

Přežití pacientů s novotvarem v ČR

VERZE 2024-02.0

Metodický popis

Autorský kolektiv:	Jan Mužík ¹ , Denisa Krejčí ¹ , Michaela Cvanová ¹ , Petr Klika ¹ , Jiří Jarkovský ¹ , Klára Benešová ¹ , Daniel Klimeš ¹ , Martin Komenda ¹ , Ladislav Dušek ¹ ¹ Ústav zdravotnických informací a statistiky ČR
Recenzenti:	
Datum vytvoření dokumentu:	3. 8. 2023
Datum poslední aktualizace:	27. 11. 2024
Datum schválení pověřencem pro ochranu osobních údajů ÚZIS:	18. 3. 2024
Datum schválení poradou vedení ÚZIS:	30. 4. 2024
Typ výstupu:	Datová sada
ID výstupu:	NR-07-03
Klíčová slova:	NOR, Národní onkologický registr, novotvary, populace, incidence, prevalence, přežití
EuroVoc:	NT1 zdravotní statistika, NT1 rakovina
Citace:	Mužík J., Krejčí D., Cvanová M., Klika P., Jarkovský J., Benešová K., Klimeš D., Komenda M., Dušek L. Přežití pacientů s novotvarem v ČR. <i>Národní zdravotnický informační portál</i> [online]. Praha: Ministerstvo zdravotnictví ČR a Ústav zdravotnických informací a statistiky ČR, 2023. Dostupné z: http://www.nzip.cz/data/1772-novotvary-preziti-otevrena-data . ISSN 2695-0340
Datová sada dostupná na:	http://www.nzip.cz/data/1772-novotvary-preziti-otevrena-data
Kontakt:	analyzy@uzis.cz

1 Popis datové sady

1.1 Základní informace

Datová sada „Přežití pacientů s novotvarem v ČR“ umožňuje podrobné analýzy přežití pacientů s novotvarem v ČR. Zdrojem dat jsou individuální anonymizované záznamy Národního onkologického registru (NOR) od roku 1994 do současnosti a obsahují základní sociodemografické údaje pacienta, údaje o diagnóze novotvaru a rozsahu onemocnění a údaj o případném úmrtí pacienta.

1.2 Primární cíl

Primárním cílem je poskytnout data pro komplexní hodnocení přežití onkologických pacientů na úrovni ČR.

1.3 Interpretace a informační hodnota

Datová sada umožňuje hodnocení přežití pacientů s onkologickým onemocněním na úrovni celé ČR, včetně hodnocení trendů od roku 1994 do současnosti.

Pro hodnocení přežití pacientů obsahuje datová sada údaje o pořadí novotvaru u osoby, údaje o délce sledování ve dnech a údaje o stavu pacienta (žije/zemřel).

Ve spojení s populačními demografickými údaji a úmrtnostními tabulkami (nejsou součástí této datové sady) umožňují tato data hodnocení relativního populačně standardizovaného přežití.

1.4 Informační limity

Z důvodu ochrany osobních údajů fyzických osob je limitována úroveň detailu popisných charakteristik na úroveň ČR, diagnózy jsou agregovány do skupin diagnóz podle MKN10, rok stanovení diagnózy a rok úmrtí, nikoliv celé datumy), není uváděno anonymní ID osoby, je uvedeno pouze pořadí novotvaru u pacienta.

1.5 Příklad použití

Typickým příkladem využití dat je hodnocení relativního přežití pacientů s danou diagnózou, stratifikovaného podle stadií onemocnění.

Příklad konkrétního výstupu:

U zhoubného novotvaru průdušnice, průdušky a plíce (C33–C34) dosahovalo v období 2010–2014 5leté relativní přežití léčených pacientů ve stadiu I 61,3 %, ve stadiu II 34,5 %, ve stadiu III 11,4 % a ve stadiu IV 2,9 %, celkově pak 17,7 %, hodnoceno kohortní analýzou.

