

VIZUALIZACE DAT O ÚMRTNOSTI V ANGLII POMOCÍ SOFTWARE R A INTERPRETACE SROVNÁNÍ MĚR ÚMRTNOSTI U OČKOVANÝCH A NEOČKOVANÝCH PROTI ONEMOCNĚNÍ COVID-19

VENCÁLEK Ondřej (CZ)

VYŽIADANÁ PREDNÁŠKA

Abstrakt: Mohlo by se zdát, že epidemie onemocnění COVID-19, která silně po několik let ovlivňovala naše životy, patří definitivně minulosti. Zatímco se mediální pozornost přesunula k jiným, aktuálnějším problémům, vědecká komunita nadále zkoumá různé aspekty týkající se efektivity jak nefarmakologických opatření typu uzavření škol či sportovišť, tak také léčebných postupů a zejména prevence onemocnění. Hlavní pozornost je přitom zcela pochopitelně věnována efektivitě vakcín, o nichž se hovořilo jako o „jízdence ven z pandemie“.

V příspěvku se zamyslíme nad tím, co můžeme říci z veřejně dostupných dat o efektivitě vakcín proti onemocnění COVID-19. Efektivitu přitom můžeme uvažovat proti samotné nákaze, proti těžkému průběhu vyžadujícímu hospitalizaci a proti úmrtí na COVID-19. Zaměříme se přitom právě na efektivitu, s jakou vakcíny brání úmrtí. Prozkoumáme data o úmrtnosti z Anglie, která zveřejnil Britský národní statistický úřad (Office for National Statistics) v srpnu 2023 [1]. Ukážeme několik vizualizací těchto dat, která se objevila na sociálních sítích. Vysvětlíme, v čem je úskalí interpretace těchto grafů.

Ukážeme, jak data vizualizovat pomocí softwaru R [2], a to zejména s využitím balíčku ggplot2! [3], který v současnosti představuje zřejmě nejpopulárnější grafický nástroj softwaru R. K práci s daty využijeme balíček dplyr [4].

V diskusi vysvětlíme rozdíl mezi dvěma základními typy studií, pomocí nichž zkoumáme efekty různých faktorů na lidské zdraví. Jde o experimentální a observační studie. Velká část (ne-li většina) našich poznatků týkajících se efektivity vakcín proti onemocnění COVID-19 je vyvozena ze srovnání nemocnosti či úmrtnosti očkované a neočkované populace. Například výše zmíněná data o úmrtnosti v Anglii jsou observační povahy. To s sebou ovšem nese riziko různých zkreslení.

V současnosti se živě diskutuje o možném zkreslení známém jako Healthy Vaccine Bias (zkreslení zdravého vakcinovaného), viz např. [5]. Objevuje se tento efekt i v datech o úmrtnosti v Anglii?



doc. Mgr. Ondřej Vencálek, Ph.D. Matematik, statistik, učitel. Pro praktiky teoretik, pro teoretiky praktik. Vystudoval obor pravděpodobnost a matematická statistika na Matematicko-fyzikální fakultě Univerzity Karlovy, kde po magisterském studiu (zakončeném v roce 2007) pokračoval ve studiu doktorském, které ukončil v prosinci 2011 obhájáním disertace z oblasti neparametrické statistiky a klasifikace. Zájem o praktické použití statistiky v oblasti medicíny ho vedl k působení na oddělení biostatistiky Státního zdravotního ústavu, v roce 2009 se však vrátil z Prahy na Moravu a začal působit na Univerzitě Palackého v Olomouci.

ního zdravotního ústavu, v roce 2009 se však vrátil z Prahy na Moravu a začal působit na Univerzitě Palackého v Olomouci.

Těší ho, má-li možnost proslovit popularizační přednášku, jako třeba tu v Planetáriu v Brně (https://www.youtube.com/watch?v=4jE_b3nenQo) nebo sepsat své myšlenky v článku s lehce filosofickým nádechem (<https://www.sisyfos.cz/old/files/Zpravodaj-1-2017.pdf>, https://www.statapol.cz/wp-content/uploads/2022/05/IB_2_2022.pdf).

Od roku 2019 je předsedou České statistické společnosti, jejíž činnost se snaží propagovat. Patří mezi popularizátory bayesovských metod, spoluzakládal v současnosti již neexistující uskupení 4BIN. V průběhu epidemie onemocnění Covid-19 se snažil poukazovat na potřebu kvalitního sběru dat a následných nezaujatých analýz.

Literatúra

- [1] Deaths by vaccination status, England. Office for National Statistics [online]. [cit. 2024-06-07]. Dostupné z <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/deathsbyvaccinationstatusengland>.
- [2] R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- [3] Wickham, H.: *ggplot2: Elegant Graphics for Data Analysis*, Springer-Verlag New York, 2016. [cit. 2024-06-07]. Dostupné z <https://ggplot2-book.org/>.
- [4] Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D.: *A Grammar of Data Manipulation*, (2023), R package version 1.1.3.
- [5] Fürst, T., Bazalová, A., Fryčák, T., Janošek, J.: *Does the healthy vaccinee bias rule them all?* Association of COVID-19 vaccination status and all-cause mortality from an analysis of data from 2.2 million individual health records. *Int J Infect Dis.* 2024 May; 142:106976. doi: 10.1016/j.ijid.2024.02.019. Epub 2024 Feb 22. PMID: 38401782.