

Metodika identifikace onkologických pacientů z dat NZIS a příprava datových analytických souborů

Metodický materiál

Autoři:	RNDr. Jan Mužík, Ph.D. RNDr. Jiří Jarkovský, Ph.D. Mgr. Petr Klika MUDr. Miroslav Zvolský RNDr. Daniel Klimeš, Ph.D. RNDr. Martin Komenda, Ph.D., MBA Ing. Tomáš Májek RNDr. Ondřej Májek, Ph.D. Mgr. Jakub Gregor, Ph.D. prof. RNDr. Ladislav Dušek, Ph.D.
Verzování dokumentu:	Verze 1 (19.8.2025)

Obsah

1	Úvod	5
1.1	Cíl dokumentu	5
1.2	Důležité dokumenty	5
2	Použité datové zdroje	6
2.1	Národní registr hrazených zdravotních služeb	6
2.2	Národní onkologických registr	7
2.3	List o prohlídce zemřelého	7
3	Identifikace onkologické léčby v datech NRHZS	7
3.1	Identifikace fází probíhající protinádorové léčby	8
3.2	Určení diagnózy léčebné fáze	8
4	Dohledání odpovídajícího záznamu v NOR	8
5	Vytvoření výsledného záznamu pro <i>novotvar a jeho léčbu</i> pro otevřená onkologická data	9
6	Interpretace a informační hodnota publikovaných dat	10
7	Parametrizace údajů o novotvaru a jeho léčbě	10
7.1	Úroveň 1	10
7.2	Úroveň 2	11
7.3	Úroveň 3	11
8	Informační limity publikovaných dat	13
9	Ochrana osobních údajů a identity právnických osob	14

Historie verzí dokumentu

Verze	Datum	Popis změn
1	19.8.2025	První verze metodiky

Použité zkratky

HVLP	Hromadně vyráběné léčivé přípravky
IČO (ICO)	Identifikační číslo poskytovatele zdravotních služeb
IČP (ICP)	Identifikační číslo pracoviště (dle ZP)
IČZ (ICZ)	Identifikační číslo zdravotnického zařízení (dle ZP)
IPLP	Individuálně připravované léčivé přípravky
LPZ	List o prohlídce zemřelého
NRHVS	Národní registr hrazených zdravotních služeb
NOR	Národní onkologický registr
NZIS	Národní zdravotnický informační systém
PZS	Poskytovatel zdravotních služeb
RÚIAN	Registr územní identifikace, adres a nemovitostí
SÚKL	Státní úřad pro kontrolu léčiv
ÚZIS	Ústav zdravotnických informací a statistiky ČR
v.z.p.	Veřejné zdravotní pojištění
ZP	Zdravotní pojišťovna

1 Úvod

1.1 Cíl dokumentu

Tato metodická dokumentace popisuje tvorbu dat v oblasti onkologie, připraveným na základě struktury a obsahu Národního zdravotnického informačního systému (NZIS). Popisuje způsob identifikace a generování popisných parametrů onkologických pacientů a jejich léčby a možnosti jejich použití. Popisuje také předpoklady a limitace použití těchto dat a výstupů z nich připravených.

1.2 Důležité dokumenty

Další informace naleznete v následujících dokumentech:

Název dokumentu	Odkaz
Kompletní metodika sběru dat do NRHZS	https://www.uzis.cz/index.php?pg=registry-sber-dat--narodni-registr-hrazenych-zdravotnich-sluzeb#dokumenty
Metodika NZIS – Národní onkologický registr	https://www.uzis.cz/index.php?pg=registry-sber-dat--narodni-zdravotni-registry--narodni-onkologicky-registr#dokumenty
Metodika NZIS – Informační systém List o prohlídce zemřelého	https://www.uzis.cz/index.php?pg=registry-sber-dat--ostatni-rezortni-registry--list-o-prohlidce-zemreleho#dokumenty

2 Použité datové zdroje

Při tvorbě podkladů pro publikované datové sady jsou používány zejména následující datové zdroje.

2.1 Národní registr hrazených zdravotních služeb

Národní registr hrazených zdravotních služeb (NRHZS) obsahuje primárně záznamy o poskytnutých zdravotních službách, které byly poskytovatelem zdravotních služeb (PZS) vykázány zdravotní pojišťovně (ZP) a tou uznány k úhradě z veřejného zdravotního pojištění (v.z.p.). Sběr dat do NRHZS je dán zákonem č. 372/2011 Sb. („zákon o zdravotních službách“) a je řízen metodikou a datovým rozhraním ÚZIS ČR (viz <https://www.uzis.cz/index.php?pg=registry-sber-dat--narodni-registr-hrazenych-zdravotnich-sluzeb#dokumenty>).

