0,001
Fyzická aktivita	6	-3711,71	104,97	<0,001
Technologie	6	-3684,31	54,79	<0,001

Souhrnná tabulka výsledků (viz Tabulka 2) naznačuje, že všechny tři faktory jsou významné, ale rodinná anamnéza je dominantním prediktorem, který vysvětluje největší část rozdílů v hmotnosti respondentů. U proměnné rodinná anamnéza nebyla v souhrnném testu vyčíslena jednotná hodnota Waldovy statistiky. Tento jev, v odborné literatuře známý jako kvazi-kompletní separace dat, nastává v situacích, kdy prediktor (rodinná anamnéza) téměř dokonale předpovídá závislou proměnnou (stupeň obezity). V našem souboru to znamená, že u vyšších stupňů obezity se prakticky nevyskytují respondenti bez rodinné zátěže, což vede k matematické nestabilitě globálního odhadu, přestože dílčí vlivy na jednotlivé kategorie jsou extrémně vysoké a vysoce signifikantní ($p < 0,001$). Při detailnějším rozboru jednotlivých kategorií je patrné, že u všech stupňů nadváhy a obezity dosahuje rodinná anamnéza výrazně vyšších hodnot Waldovy statistiky než ostatní faktory. Rodinná anamnéza je tedy v modelu faktorem s nejkonzistentnějším a nejrobustnějším vlivem na zvýšení BMI.

Dalším významným prediktorem je dle našich výsledků také míra fyzické aktivity ($p < 0,001$). Detailní analýza koeficientů potvrzuje její protektivní roli, kdy se zvyšující se frekvencí pohybu klesá pravděpodobnost výskytu vyšších stupňů obezity.

Posledním sledovaným faktorem je čas strávený u digitálních technologií. I tento prediktor se ukázal jako statisticky významný ($p < 0,001$), což potvrzuje, že digitální návyky hrají v tělesné hmotnosti respondentů svou roli. Ve srovnání s rodinnou anamnézou a fyzickou aktivitou je však vliv technologií v rámci celého modelu nejslabší.

Tabulka 2: Výsledky multinomiální logistické regrese pro faktory obezity

Proměnná (prediktor)	Stupně volnosti (Df)	Waldova statistika	p-hodnota
Rodinná anamnéza	6	—*	—*
Fyzická aktivita	6	92,915	<0,001
Technologie	6	52,385	<0,001

* U proměnné rodinná anamnéza nebylo jednotné číslo pro celou proměnnou modelem vyčísleno z důvodu extrémně silné závislosti (separace dat) u vyšších hmotnostních kategorií.

6.2. Analýza vlivu jednotlivých prediktorů a interpretace poměrů šancí

Následně jsme provedli detailnější analýzu dat, jejíž výsledky se nachází v Tabulce 3. Analýza koeficientů B umožňuje identifikovat charakter a sílu vlivu jednotlivých faktorů u konkrétních hmotnostních kategorií ve srovnání s referenční skupinou respondentů s normální hmotností.

Klíčovým ukazatelem v této analýze je poměr šancí neboli Odds Ratio (OR). Hodnota $OR > 1$ indikuje zvýšení šance na výskyt dané hmotnostní kategorie (rizikový faktor), zatímco hodnota $OR < 1$ znamená snížení této šance (protektivní faktor). Například u proměnné „fyzická aktivita“, kde $OR = 0,39$, dochází k poklesu šance na rozvoj obezity 3. typu o 61 % ve srovnání s referenční skupinou. Naopak u rodinné anamnézy dosahují hodnoty OR u nejtěžších forem obezity extrémních hodnot, což potvrzuje její dominantní predikční roli v modelu.

