

Reprodukční a vybrané zdravotní události u žen starších 15 let, k nimž došlo na území ČR od roku 1994

VERZE 2025-02.01

Metodický popis

Autorský kolektiv:	Hana Davidová ¹ , Jitka Jírová ¹ , Ondřej Šanca ¹ , Jiří Jarkovský ¹ , Daniel Klimeš ¹ , Jan Mužík ¹ , Martin Komenda ¹ , Ladislav Dušek ¹ ¹ Ústav zdravotnických informací a statistiky ČR
Recenzenti:	
Datum vytvoření dokumentu:	15.09.2025
Datum poslední aktualizace:	04.12.2025
Datum vyjádření pověření pro ochranu osobních údajů UZIS:	04.12.2025
Datum schválení poradou vedení UZIS:	23.10.2025
Typ výstupu:	Otevřená data
ID výstupu:	NR-12-17
Klíčová slova	NRRZ, Národní registr reprodukčního zdraví, LPZ, List o prohlídce zemřelého, ISIN, Informační systém infekční nemoci, úmrtí, porod, potrat, asistovaná reprodukce, COVID-19, očkování
EuroVoc	státní zdravotní služba, porodnost, očkování, nemoc
Citace:	Davidová H., Jírová J., Šanca O., Jarkovský J., Klimeš D., Mužík J., Komenda M., Dušek L. Reprodukční a vybrané zdravotní události u žen starších 15 let, k nimž došlo na území ČR od roku 1994. <i>Národní zdravotnický informační portál</i> [online]. Praha: Ministerstvo zdravotnictví ČR a Ústav zdravotnických informací a statistiky ČR, 2025. Dostupné z: http://www.nzip.cz/data/2547-reprodukci-zdravotni-udalosti-otevrena-data . ISSN 2695-0340
Datová sada dostupná na:	http://www.nzip.cz/data/2547-reprodukci-zdravotni-udalosti-otevrena-data .
Kontakt:	analzy@uzis.cz

1 Popis datové sady

1.1 Primární cíl

Datová sada umožňuje sledovat reprodukční a zdravotní události žen v časovém kontextu. Obsahuje informace o úmrtích, potratech, porodech a onemocněních a očkování proti COVID-19. Každá událost je přiřazena konkrétní ženě a je označena rokem dané události a kategorizovaným věkem ženy v době dané události. Díky tomu lze analyzovat jednotlivé životní události žen, jejich četnost a časové rozložení.

Datová sada je strukturována jako řádková data, kde každý řádek odpovídá jedné konkrétní události dané ženy. Ženy jsou v sadě identifikovány unikátním identifikátorem, takže je možné sledovat všechny jejich události v chronologickém pořadí a propojit je s informacemi o věku, typu události a dalších charakteristikách. Tento přístup umožňuje analyzovat reprodukční a zdravotní historii žen v návaznosti na prodělaná očkování a onemocnění COVID-19.

1.2 Základní informace

Pro umožnění analýzy reprodukčních a zdravotních událostí žen (porody, potraty, úmrtí), onemocnění a očkování proti COVID-19 vznikla datová sada popisující jednotlivé události v řádkových datech.

Datová sada vzniká v souladu s národní metodikou tvorby otevřených dat a je dostupná v národním katalogu.

Obsah datové sady zahrnuje základní údaje o ženě a konkrétní události (rok narození kategorizovaný do desetiletí, kategorizovaný věk ženy v době události, typ, pořadí a rok události) a následně další specifikace jednotlivých událostí (diagnóza úmrtí; stáří plodu při samovolném potratu; informaci, zda umělý potrat byl proveden s nebo bez zdravotních důvodů; četnost porodu; pořadí a typ očkovací dávky a pořadí infekce COVID-19).

Z technického hlediska se jedná o údaje z NRRZ, blíže viz dokumentace [Národního registru reprodukčního zdraví](#); údaje z LPZ, blíže viz dokumentace [Listu o prohlídce zemřelého](#) a údaje z ISIN, blíže viz dokumentace [Informačního systému infekční nemoci](#).

1.3 Interpretace a informační hodnota a limitace

Datová sada vychází z dat Národního registru reprodukčního zdraví (NRRZ), Listu o prohlídce zemřelého (LPZ) a Informačního systému infekční nemoci (ISIN).

Cílem datové sady je poskytnout informace o reprodukčních a zdravotních událostech žen a o onemocnění a očkování těchto žen proti COVID-19.

Z důvodu ochrany osobních údajů fyzických osob je limitována úroveň detailu popisných charakteristik (uveden je pouze rok narození kategorizovaný do desetiletí a nejsou blíže specifikovány další demografické údaje). Zahrnuty jsou pouze ženy starší 15 let. V případě reprodukčních událostí (porody a potraty) jsou pak zahrnuty pouze ženy ve věku 15–55 let. Věk je kategorizován do dvou kategorií: do 34 let a nad 35 let včetně. Zahrnuty jsou události od roku 1994 až do posledního uzavřeného roku.