NRHZS obsahuje údaje o všech formách zdravotních služeb:

- ambulantní (primární i sekundární péče),
- lůžkové (akutní, následné včetně lázeňské a dlouhodobé),
- jednodenní chirurgie a denní léčebné stacionáře,
- péče u poskytovatelů sociálních služeb a domácí péče,
- vydané léčivé prostředky a zdravotní pomůcky a materiál,
- diagnostický komplement,
- dopravní služby ve zdravotnictví (ZZS, přeprava raněných, nemocných a rodiček, přeprava zdravotnického personálu).

Mezi typy zdravotních služeb, které jsou do NRHZS předávány, patří zejména:

- zdravotní a dopravní výkony,
- aplikovaná nebo vydaná léčiva a léčivé přípravky, zdravotnický materiál a pomůcky,
- údaje o kategoriích náročnosti ošetrovatelské péče o hospitalizované pacienty,
- náklady na pobyt v lázních a ozdravovnách.

K provedeným zdravotním službám se evidují následující údaje:

- Forma a typ zdravotní služby (viz výše)
- Identifikace poskytovatele podle příslušných číselníků (IČZ, IČP a odbornost poskytujícího a žádajícího pracoviště)
- Identifikace pacienta
- Identifikace zdravotní pojišťovny
- Datum poskytnutí zdravotní služby
- Kód konkrétní zdravotní služby dle příslušných číselníků
- Množství a cena, resp. hodnota služby vykázané pojišťovně k úhradě
- Diagnózy související s poskytnutou službou (hlavní a případné doplňující/vedlejší)
- U hospitalizací dále datum zahájení a ukončení hospitalizace, důvod přijetí a způsob ukončení hospitalizace podle příslušných číselníků
- Apod.

Předávaná data a jejich interpretace je ovlivněna zejména následujícími faktory:

1. Účelem pořizování dat, kterým je vyúčtování zdravotních služeb PZS zdravotní pojišťovně za účelem získání úhrady z v.z.p. Pravidla pro jejich vykazování jsou dána právními předpisy,

metodickými dokumenty a smlouvou mezi PZS a ZP. Jsou dány limity pro místo, situace a četnosti provádění zdravotních služeb. Tedy ne vše, co PZS provede, lze také (přesně) vykázat zdravotní pojišťovně. V naprosté většině případů se předávají pouze informace o tom, jaká zdravotní služba byla poskytnuta, nikoliv co bylo jejím účelem a/nebo jaký měla výsledek. Zdravotní služby hrazené mimo v.z.p. se do registru nedostávají.

2. Kvalitou vykazování na straně poskytovatele. Různí poskytovatelé mají různé možnosti a různou motivaci důsledně vykazovat zdravotní služby. Různé předávané údaje mají různý vliv na úhradu a jejich kvalita nemusí být vždy dostatečná.
3. Validacním a revizním procesem na straně ZP. Záleží na motivaci a možnostech jednotlivých zdravotních pojišťoven kontrolovat předávaná data od poskytovatelů, kdy některé záznamy mohou kontrolovat přísněji a jiné více benevolentně.
4. Kvalitou předávání dat do NRHZS ze strany ZP a jejich následné zpracování. Do registru se dostávají pouze klíčové vybrané údaje z dat poskytovatelů podle metodiky a datového rozhraní registru. Každá pojišťovna pak implementuje svoje vlastní postupy pro předávání dat, které nemusejí být vždy zcela korektní a úplná. Dalším možným zkreslením mohou data projít při jejich zpracování v registru a přípravě analytických výstupů.

2.2 Národní onkologický registr

Národní onkologický registr ČR (NOR) je součástí Národního zdravotnického informačního systému (NZIS) podle zákona č. 372/2011 Sb. („zákon o zdravotních službách). NOR je celoplošným populačním registrem, jehož účelem je registrace onkologických onemocnění a periodické sledování jejich dalšího vývoje. NOR poskytuje souhrnné údaje o nově zjištěných onemocněních ze skupiny novotvary dle Mezinárodní statistické klasifikace nemocí a přidružených zdravotních problémů, 10. revize (dále jen MKN-10): zhoubné novotvary (C00–C97), novotvary in situ (D00–D09), novotvary nejistého nebo neznámého chování (diagnózy D37–D48), vybrané nezhoubné / benigní novotvary s morfologií 8683/0 gangliocytický paragangliom nebo 8936/0 gastrointestinální stromální tumor – benigní a dále pozitivní biotické či cytologické nálezy D12 - nezhoubný novotvar tlustého střeva, konečníku, řitního kanálu a řiti a N87 - dysplazie hrdla děložního.

