

PREDIKTORY MOZGOVEJ (CMP) PRÍHODY – LOGISTICKÁ REGRESIA¹

Cievna mozgová príhoda (CMP) je známa ako druhá najčastejšia príčina smrti (1. miesto obsadzuje ochorenia srdca). Ročne postihuje mozgová príhoda viac ako 17 miliónov ľudí na celom svete a je spojená s obrovskými spoločenskými nákladmi. Každému šiestemu z nás cievna mozgová príhoda hrozí, a to bez ohľadu na náš vek (ročne postihuje viac než 80 000 ľudí mladších ako 20 rokov) (Bandi et al., 2020).

Zaujímavosťou je, že hoci incidencia cievnej mozgovovej príhody je výraznejšia medzi dospelými, odborníci sa v praxi stretávajú i s prípadmi mŕtvice u detí. Incidencia mozgovovej príhody je približne 1,2 až 13 prípadov na 100 000 detí mladších ako 18 rokov (Rawanduzy et al., 2023).

CMP, ľudovo povedané mŕtvica, je náhla porucha krvného obehu mozgu, ktorá vedie k nezvratnému poškodeniu mozgového tkaniva. Pozostáva z neurologického ochorenia, ktoré môže byť výsledkom ischémie alebo krvácania mozgových tepien a zvyčajne vedie k heterogénnym motorickým a kognitívnym poruchám, ktoré ohrozujú funkčnosť nášho organizmu. Včasná detekcia mozgovovej príhody je kľúčovým krokom pre efektívnu liečbu a môže mať v tomto procese veľkú hodnotu, nakoľko je spájaná s častou mortalitou a morbiditou (Sirsat et al., 2020).

Podľa mechanizmu vzniku sa delí a *ischemickú* a *hemoragickú* mozgovú príhodu. *Ischemická* vzniká následkom upchatia mozgovovej tepny, ktorá spôsobí zastavenie alebo zníženie prietoku krvi do mozgu. *Hemoragická* vzniká v dôsledku prasknutia mozgovovej tepny, ktorá spôsobí buď akútne zakrvácanie do mozgu alebo medzi mozgové obaly (Sirsat et al., 2020).

Medzi ovplyvniteľné rizikové faktory zaraďujeme vysoký krvný tlak (hypertenzia), najrôznejšie ochorenia srdca, chlopní alebo nepravidelný srdcový rytmus (arytmia), cukrovka (diabetes mellitus), zvýšená hladina cholesterolu (hyperlipidémia), obezita, fajčenie, nezdravá životospráva alebo konzumácia alkoholu. V súčasnosti sa lekári snažia o kompletnú anamnézu pacienta. Údaje by tak mohli dopomôcť k tomu, aby osoby zasiahnuté touto chorobou mohli rýchlejšie dokázať predísť nákladom, ktoré ochorenia so sebou prináša (Alanazi et al., 2021). Čím skorej sú príznaky rozpoznané, tým skorej môže byť zahájená účinná liečba.

Data:

V tejto správe sa budeme zaoberať niekoľkými údajmi, ktoré v sebe zahŕňajú prediktívny potenciál určiť riziko mŕtvice. Dáta s ktorými sme v tejto správe pracovali, vychádzajú zo štúdie Pathan et al. (2020), ktorá si kládla za cieľ nahrnúť novú techniku založenú na súbore dát, vychádzajúcich zo zdravotníckych anamnéz pacientov pre hodnotenie dôležitosti jednotlivých údajov v súvislosti detekovania mŕtvice. Pôvodne dáta obsahovali 10 nezávislých premenných. Pre účely tejto správy sme sa rozhodli ponechať 8 nezávislých premenných, ktoré popisujeme nižšie.

Výskumu sa zúčastnilo 4873 ľudí vo veku od 1 do 82 rokov, z toho 59 % žien a 41 % mužov. Súbor tvorilo skoro 12 % detí a 88 % dospelých. Dáta boli pre účely tejto správy mierne upravené.

¹ Táto správa vznikla v rámci predmetu Viacrozmerné štatistické metódy (UPOL, Katedra psychologie, 2023/2024).

Nezávislé premenné:

- **Pohlavie:** muž, žena
- **Vek:** vek pacientov
- **Hypertenzia:** 0 neprítomná, 1 prítomná
- **Kardiovaskulárne ochorenie:** 0 neprítomné, 1 prítomné
- **Typ zamestnania:** štátna správa; práca v súkromnom sektore; živnostník; dieťa
- **Ø hladina glukózy:** priemerná hladina glukózy v krvi
- **BMI – Body Mass Index**
- **Fajčenie:**
 - nikdy
 - v minulosti
 - fajčiar
 - nezistené

Závislá premenná:

Závislá premenná je v tomto prípade predstavuje odpoveď na otázku, aká je pravdepodobnosť výskytu mŕtvice na základe zozbieraných údajov. 1 kóduje výskyt, 0 kóduje jej neprítomnosť.

Logistická regresia

Tabuľka č. 1 Logistická regresia

Prediktory	Regresný koeficient	Šanca	Waldova štatistika	p-hodnota
Pohlavie	-0,03	0,97	0,03	0,85
vek	0,05	1,05	100,18	0,00
BMI	0,00	1,00	0,00	0,96
Ø hladina glukózy	0,00	1,00	200,79	0,00
Fajčil v minulosti	-0,21	0,81	0,88	0,35
Nikdy nefajčil	0,21	1,24	0,99	0,32
Aktívny fajčiar	0,09	1,09	0,12	0,73
Hypertenzia	0,63	1,88	11,83	0,00
Kardiovaskulárne ochorenie	0,60	1,82	7,66	0,01
Súkromný sektor	-1,53	0,22	2,20	0,14
Živnostník	-1,54	0,21	2,18	0,14
Práca v štátnej správe	-1,50	0,22	2,04	0,15

*Signifikatné hodnoty sú zvýraznené červenou farbou.