Sloupec s kategorizovanou hodnotou věku ženy v době dané události hodnotu obsahuje pouze tehdy, pokud jej bylo možné určit z dostupných dat. V případech, kdy byl znám pouze rok narození, ale ne celé datum, nebyl věk dopočítáván (typicky pro očkování a onemocnění COVID-19), jelikož by šlo pouze o orientační rozdíl mezi roky, nikoli o přesný věk v době dané události. Pokud však byl věk v datech již

uveden (např. vypočtený z kompletních dat o narození a události), zůstal zachován a byl následně kategorizován.

U části žen rok narození v datech zcela chybí, a proto u těchto záznamů nelze určit věk ani jej nijak dopočítat. Tyto případy zůstaly v datové sadě zachovány, protože i přes neúplné informace o věku mohou poskytovat důležité údaje o výskytu a načasování jednotlivých událostí.

Data obsahují sloupec s názvem unikátní_overene_RC. Jde o sloupec nesoucí informaci, zda bylo u dané ženy možné ověřit unikátní rodné číslo. Pokud toto ověření možné nebylo (hodnota 0 ve sloupci unikátní_overene_RC), nelze s jistotou určit, zda dané ID odpovídá pouze jedné unikátní ženě.

1.4 Příklady použití

Typickým příkladem využití datové sady je analýza vztahů mezi reprodukčními a zdravotními událostmi, výskytem onemocnění COVID-19 a očkováním proti COVID-19. Datová sada umožňuje zkoumat, jak se tyto různé typy událostí časově prolínají, zda na sebe nějak navazují, a jak se jejich rozložení mění v čase nebo mezi různými skupinami žen (např. generacemi žen podle roku narození).

Díky tomu, že každý řádek představuje konkrétní událost přiřazenou jedné ženě a že jednotlivé události jsou chronologicky seřazeny, lze data využít i pro analýzy průběhu života žen. To umožňuje například sledovat sled událostí (např. očkování → těhotenství → porod) nebo vytvářet přehledové statistiky o výskytu jednotlivých typů událostí a jejich vzájemných souvislostech.

1.5 Ochrana osobních údajů

Datová sada je připravována v souladu se zákonem 372/2011 Sb. (Zákon o zdravotních službách a podmínkách jejich poskytování) §73 odst. 8 a zákonem 110/2019 Sb. §16.

Data jsou anonymizovaná a pro propojení řádků mezi sebou (identifikace jednotlivých žen) jsou využity uměle vytvořené identifikátory.

1.6 Klíčová slova

NRRZ, Národní registr reprodukčního zdraví, LPZ, List o prohlídce zemřelého, ISIN, Informační systém infekční nemoci, úmrtí, porod, potrat, asistovaná reprodukce, COVID-19, očkování

1.7 Klasifikace dle EuroVoc

státní zdravotní služba, porodnost, očkování, nemoc

2 Technická charakteristika

- Periodicita aktualizace: nepravidelná
- Formát datové sady: csv
- Počet záznamů: 17 662 487
- Počet popisných atributů: 15
- Velikost: 850 MB

3 Zdroj dat

Zdroji dat pro datovou sadu jsou Národní registr reprodukčního zdraví (NRRZ), List o prohlídce zemřelého (LPZ) a Informační systém infekční nemoci (ISIN), které jsou ve správě ÚZIS ČR dle zákona § 74 a § 75 zákona č. 372/2011 Sb., o zdravotních službách.

4 Agregace a transformace dat

Původní záznamy událostí byly sloučeny do jednotné datové sady, v níž je každé ženě přiřazeno unikátní identifikační číslo. Každý řádek představuje jednu událost, přičemž jednotlivé události jsou chronologicky seřazeny podle data jejich výskytu. V rámci transformace byla data sjednocena do společné struktury, která umožňuje propojit různé typy událostí (např. porod, potrat, očkování, úmrtí) a analyzovat je v čase. Věk ženy při dané události byl dopočítán, a následně kategorizován, pouze pokud bylo k dispozici přesné datum narození.

5 Atributy datové sady

1. ID ženy

- Název parametru: id_zeny
- Datový typ: numeric
- Popis parametru: unikátní ID popisující jednu konkrétní ženu v datech
- Možné hodnoty: číselné hodnoty od 1 do počtu unikátních žen v datech

2. Unikátní ověřené rodné číslo

- Název parametru: unikatni_overene_RC
- Datový typ: numeric
- Popis parametru: číselná hodnota určující, zda je dané id_zeny možné brát jako unikátní, na základě ověřeného rodného čísla
- Možné hodnoty:
 - 0 – neověřené rodné číslo
 - 1 – ověřené rodné čísloPřípadně prázdná hodnota v případě chybějící informace

3. Desetiletí narození

- Název parametru: desetileti
- Datový typ: numeric
- Popis parametru: desetiletí roku narození ženy; např. hodnota 1980 odpovídá rokům narození 1980–1989 apod.
- Možné hodnoty: číselné hodnoty desetiletí roku narození, případně prázdná hodnota při chybějící informaci

4. Věk

- Název parametru: vek
- Datový typ: numeric
- Popis parametru: kategorizovaný věk ženy v době události vycházející [Číselníku věkových skupin NZIS](#)
- Možné hodnoty:
 - 66000034 – odpovídá věku 15–34 let
 - 66035999 – odpovídá věku 35 let a více