2.3 List o prohlídce zemřelého

Registr / informační systém List o prohlídce (IS LPZ) zemřelého slouží ke sběru dat o zemřelých a okolnostech úmrtí, zejména pak o příčinách smrti. Postup při úmrtí a způsob předávání Listu o prohlídce zemřelého je ošetřen legislativně (zákon č. 372/2011 Sb., o zdravotních službách a podmínkách jejich poskytování, a vyhláška č. 297/2012 Sb., o Listu o prohlídce zemřelého).

Data IS LPZ metodicky zpracovává Český statistický úřad (ČSÚ) dle mezinárodních pravidel na základě údajů z (LPZ). Zemřelému je určena jedna hlavní příčina úmrtí, z těchto dat jsou pak zpracovány oficiální statistické výstupy o populační mortalitě dle příčin úmrtí.

3 Identifikace onkologické léčby v datech NRHZS

Jak bylo uvedeno výše, do NRHZS se předávají informace o tom, jaké zdravotní služby byly poskytnuty, ale nikoliv co bylo jejich účelem a jakého výsledku bylo dosaženo. Klinické údaje ve většině případů nejsou evidovány. Výjimkou jsou některé vybrané „markery“, které evidují např. stádium onkologického onemocnění při hospitalizaci (sledované v posledních několika málo letech od roku 2019) a údaje o hlavních a vedlejších diagnózách.

Přitom diagnóz se může v průběhu času u pacienta vyskytovat celá řada, a to více či méně specifických, potvrzených či jen suspektních. Také velká část terapeutických postupů (chirurgické zákroky, některá radioterapie, vybrané chemoterapeutické léky atd.) lze použít pro onkologické i jiné příčiny, resp. pro různé onkologické diagnózy. Záznam o tom, že pacient trpí nějakým onkologickým onemocněním a jakým, jak je léčen a případně s jakým výsledkem je tedy třeba z dat interpretovat nepřímo.

3.1 Identifikace fází probíhající protinádorové léčby

Základním postupem při identifikaci onkologických pacientů je identifikace úseků neboli fází (možné) protinádorové léčby z vykázaných zdravotních služeb. Za tímto účelem je udržován pravidelně aktualizovaný a odborníky oponovaný seznam terapeutických zdravotních výkonů a léčivých přípravků, které mohou být s velkou pravděpodobností použity při protinádorové léčbě onkologických pacientů. Je členěn do jednotlivých skupin (chirurgické výkony, radioterapeutické výkony dle jednotlivých typů záření, léčiva pro chemoterapii, imunoterapii, hormonální terapii, cílenou protinádorovou léčbu atd.). Seznam sledovaných zdravotních služeb je uveden v samostatné příloze 1.

V předávaných datech jsou takové záznamy vyhledány (zvláště pro každého pacienta a poskytovatele zdravotních služeb (PZS)) a seskládány do úseků (fází), pro které platí následující:

- Po sobě jdoucí záznamy o (možné) protinádorové léčbě nejsou uvnitř jedné léčebné fáze vzdáleny více než 3 měsíce (bez ohledu na to, o jaký typ léčby jde). U léčiv v tabletové formě, vydávaných na recept, se dopočítává pravděpodobná doba na léčbě podle množství vydaného léčiva a obvyklé denní terapeutické dávky přípravku.
- Pokud je u jednoho pacienta a jednoho PZS evidováno více léčebných fází, jsou od sebe odděleny více než 3 měsíce trvajícím obdobím bez (možné) protinádorové léčby.

