



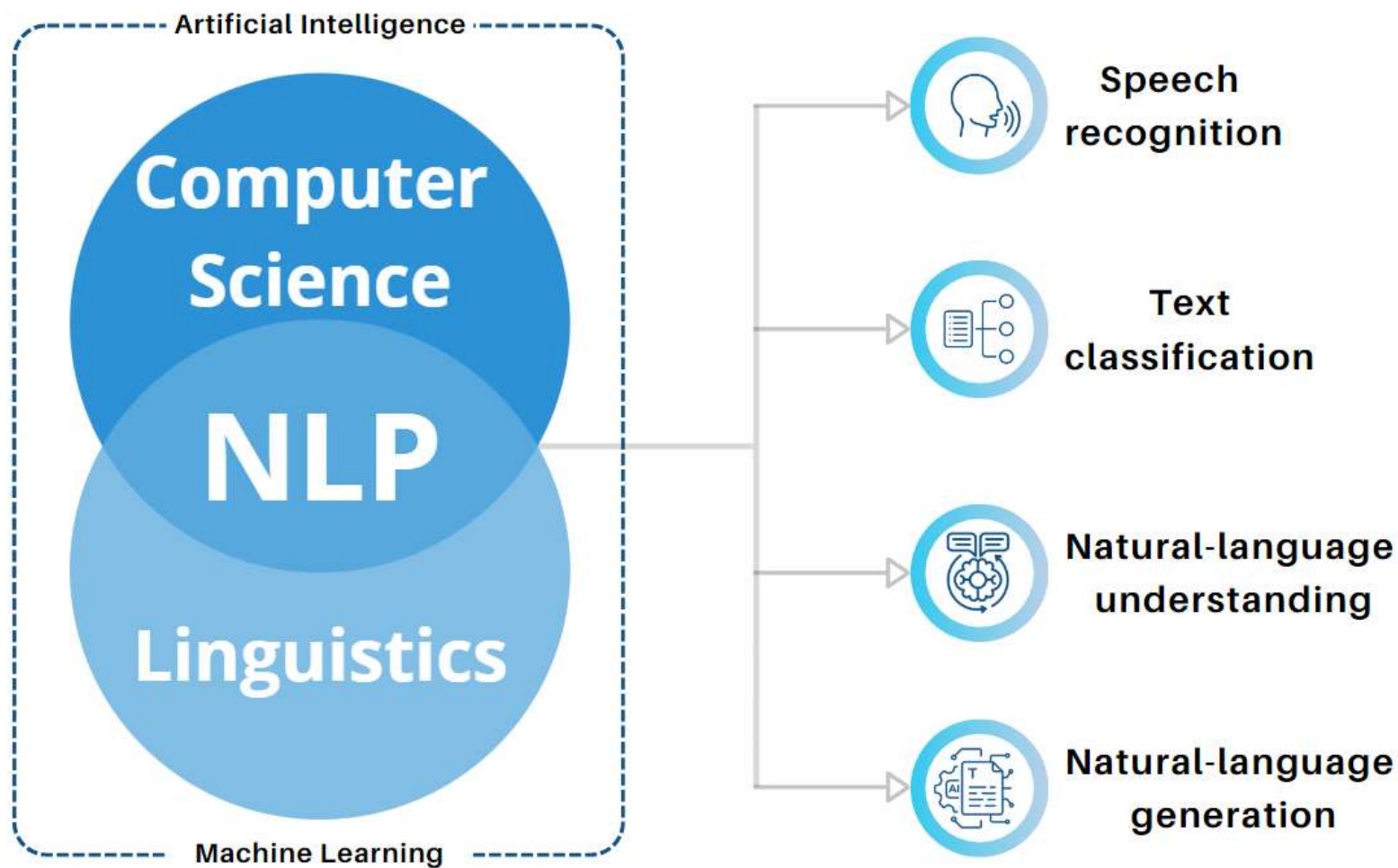
**European Network
of Cancer Registries**

Utilizing Natural Language Processing in eKRN+ application to support cancer registration.