5. Událost

- Název parametru: udalost
 - Datový typ: string
 - Popis parametru: typ události
 - Možné hodnoty:
 - umrti
 - samovolny potrat
 - umely potrat
 - porod
 - covid ockovani
 - covid onemocneni
- 6. Pořadí události**
- Název parametru: poradi_udalosti
 - Datový typ: numeric
 - Popis parametru: pořadí, v jakém se daná událost vyskytla u dané ženy
 - Možné hodnoty: číselné hodnoty od 1 do počtu událostí dané ženy
- 7. Rok události**
- Název parametru: rok_udalosti
 - Datový typ: numeric
 - Popis parametru: rok, v jakém se daná událost vyskytla u dané ženy
 - Možné hodnoty: číselné hodnoty od roku 1994 do posledního uzavřeného roku
- 8. Úmrtí – diagnóza**
- Název parametru: umrti_diagnoza
 - Datový typ: string
 - Popis parametru: hlavní diagnóza úmrtí dle MKN-10 kódu v délce 3 znaků
 - Možné hodnoty: viz [číselník diagnóz MKN-10](#) o délce 3 znaky, případně prázdná hodnota, nebylo-li danou událostí úmrtí, nebo při chybějící informaci
- 9. Samovolný potrat – stáří plodu v týdnech**
- Název parametru: s_potrat_stari_plodu_v_tydnech
 - Datový typ: numeric
 - Popis parametru: stáří plodu v týdnech při samovolném potratu
 - Možné hodnoty: číselné hodnoty označující stáří plodu v týdnech v rozmezí 3 až 28 týdnů, případně prázdná hodnota, nebyl-li danou událostí samovolný potrat, nebo při chybějící informaci
- 10. Umělý potrat – s/bez zdravotního důvodu**
- Název parametru: u_potrat_s_bez_zdrav_duvodu
 - Datový typ: numeric
 - Popis parametru: typ umělého potratu
 - Možné hodnoty:
 - 0) 0 – bez zdravotního důvodu
 - 1) 1 – se zdravotním důvodem
- Případně prázdná hodnota, nebyl-li danou událostí umělý potrat, nebo při chybějící informaci
- 11. Porod – četnost**
- Název parametru: porod_cetnost
 - Datový typ: numeric
 - Popis parametru: počet narozených dětí při daném porodu
 - Možné hodnoty: číselné hodnoty od 1 do počtu narozených dětí při daném porodu, případně prázdná hodnota, nebyl-li danou událostí porod, nebo při chybějící informaci

12. Porod – gestační stáří

- Název parametru: porod_gestacni_stari
- Datový typ: numeric
- Popis parametru: gestační stáří v týdnech
- Možné hodnoty: číselné hodnoty označující gestační stáří v týdnech v rozmezí 20 až 49 týdnů, případně prázdná hodnota, nebyl-li danou událostí porod, nebo při chybějící informaci

13. COVID-19 očkování – pořadí dávky

- Název parametru: covid_ocko_poradi_davky
- Datový typ: numeric
- Popis parametru: pořadí očkovací dávky u dané ženy
- Možné hodnoty: číselné hodnoty od 1 do počtu dávek u dané ženy, případně prázdná hodnota, nebylo-li danou událostí COVID-19 očkování, nebo při chybějící informaci

14. COVID-19 očkování – typ

- Název parametru: covid_ocko_typ
- Datový typ: string
- Popis parametru: typ COVID-19 očkování
- Možné hodnoty:
 - CO01 – Comirnaty
 - CO02 – SPIKEVAX
 - CO03 – VAXZEVRIA
 - CO04 – COVID-19 Vaccine Janssen
 - CO05 – Sputnik V
 - CO07 – Nuvaxovid
 - CO08 – Comirnaty Original/Omicron BA.1
 - CO09 – Comirnaty Original/Omicron BA.4/BA.5
 - CO10 – Covovax
 - CO11 – Sinopharm
 - CO12 – Sinovac
 - CO13 – COVAXIN
 - CO14 – Covishield
 - CO15 – Spikevax bivalent Original/Omicron BA.1
 - CO16 – Comirnaty 6m-4
 - CO19 – SPIKEVAX BIVALENT ORIGINAL/OMICRON BA.4-5
 - CO20 – Comirnaty Omicron XBB.1.5.
 - CO21 – Comirnaty Omicron XBB.1.5. 5-11
 - CO22 – Nuvaxovid XBB 1.5
 - CO24 – Comirnaty JN.1

Případně prázdná hodnota, nebylo-li danou událostí COVID-19 očkování, nebo při chybějící informaci

15. COVID-19 onemocnění – pořadí infekce

- Název parametru: covid_onemoc_poradi_infekce
- Datový typ: numeric
- Popis parametru: pořadí infekce COVID-19 onemocnění u dané ženy
- Možné hodnoty: číselné hodnoty od 1 do počtu onemocnění COVID-19 u dané ženy, případně prázdná hodnota, nebylo-li danou událostí COVID-19 onemocnění, nebo při chybějící informaci