3.2 Určení diagnózy léčebné fáze

Pro takto sestavené léčebné fáze se následně dopočítává předpokládaná diagnóza léčby. Algoritmus prochází hlavní a vedlejší diagnózy na záznamech z období léčebné fáze (u daného PZS) a období 3 měsíce před jejím zahájením a po jejím ukončení (u všech PZS). Počítá se jejich frekvence podle váhy: větší váha je pro hlavní diagnózy v období léčebné fáze, menší váha je pro vedlejší diagnózy v období léčebné fáze a hlavní diagnózy v období před začátkem a po skončení léčebné fáze. Výsledné hodnoty se seřadí sestupně a vybere se:

- První nejčtenější diagnóza, pokud není uvedena Z51* (diagnóza dříve používaná pro hospitalizace s chemoterapií nebo radioterapií)
- Druhá nejčtenější diagnóza, pokud první nebyla vybrána
- Pokud žádná diagnóza u léčebné fáze nebyla dohledána (pokud je evidována pouze léčba vydávaná na recept, kde uvedení diagnózy není povinné), zůstane diagnóza léčebné fáze prázdná

Ostatní nalezené diagnózy jsou evidovány jako vedlejší a mohou pomoci např. při dohledání odpovídajícího záznamu v jiném datové zdroji (zpravidla NOR).

4 Dohledání odpovídajícího záznamu v NOR

Výsledkem předchozího procesu je určení období, kdy probíhala (možná) protinádorová léčba a její nejpravděpodobnější diagnóza. K určení klíčových klinických parametrů ale bohužel nejsou údaje z NRHZS dostatečné, neboť zde není evidováno datum stanovení diagnózy, morfologická diagnostika

atd. Dalším krokem je tedy dohledání odpovídajícího záznamu v Národním onkologickém registru, je-li v něm evidován.

Procházejí se všechny záznamy v NOR u daného pacienta. Pro každou dvojici léčebné fáze a záznam v NOR se porovnává shoda časová a shoda diagnóz, vyjádřena pravděpodobností této shody. Záznamy s datem stanovení diagnózy přímo v období léčebné fáze mají pravděpodobnost nejvyšší a snižuje se s rostoucí vzdáleností data diagnózy od období léčby.

Pro ověřování shody diagnóz se používají pomocné číselníky diagnostických skupin a skupin lokalizací pro onkologická onemocnění, uvedených v samostatné příloze 2. Porovnávají se všechny diagnózy dohledané u léčebné fáze podle její váhy a určuje se pravděpodobnost shody s diagnózou v NOR. Ta je nejvyšší, pokud se shodují přesně na 4 resp. 3 místa (např. C18.1 nebo C18) a klesá pro shodu alespoň u diagnostické skupiny (např. C18–C20) nebo skupiny lokalizací (např. C18–C20, D01.0–D01.2, D12, D13.2–D13.5).

Výsledná pravděpodobnost přiřazení záznamu NOR k léčebné fázi je dána vynásobením pravděpodobností časové shody a shody diagnóz. Vybere se záznam s nejvyšší výslednou pravděpodobností, pokud tato překročí minimální prahovou hodnotu. Pokud je ekvivalentních záznamů více, vybírají se podle data stanovení diagnózy, kódu diagnózy a data nahlášení.

5 Vytvoření výsledného záznamu pro *novotvar a jeho léčbu* pro otevřená onkologická data

Výsledný záznam je kombinací záznamu v NOR a všech léčebných fází, které k němu byly dohledány. Záznam se doplňuje o další parametry z dat NRHZS, které se týkají primární diagnostiky, dispenzární péče, hospitalizací atd.

Vstupem jsou řádková data nově zjištěných novotvarů z registru NOR pro vybrané diagnostické skupiny novotvarů. Jednomu záznamu v NOR odpovídá jeden záznam v datové sadě (novotvar). Onkologická léčba pacientů, ke kterým nebyl záznam NOR dohledán nebo nepatří mezi vybrané diagnostické skupiny, není uvedena. K vybraným záznamům NOR je vybírána specifická léčba nebo diagnostické výkony podle výčtu, který je uveden v samostatných dokumentech s popisem otevřených datových sad.

Jednotlivé léčebné fáze vybraného novotvaru a jejich konkrétní terapeutické postupy (modality) jsou pak rozděleny na primární a následnou protinádorovou léčbu. Pro primární léčbu platí, že nezačíná dříve než 1 rok před stanovením finální diagnózy a později než půl roku po datu stanovení diagnózy. Pokud jde o první takový blok léčby u daného novotvaru, počítá se i léčba zahájená do 1 roku od stanovení diagnózy (odložená primární léčba) a ukončená později. Do primární léčby se nezahrnuje terapie (např. chemoterapie konkrétní léčivou látkou), která začala později než 1 rok od stanovení diagnózy, ani pokud bezprostředně navazuje na jinou protinádorovou terapii, která začala dříve a je zahrnuta do primární léčby.