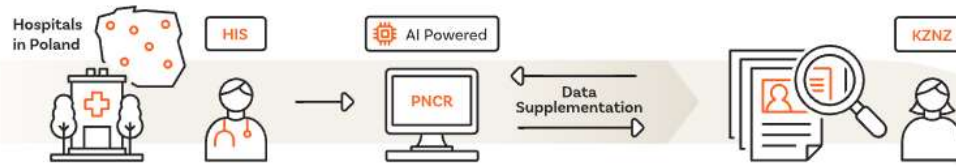
Klaudia Barańska

Polish National Cancer Registry, Maria
Sklodowska-Curie National Research
Institute of Oncology





eKRN+



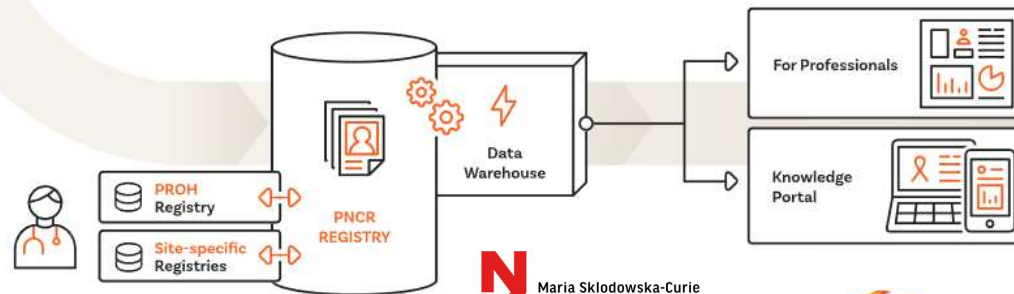
01 DATA SOURCES FOR PNCR
Data on new cancer cases can be reported to the Polish National Cancer Registry and site-specific registries both by physicians, through PNCR platform, and be directly downloaded from the PNCR-integrated hospital information systems (HIS).

02 DATA VERIFICATION
Reported/downloaded data is verified for its correctness and compliance with registration standards. Missing data are being filled in and the whole record registered or added to previously collected data.