Použitím logistickej regresie sme schopní predikovať pravdepodobnosť výskytu mŕtvice na základe jeho pohlavia, veku, vysokého krvného tlaku, kardiovaskulárnych chorôb, priemeru hladiny glukózy a BMI. Dodatočne sa pozrieme na to, či aj typ zamestnania a teda i nikotínová závislosť predstavujú rizikový faktor pre mozgovú príhodu. Naším hlavným cieľom bude určiť, ktorý spomedzi vymenovaných prediktorov je najčastejšie spájaný s výskytom mŕtvice.

K výpočtu logistickej regresie sme využili program Statistica. Jej výsledky sme uviedli v tabuľke č. 1. Ku každému prediktoru sme pridali jeho hodnotu regresného koeficientu s hodnotou šance a príslušnou p-hodnotou. Za štatisticky významné považujeme výsledky s hodnou $p < 0,05$.

Štatisticky významná p-hodnota sa potvrdila pri 4 regresoroch – vek, priemerná hladina glukózy, vysoký krvný tlak a prítomnosti kardiovaskulárneho ochorenia. Naopak pohlavie ($p=0,85$) a BMI ($p=0,96$) nie sú štatisticky významné. Našou referenčnou skupinou v prípade „zamestnania“ tvorili deti. V porovnaní s dospelými pracujúcich v súkromnom sektore ($p=0,14$) ako živnostníci ($p=0,14$) alebo vo vláde ($p=0,15$) sme nenašli štatisticky významný rozdiel. Rovnako sme nenašli významný vplyv ani v prípade, či človek nefajčil ($p=0,32$), v minulosti fajčil ($p=0,35$) alebo je v súčasnosti fajčiarom ($p=0,73$).

Najsilnejšími a teda i najpodstatnejšími prediktorom tvorí prítomnosť hypertenzia ($p < 0,001$). Pokiaľ teda človek trpí vysokým krvným tlakom, je pravdepodobnosť výskytu mŕtvice o 1,88-krát vyššia. Hypertenziu teda môžeme označiť ako faktor, ktorý najviac ovplyvňuje, či sa u dotyčného pacienta mozgová príhoda objaví.

Ukazovatelia kvality modelu

Ukazovatelia	Hodnota	Hodnota (%)
Cox-Snell R ²	0,13	13%
Nagelkerke R ²	0,40	40%

Kvalitu nášho modelu sme overovali prostredníctvom ukazovateľov Cox-Snell R² a Nagelkerke R². Umožňujú nám určiť koľko % rozptylu sme schopní vysvetliť pomocou našich prediktorov. Hodnota Cox-Snell R² je rovná 0,13 %, teda vysvetľuje celkom 13 % a hodnota Nagelkerke R² sa rovná 0,40 teda vysvetľuje 40 % rozptylu. Celkovo sú hodnoty ukazovateľov kvality modelu stredné až dobré, čo značí, že binomálny regresný model ma dobrú prispôsobivosť k dátam a môže byť použitý k vytvoreniu pomerne presných predikcii.

Diskusia

Výsledky tejto správy sa zhodujú s výsledkami pôvodnej štúdie Pathon et al. (2020). V ich štúdií rovnako podčiarkujú 4 významné prediktory, ktorými sú vek, hladina glukózy, hypertenzia a výskyt srdcového ochorenia. Zhodu nachádzame s ďalšími štúdiami, ktoré za najdôležitejšie faktory, s potenciálom väčšej náchylnosti na mŕtvicu považujú práve hypertenziu a prítomnosť ochorenia srdca (Bandi et al., 2020; Alanazi, et al., 2020).

Tieto výsledky nám môžu pomôcť sústrediť sa na jednotlivé údaje zo zdravotného záznamu pacientov, nepodceňovať riziká a tak predchádzať nákladom spojených s ochorením ako je mozgová príhoda. Výsledky by mohli byť zaujímavé nie len pre lekárov, ale aj ľudí mladších i starších.

Zoznam zdrojov:

Alanazi, E. M., Abdou, A., & Luo, J. (2021). Predicting Risk of Stroke From Lab Tests Using Machine Learning Algorithms: Development and Evaluation of Prediction Models. (2021). *Journal of medical Internet research*, 5(12), e23440. doi: 10.2196/23440

Bandi, V., Bhattacharyya, D., & Midhunchakkravarthy, D. (2020). Prediction of Brain Stroke Severity Using Machine Learning. *Revue d'Intelligence Artificielle*, 34(6), 753-761. <https://doi.org/10.18280/ria.340609>

Pathon, M. S., Jianbiao, Z., John, D., Nagm A., & Dev, S. (2020). Identifying Stroke Indicators Using Rough Sets. *IEEE Access*, 8, 210318-210327. doi: 10.1109/ACCESS.2020.3039439

Rawanduzy, C. A., Earl, E., Mayer, C., & Lucke-Wold, B. (2023). Pediatric Stroke: A Review of Common Etiologies and Management Strategies. *Biomedicines*, 11(1). doi: 10.3390/biomedicines11010002

Sirsat, M. S., Fermé, E., & Câmara, J. (2020). Machine learning for brain stroke: a review. *Journal of Stroke and Cerebrovascular Diseases*, 29(10), 105162. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105162>