Doba trvání jednotlivých terapií, které začaly v rámci primární terapie, není omezena, tj. primární léčba může zahrnovat i dlouhodobou adjuvantní léčbu, pokud začala za podmínek uvedených výše. V některých případech může primární léčba nesprávně zahrnovat i léčbu relapsu/progrese onemocnění, pokud je vývoj tohoto novotvaru velmi rychlý a následná léčba bezprostředně navazuje na léčbu primární – tyto situace nelze vždy obecně spolehlivě rozlišit.

Následná léčba je definována terapeutickými postupy uvedenými v samostatné dokumentaci a současně nesplňuje časové omezení pro primární léčbu novotvaru. Začala tedy později než 6 měsíců

(resp. 12 měsíců u první protinádorové léčby novotvaru) po datu stanovení diagnózy. V některých případech může následná léčba nesprávně zahrnovat i dlouhodobou adjuvantní léčbu, pokud nezačala za podmínek uvedených u primární léčby. Ve výsledné datové sadě jsou všechny následné léčebné fáze popsány v rámci jednoho bloku, bez ohledu na jejich počet a případné opakované remise onemocnění.

Na základě výše definovaných období, daných datem primární diagnózy, zahájením a ukončením primární léčby a zahájením následné léčby, jsou pak doplňovány další informace o provedených výkonech z NRHZS. Popisují provedení konzilia multioborového týmu, jednotlivé typy zobrazovacích nebo endoskopických vyšetření, dispenzární péči o pacienta po primární léčbě nebo konzumaci služeb ambulantních specialistů.

V případě úmrtí pacienta je doplněna doba do úmrtí a příčiny úmrtí podle údajů uvedených v IS LPZ.

6 Interpretace a informační hodnota publikovaných dat

Publikovaná data obsahují všechny záznamy z Národního onkologického registru (NOR) od roku 2011 do posledního uzavřeného roku, které mají uvedenu některou z vybraných diagnóz zhoubného novotvaru dospělých pacientů od 18 let věku. Obsahuje jejich detailní diagnostické informace včetně topografické a morfologické diagnózy, TNM klasifikace včetně výsledného stádia a lokalizace metastáz.

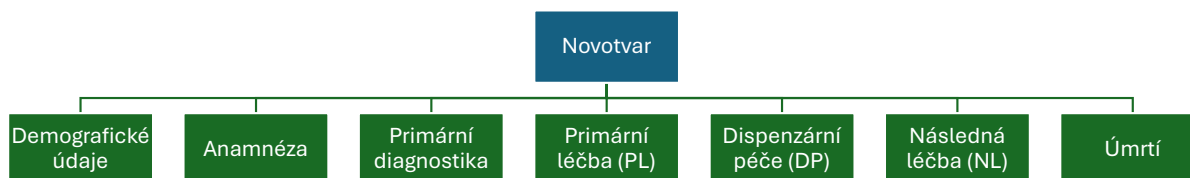
K těmto novotvarům jsou pak z Národního registru hrazených zdravotních služeb (NRHZS) doplněny detailní údaje o poskytnutých hrazených zdravotních službách od roku 2010 do posledního uzavřeného kalendářního roku. Zahrnuje údaje o primární diagnostice s využitím vybraných zobrazovacích nebo endoskopických metod, stratifikovaných podle typu (např. CT, PET/CT, RTG, SONO...), provedení konzultace multioborovým týmem před zahájením léčby, primární a následnou protinádorovou léčbou členěnou podle typu (operační zákrok, radioterapie, chemoterapie, „centrová“ protinádorová léčba...) nebo pobytu na lůžku akutní lůžkové péče během protinádorové léčby. Zahrnuje i údaje o dispenzarizaci pacienta nebo konzumaci služeb ambulantních specialistů po primární protinádorové léčbě.

Nakonec jsou k pacientům doplněny základní údaje o případném úmrtí, včetně doby do úmrtí od data primární diagnózy novotvaru a úmrtních diagnóz.