Tabulka 3: Analýza koeficientů B

Kategorie	Prediktor	Koeficient (B)	Odds Ratio	Směrodatná chyba	Waldova statistika	p-hodnota
Nadváha 1. stupně	Rodinná anamnéza	0,86	2,36	0,18	23,00	<0,001
	Fyzická aktivita	-0,23	0,79	0,09	6,39	0,011
	Technologie	-0,26	0,77	0,13	3,98	0,046
Nadváha 2. stupně	Rodinná anamnéza	2,62	13,74	0,27	91,95	<0,001
	Fyzická aktivita	-0,42	0,66	0,10	18,44	<0,001
	Technologie	-0,03	0,97	0,13	0,05	0,829

Data a další informace o této zprávě jsou dostupné na adrese [https://dostal.vyzkum-
psychologie.cz/stat4?i=736](https://dostal.vyzkum-psychologie.cz/stat4?i=736).

Obezita typu 1	Rodinná anamnéza	3,81	45,15	0,40	89,93	<0,001
	Fyzická aktivita	-0,41	0,66	0,10	18,74	<0,001
	Technologie	-0,09	0,91	0,13	0,49	0,482
Obezita typu 2	Rodinná anamnéza	5,67	290,0 3	1,01	31,52	<0,001
	Fyzická aktivita	-0,39	0,68	0,10	14,98	<0,001
	Technologie	-0,59	0,55	0,14	17,99	<0,001
Obezita typu 3	Rodinná anamnéza	17,28	∞	—*	—*	—*
	Fyzická aktivita	-0,93	0,39	0,11	76,66	<0,001
	Technologie	-0,13	0,88	0,13	1,03	0,309
Podváha	Rodinná anamnéza	-0,37	0,69	0,17	4,69	0,030
	Fyzická aktivita	0,00	1,00	0,09	0,00	0,991
	Technologie	0,38	1,46	0,13	9,39	0,002

* U proměnné rodinná anamnéza v kategorii Obezita 3. typu nebyla Waldova statistika ani p-hodnota přímo vyčíslena z důvodu tzv. kompletní separace dat (všichni respondenti v této podskupině vykazují rodinnou zátěž, což vede k matematické nestabilitě výpočtu dílčího testu). Vzhledem k výši koeficientu ($B = 17,28$) a výsledkům souhrnného testu modelu ($p < 0,001$) je však tento vliv považován za vysoce statisticky významný. Referenční kategorií pro všechny výpočty je skupina respondentů s normální hmotností.

Rodinná anamnéza se ukázala jako nejsilnější rizikový faktor. U všech kategorií nadváhy a obezity (1. až 3. stupeň) nabývá koeficient B vysokých kladných hodnot, které navíc s rostoucí závažností obezity prudce stoupají (B od 0,86 u nadváhy až po extrémních 17,28 u obezity 3. stupně).

Fyzická aktivita vystupuje v modelu jako konzistentní ochranný (protektivní) faktor. U všech kategorií nadváhy a obezity jsou koeficienty záporné. Nejsilnější ochranný účinek pohybu je patrný právě u nejtěžšího stupně obezity ($B = -0,93$).

Vliv času stráveného u technologií je v modelu nejvíce diferencovaný. Zatímco u některých forem obezity je vztah spíše slabý nebo nesignifikantní (např. obezita 1. typu, $p = 0,482$), u kategorie podváhy představuje vysoký čas u technologií statisticky významný faktor mírně zvyšující šanci na výskyt v této kategorii.

6.3. Hodnocení úspěšnosti predikce

Ačkoliv celková úspěšnost klasifikace činí 28,7 %, je nutné tento výsledek interpretovat v kontextu vysokého počtu kategorií (7). Model vykazuje výrazně vyšší predikční schopnost, než jaké by bylo dosaženo náhodným rozdělením (14,3 %).

Data a další informace o této zprávě jsou dostupné na adrese <https://dostal.vyzkum-psychologie.cz/stat4?i=736>.