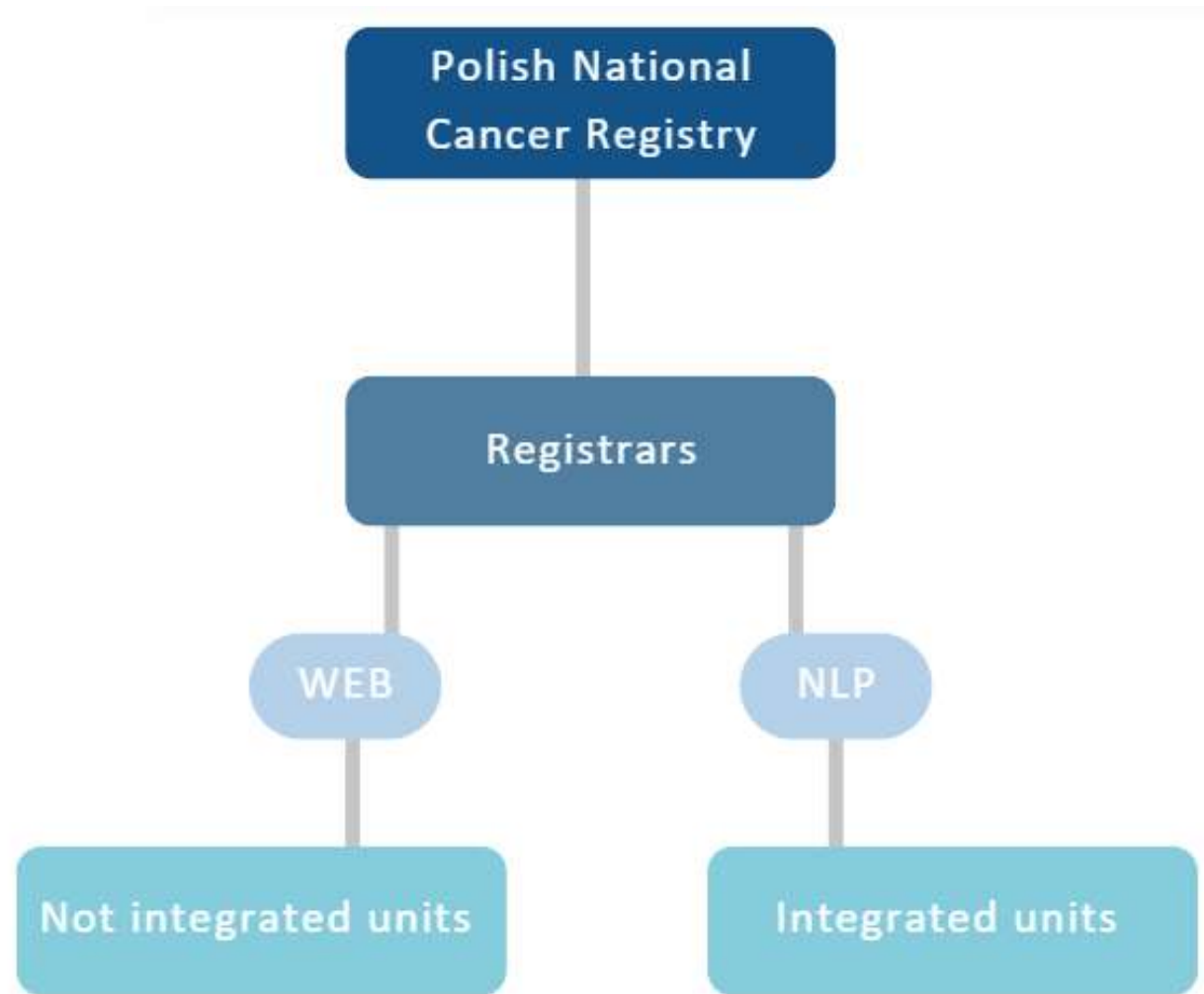
AFTER VERIFICATION DATA IS TRANSFERRED TO CENTRAL DATABASE OF REGISTRIES

03 CENTRAL DATABASE OF REGISTRIES
Central Database of Registries is a technical environment, where all cancer registries are gathered (both PNCR and site-specific registries). Thanks to such solution data is not duplicated, but the common part of all registers is kept in one place. Due to the unification of forms, the person submitting data on the case simultaneously supplies several registers with a single notification.

04 PORTAL, REPORTS, OPEN ACCESS DATA
The portal is a window on the world of the Polish National Cancer Registry, offering analyzes and reports not only for health professionals, but also presenting studies on cancer epidemiology at the national and voivodship level.



Data flow



List of fields in cancer registration form

Fields searched using the NLP mechanism

Metadata

Patient

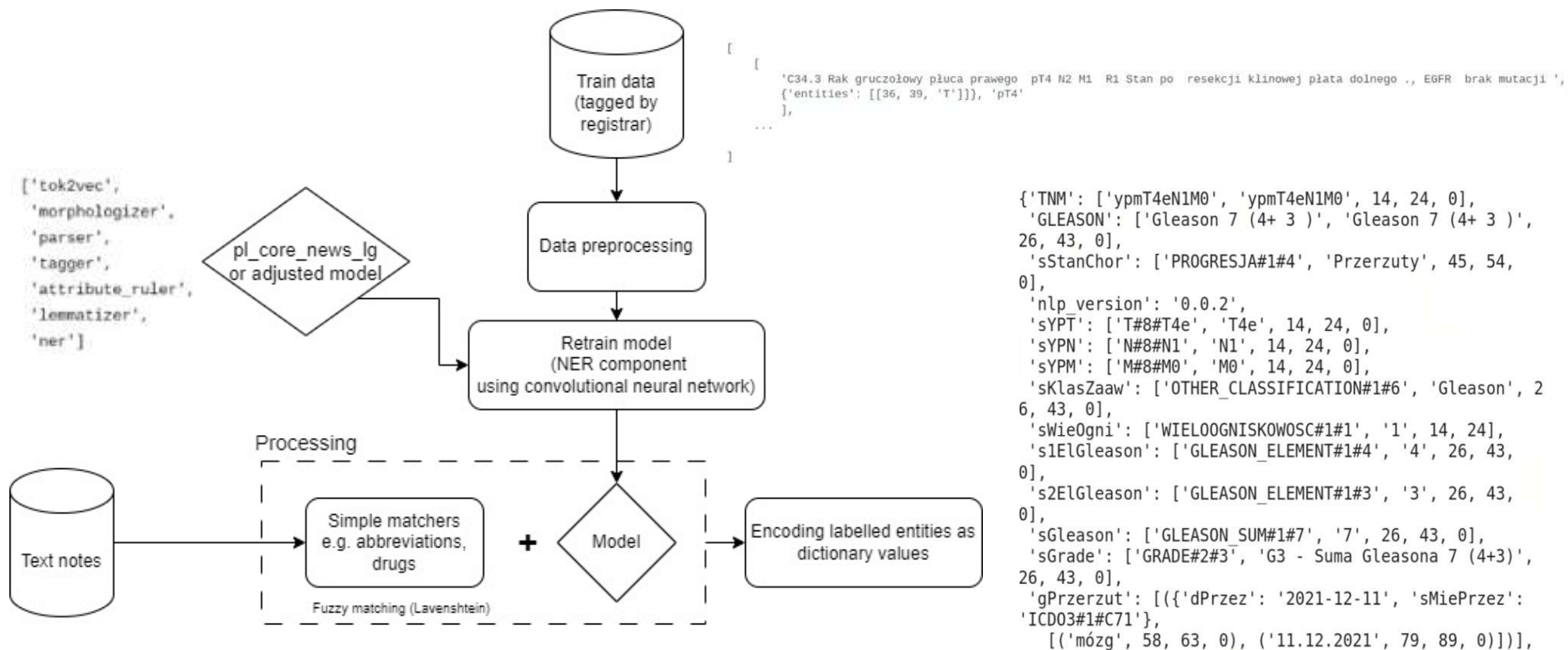
1. Personal ID (PESEL)
2. Date of birth
3. Date of death
4. Sex
5. Name of patient
6. Territorial code (place of residence of patient)
7. Address (zip code, street)