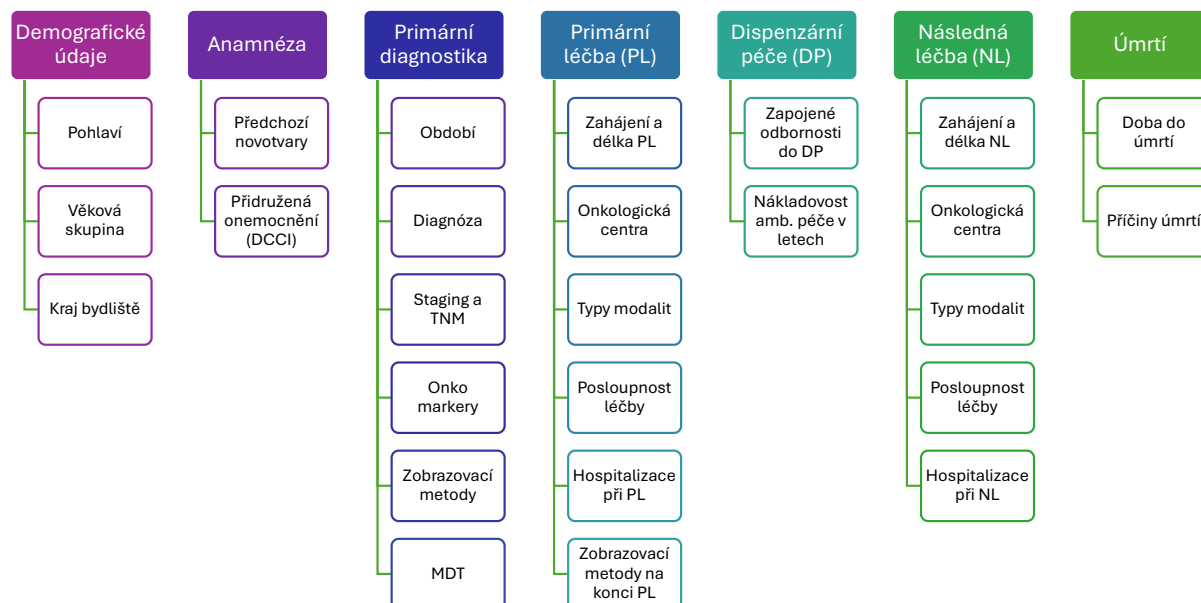
7 Parametrizace údajů o novotvaru a jeho léčbě

Jednotlivé okruhy generovaných parametrů jsou zobrazeny na následujících diagramech.