1.6 Ochrana osobních údajů a identity právnických osob

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.

Pro odstranění rizika identifikace fyzických osob jsou data kompletně anonymizována a neobsahují žádné identifikátory umožňující propojení řádků mezi sebou, jednotlivé řádky souboru jsou tak zcela nezávislé.

1.7 Klíčová slova

NOR, Národní onkologický registr, novotvary, populace, incidence, prevalence, přežití

1.8 Klasifikace dle EuroVoc

NT1 zdravotní statistika, NT1 rakovina

2 Technická charakteristika

- Periodicita aktualizace
 - 1x ročně za uzavřený kalendářní rok.
- Formát datové sady: csv
- Počet záznamů: 2 320 000
- Počet popisných atributů: 14
- Velikost: 134 MB
- Odborný garant: Jan Mužík

3 Zdroj dat

Zdrojem dat pro datovou sadu je databáze Národního onkologického registru (NOR) a data dalších databází Národního zdravotnického informačního systému (NZIS), které jsou ve správě ÚZIS ČR dle zákona 372/2011 Sb. (Zákon o zdravotních službách a podmínkách jejich poskytování). Výše uvedené zdroje jsou zpracovány do analytické databáze NOR, která je sestavena z vypořádaných dat primárních hlášení do NOR a doplněna z dat Národního registru hrazených zdravotních služeb (NRHZZS), Národního registru hospitalizací (NRHOSP) a Listu o prohlídce zemřelého (LPZ). Zpravodajskými jednotkami primárních dat jsou jednotliví poskytovatelé zdravotních služeb (PZS), kteří diagnostikují anebo léčí pacienty s nově zjištěnými novotvarů.

4 Agregace a transformace dat

Datová sada obsahuje řádková data nově zjištěných novotvarů. Použité transformace vybírají z datumu diagnózy a úmrtí pouze rok, je uváděno pořadí onemocnění u pacienta, nikoliv identifikátor pacienta; vše s cílem ochrany osobních údajů.

5 Atributy datové sady

1. ID řádku

- Název parametru: id
- Datový typ: numeric
- Popis parametru: anonymní pořadové id záznamu
- Možné hodnoty: 1 až n

2. Pohlaví

- Název parametru: pohlaví
- Datový typ: numeric
- Popis parametru: pohlaví pacienta
- Možné hodnoty:
 - 1 – muž
 - 2 – žena

3. Skupina diagnózy novotvaru

- Název parametru: diagnoza_skupina
- Datový typ: numeric
- Popis parametru: skupiny diagnóz novotvarů pro statistická hodnocení
- Možné hodnoty:
 - 1 – ZN hlavy a krku (C00–C14, C30–C31)
 - 2 – ZN jícnu (C15)
 - 3 – ZN žaludku (C16)
 - 4 – ZN tlustého střeva a konečníku (C18–C20)
 - 5 – ZN jater a intrahepatálních žlučových cest (C22)
 - 6 – ZN žlučníku a žlučových cest (C23, C24)
 - 7 – ZN slinivky břišní (C25)
 - 8 – ZN hrtanu (C32)
 - 9 – ZN průdušnice, průdušky a plíce (C33, C34)
 - 10 – zhoubný melanom kůže (C43)
 - 11 – nemelanomový kožní ZN (C44)
 - 12 – ZN pojivových a měk. tkání a perif. nervů (C47, C49)
 - 13 – ZN prsu (C50)
 - 14 – ZN hrdla děložního (C53)
 - 15 – ZN dělohy (C54, C55)
 - 16 – ZN vaječníku (C56)
 - 17 – ZN prostaty (C61)
 - 18 – ZN varlete (C62)
 - 19 – ZN ledviny (C64)
 - 20 – ZN močového měchýře (C67)
 - 21 – ZN mozku, míchy a jiných částí CNS (C70–C72)
 - 22 – ZN štítné žlázy (C73)
 - 23 – Hodgkinův lymfom (C81)
 - 24 – non-Hodgkinův lymfom (C82–C86)
 - 25 – mnohočetný myelom (C90)
 - 26 – leukémie (C91–C95)