Klíčovým zjištěním je vysoká úspěšnost u kategorií s vysokým BMI (Obezita 3. typu – 57,7 %), což dokazuje, že zvolené faktory (rodinná anamnéza, fyzická aktivita a využívání technologií) jsou vysoce efektivními prediktory pro identifikaci těžkých forem obezity. Detailní výsledky jsou uvedeny v tabulce níže.

Tabulka 4: Klasifikační matice úspěšnosti predikce modelu

Hmotnostní kategorie	Pozorováno (N)	Klasifikováno správně	Úspěšnost (%)
Normální váha	287	62	21,6 %
Nadváha 1. stupně	290	0	0,0 %
Nadváha 2. stupně	290	0	0,0 %
Obezita 1. typu	351	114	32,5 %
Obezita 2. typu	297	148	49,8 %
Obezita 3. typu	324	187	57,7 %
Podváha	272	95	34,9 %
Celkem	2111	606	28,7 %

Na základě provedené multinomiální logistické regrese a výsledků Likelihood Ratio testu ($p < 0,001$) zamítáme nulovou hypotézu o neexistenci vztahu mezi sledovanými faktory a úrovní BMI. Přijímáme alternativní hypotézu (H_A), neboť bylo statisticky prokázáno, že rodinná anamnéza, frekvence fyzické aktivity a čas strávený u technologií jsou významnými prediktory výsledného stupně hmotnosti.

7. Diskuze

Výsledky analýzy potvrzují, že obezita je multifaktoriální jev s dominantním vlivem dědičnosti. Extrémní hodnota koeficientu u rodinné anamnézy ($B = 17,28$ u Obezity 3. typu) a poměr šancí u Obezity 2. typu ($OR = 290$) naznačují, že rodinná dispozice funguje v tomto souboru jako velmi silný prediktor nejtěžších forem obezity. Tento vztah je natolik silný, že u nejvyšší kategorie obezity vedl až k matematické separaci dat (nekonečný poměr šancí).

Klíčovým zjištěním je protektivní vliv fyzické aktivity. Pravidelný pohyb dokáže dle naší analýzy snížit šanci na rozvoj nejtěžší obezity o 61 % ($OR = 0,39$). To potvrzuje, že fyzická aktivita je zásadním modifikovatelným faktorem prevence.

Vliv digitálních technologií se jeví jako sekundární a statisticky nejslabší ($\chi^2 = 54,79$). Zatímco u nejtěžší obezity je jeho role nejednoznačná, u kategorie podváhy zvyšuje šanci na její výskyt o 46 %, což může reflektovat specifické vzorce sedavého chování u této skupiny.

Celková přesnost modelu (28,7 %) je vzhledem k jemnému rozlišení 7 váhových kategorií adekvátní. Model je dvakrát úspěšnější než náhodný odhad (14,3 %) a vykazuje vysokou stabilitu v predikci extrémních stavů (Obezita 3. typu – 57,7 %). Nulová úspěšnost u mezistupňů nadváhy však naznačuje, že tyto kategorie jsou v reálných datech velmi prostupné a pravděpodobně ovlivněné dalšími faktory, jako jsou například stravovací návyky, které model nezahrnoval.

Zdroje:

Müller, M. J., & Geisler, C. (2017). From the control of energy homeostasis to a multicausal model of obesity. *The Journal of Nutrition*, 147(7), 1220–1226. <https://doi.org/10.3945/jn.117.247957>

Palechor, F. M., & Manotas, A. de la H. (2019). Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. *Data in Brief*, 25, 104344. <https://doi.org/10.1016/j.dib.2019.104344>

Swinburn, B. A., Sacks, G., Hall, K. D., McPherson, K., Finegood, D. T., Moodie, M. L., & Gortmaker, S. L. (2011). The global obesity pandemic: Shaped by global driving forces and local environments. *The Lancet*, 378(9793), 804–814. [https://doi.org/10.1016/S0140-6736\(11\)60813-1](https://doi.org/10.1016/S0140-6736(11)60813-1)

World Health Organization. (2021, June 9). *Obesity and overweight*. <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>