7.1 Úroveň 1

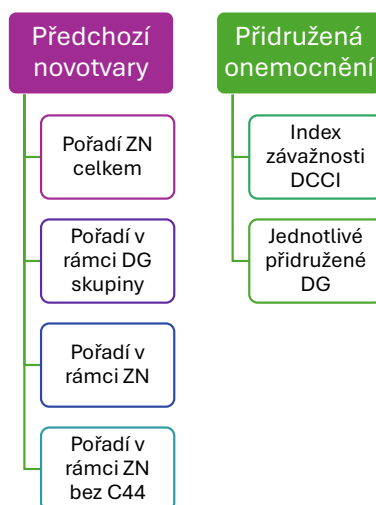


7.2 Úroveň 2

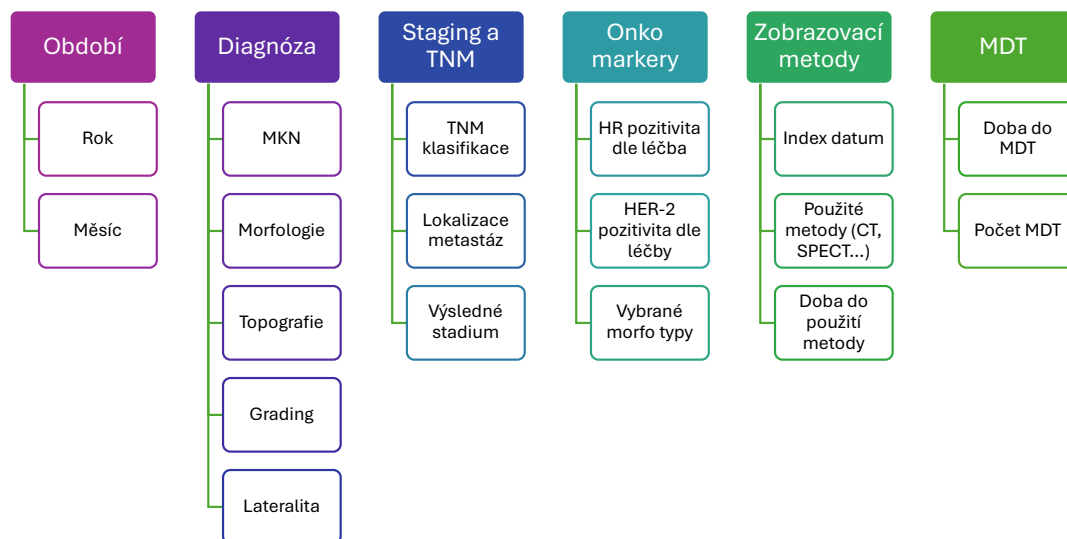


7.3 Úroveň 3

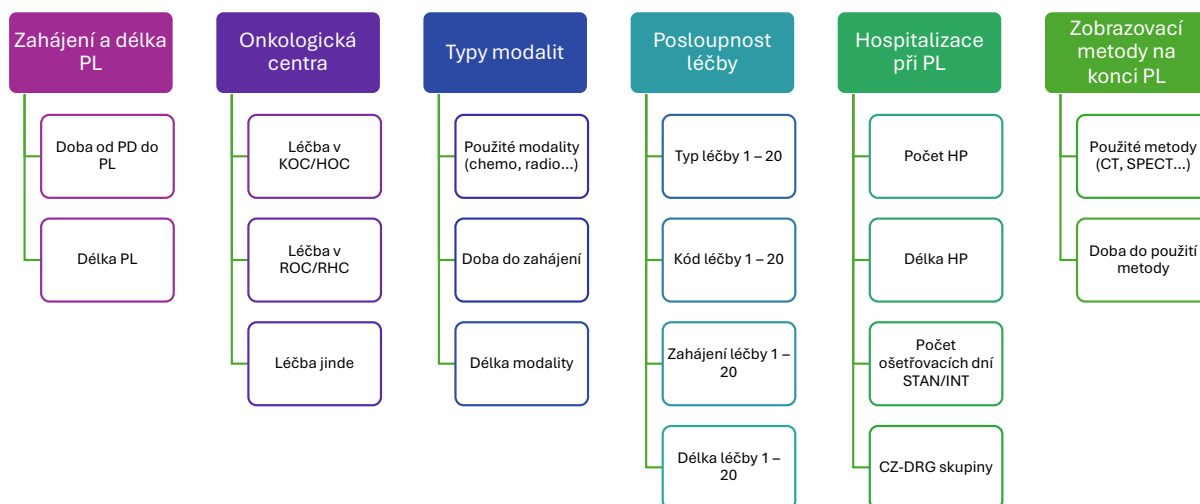
7.3.1 Anamnéza



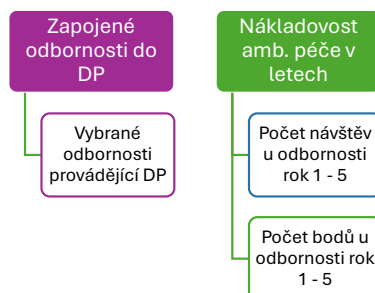
7.3.2 Primární diagnostika



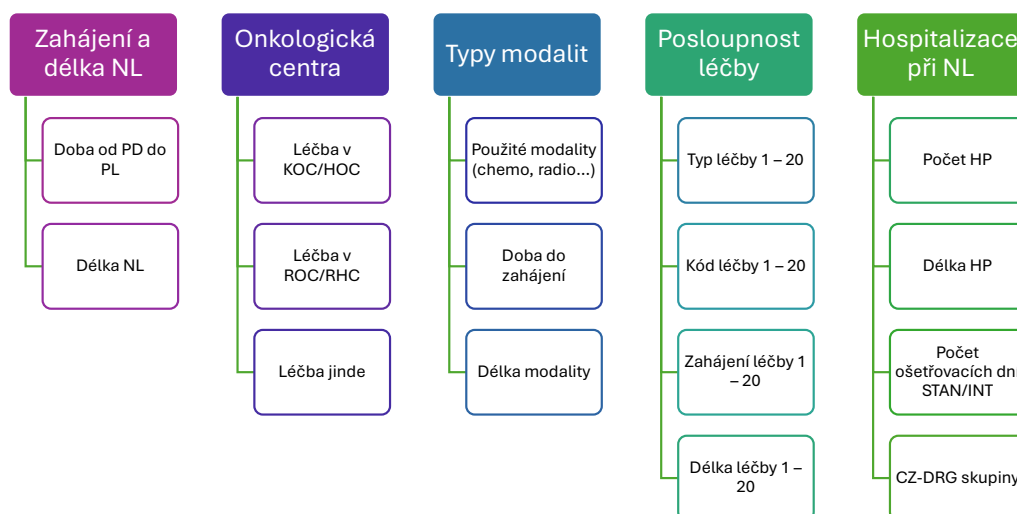
7.3.3 Primární léčba



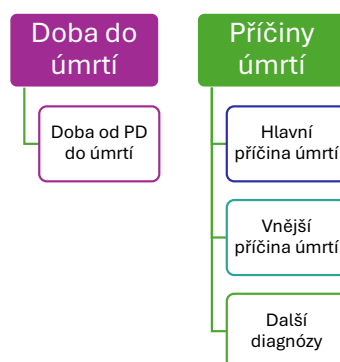
7.3.4 Dispenzární péče



7.3.5 Následná léčba



7.3.6 Úmrtí



8 Informační limity publikovaných dat

Při interpretaci publikovaných dat je třeba zohlednit zejména účel a původ dat o provedených výkonech a aplikované léčbě z dat NRHZS, jak bylo popsáno v samostatné kapitole.

V administrativních datech NRHZS se neuvádí přesný identifikátor novotvaru dle NOR, který je v daném období diagnostikován nebo léčen. Přiřazení léčby ke konkrétnímu novotvaru je dáno algoritmicky na základě četnosti vykazování a podobnosti diagnóz, časové souvislosti a konkrétních diagnostických a terapeutických postupů. Situace je komplikovaná zejména v případě opakované protinádorové léčby při relapsu/progresi původního onemocnění nebo výskytu více onkologických diagnóz u stejného pacienta (s různou nebo i stejnou diagnózou, např. další novotvar s opačnou lateralitou k předchozímu novotvaru). Může se tedy stát, že některá léčba je přiřazena k jinému novotvaru, než ke kterému ve skutečnosti přísluší.

Podobná komplikace může nastat při rozlišení primární léčby od léčby relapsu/progrese (následné léčby). V případech, kdy následná léčba bezprostředně navazuje na předchozí (neúspěšnou) léčbu primární, může dojít ke sloučení těchto léčebných fází pod primární léčbu. Naopak, pokud je primární léčba rozfázována do dlouhého období, zahrnuje dlouhodobou (adjuvantní) léčbu nebo je odložena o více než rok od data diagnózy, může být část nebo celá primární léčba nesprávně označena jako následná léčba.