- 27 – ostatní zhoubné novotvary
- 28 – novotvary in situ (D00–D09)
- 29 – novotvary nezhoubné a neznámého chování (D10–D36, D37–D48)

4. Pořadí novotvaru u osoby

- Název parametru: novotvar_poradi
- Datový typ: numeric
- Popis parametru: pořadí novotvaru u osoby
- Možné hodnoty:
 - 1 – první
 - 2 – další

5. Pořadí novotvaru v rámci diagnostické skupiny

- Název parametru: novotvar_poradi_dg_skupina
- Datový typ: numeric
- Popis parametru: pořadí novotvaru v rámci diagnostické skupiny u pacienta
- Možné hodnoty:
 - 1 – první
 - 2 – další

6. Pořadí novotvaru v rámci zhoubných novotvarů

- Název parametru: novotvar_poradi_maligni
- Datový typ: numeric
- Popis parametru: pořadí novotvaru v rámci zhoubných novotvarů (C00–C97) u pacienta
- Možné hodnoty:
 - 1 – první
 - 2 – další

7. Pořadí novotvaru v rámci zhoubných novotvarů bez dg. C44

- Název parametru: novotvar_poradi_maligni_bez_C44
- Datový typ: numeric
- Popis parametru: pořadí novotvaru v rámci zhoubných novotvarů (C00–C97) kromě nemelanomových kožních nádorů (C44) u pacienta
- Možné hodnoty:
 - 1 – první
 - 2 – další

8. Stadium onemocnění

- Název parametru: stadium
- Datový typ: string 1 znak
- Popis parametru: stadium onemocnění podle TNM klasifikace včetně neznámých stadií rozdělených podle úplnosti záznamu novotvaru
- Možné hodnoty:
 - 1 – stadium I
 - 2 – stadium II
 - 3 – stadium III
 - 4 – stadium IV
 - X – stadium neuvedeno z objektivních důvodů (časné úmrtí / bez protinádorové léčby / není zavedena TNM klasifikace)

- Y – chybějící údaj o stadiu – neúplný záznam

9. Rok diagnózy

- Název parametru: rok_dg
- Datový typ: numeric
- Popis parametru: rok stanovení diagnózy novotvaru
- Možné hodnoty: od 1994 dále

10. Kód věku při diagnóze

- Název parametru: vek_kod_dg
- Datový typ: numeric
- Popis parametru: kód věku pacienta při diagnóze novotvaru vycházející z číselníku věkových skupin
- Možné hodnoty: kódy věkových skupin vycházející z [číselníku věkových skupin](#); kódy odpovídají hodnotám věku 0–100

11. Kód věkové kategorie při diagnóze

- Název parametru: vek_kategorie_kod_dg
- Datový typ: numeric
- Popis parametru: kód věkové skupiny pacienta při diagnóze novotvaru vycházející z číselníku věkových skupin
- Možné hodnoty: kódy věkových skupin vycházející z [číselníku věkových skupin](#); kódy odpovídají hodnotám věku 0–84, 85+

12. Stav pacienta

- Název parametru: preziti
- Datový typ: numeric
- Popis parametru: stav pacienta
- Možné hodnoty:
 - 0 – žije
 - 1 – zemřel

13. Délka sledování ve dnech

- Název parametru: doba_sledovani
- Datový typ: numeric
- Popis parametru: délka sledování či přežití pacienta od data stanovení diagnózy do data posledního sledování či data úmrtí ve dnech
- Možné hodnoty: 0 až n

14. Protinádorová léčba

- Název parametru: lecen
- Datový typ: numeric
- Popis parametru: protinádorová léčba pacienta s novotvarem
- Možné hodnoty:
 - 0 – ne
 - 1 – ano