Údaje o diagnostických výkonech a protinádorové léčbě nezahrnují výkony a léčiva, která jsou velmi málo četná (protinádorové terapie použita u méně než 1 % novotvarů, diagnostika u méně než 0,5 % novotvarů). Málo četné nebo nové postupy tedy nemusejí být uvedeny, i když ve skutečnosti byly u pacienta provedeny.

Při interpretaci údajů o přidružených onemocněních pacienta (DCCI) je třeba zohlednit, že Národní registr hrazených zdravotnických služeb využívá data vykazovaná za účelem proplácení zdravotní péče. Pacient nemusí danou diagnózou nutně trpět, ale může jít pouze o podezření či nepřesně určenou diagnózu. Index je proto vhodný k hodnocení zdraví na populační úrovni (např. pro srovnání populací v různých regionech), avšak detailní rozbor na úrovni jedinců by mohl vést k zavádějícím výsledkům. Další limitací představuje to, že index nezahrnuje některé skupiny chorob, které mohou mít značný dopad na život ve zdraví, resp. na invaliditu a stupeň závislosti jedince (např. psychiatrická onemocnění). Výsledky nejsou mezinárodně srovnatelné, jelikož databáze jednotlivých států mají odlišné pokrytí populace (např. pouze hospitalizovaní), různou kvalitu a granularitu sbíraných dat a různou délku hodnoceného časového období.

9 Ochrana osobních údajů a identity právnických osob

Datové sady jsou připravovány v souladu se zákonem č. 372/2011 Sb. (Zákon o zdravotních službách a podmínkách jejich poskytování) §73 odst. 8 a zákonem č. 110/2019 Sb. (Zákon o zpracování osobních údajů) §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.

Ze stejného důvodu jsou datumové údaje prezentovány pouze na úroveň roku a měsíce diagnózy, ostatní datumové údaje se vztahují vždy jen relativně k tomuto údaji (počet dní od data diagnózy). Geografické údaje (bydliště pacienta) jsou prezentovány pouze do úrovně krajů. Konkrétní poskytovatelé zdravotních služeb nejsou uváděni, pouze jejich příslušnost ke skupině center vysoce specializované péče (KOC/ROC, resp. HOC/RHC).