Medical unit

8. Reporting unit: address, REGON
9. User ID (doctor)
10. Doctor's medical license (PWZ)
11. Date of notification

12. Outpatient clinic - date of visit
13. Hospital - date of admission, date of discharge
14. Date of incidence
15. ICD-10 topography code (C00-D09)
16. ICD-O-3 (topography)
17. Morfology code (8000/3,..., 9989/3), date
18. Grade
19. ECOG (1-5)
20. Recurrence/Progression date (local recurrence, systemic recurrence, distant metastases at diagnosis)
21. Metastasis topography code, date
22. Laterality (left, right, both)
23. cTNM, pTNM, ypTNM
24. Stage (I-IV, abc)
25. Biomarkers (p16, HPV, EBV, BRCA1, BRCA2, HER2, ER, PR)
26. Basis of diagnosis (DCO, Clinical investigation, diagnostic procedures, specific markers, cytology, histology, cytogenetic/molecular test)
27. Treatment method, date (chemotherapy, radiotherapy, surgery, hormone therapy, immunotherapy, targeted therapy, transplantation)
28. Treatment intension (radical, palliative, symptomatic, in observation, referred for treatment, refusal of treatment)

Polish National Cancer Registry NLP solution



Case study

Wszytkie Zgłoszenia Karty cząstkowe

Szukaj X Q PESEL: ✓

Status Data rozpoznania ICD-10 Morfologia ICD-O-3 Źródło REGON HIS

Nieprzetworzone karty

S...	D...	I...	Mo...	Stro...	Źr...	R...	S...	HIS
Do weryfi...	2021-02-26	C50.3	85003	Strona prawa	Moduł HIS	000288366	Nieudost...	TAK

Liczba elementów na stronie 100 1 - 1 z 1

NLP

Szukaj X Q

Tag	Klucz	Wartość	Źródło	Do sprawdzenia
Kod stopnia histolo...	sGrade	G2	G2	NIE

Notatki

Rodzaj notatki	Data notatki	ID notatki
WYPIS	2021-03-08	4502700
WYWIAD	2021-03-08	4502701

Zaznaczony fragment: [G2](#)

Treść notatki [EDYCJA TAGU]

Ewentualne inne dolegliwości:

Wywiady: Stan po mastektomii prawostronnej z powodu ca ductale inwazywnym [G2](#). X. 2007.
 Stan po [chirurgii](#) uzup. 3 X FEC-3 X TAX, od 10.2007
 Stan po [rtg](#) pooperacyjnej. zak. 05.2008 do dawki 45 [Gx](#). na obszar bliższy po mastektomii i w. [chł.](#) nadobojczykowopachowych prawych
 Stan po hormonoterapii [arozoletem](#)
 Stan po leczeniu [herceptyna](#) 06.2008 - 06.2009
 Tamoxifen przez 10 lat - leczenie odstawione 7.01.2019 r.

Choroba niedokrwienne serca.
 Neuralgia n. V - stan po leczeniu zabiegowym
 Przebieg dotychczasowego leczenia: j.w.
 Dotychczasowe leczenie onkologiczne: j.w.
 Dotychczasowa diagnostyka: EMG (20.02.19 r.) - przewlekła polineuropatia czuciowo-ruchowa demielinizacyjno-aksonalna.
 Przebyte inne choroby:.
 Przebyte operacje: Stan po operacjach neurochirurgicznych (neuralgia n V prawego w przebiegu konfliktu nerwowo-naczyniowego powikłaną [płytkotokiem](#) nosowym - odbarczenie nerwu 16.03.2016 r. W trakcie hospitalizacji stwierdzono krwiaka przymózgowego - ewakuowany 2.04.2016 r.)
 Stosowane leki: [Toluca](#) 40 mg, x 1 W
[Tolucombi](#) 40 + 12,5 - 1/2 tabl R
[Riocermark](#) 2x2,5 mg

Edytor przypadku

zgłoszenie karta cząstkowa karta wynikowa

Numer dokumentu 10365785 Status karty Do weryfikacji 0.1.15 Wersja modelu NLP 0.1.15

type (C50_)Duct adenocarcinoma NOS;Duct carcinoma NOS;Ductal carcinoma NOS;Duct cell carcinoma;Infiltrating duct adenocarcinoma (C50_)Basal-like carcinoma of breast (C50_)Adenocarcinoma of mammary gland type;Adenocarcinoma of anogenital mammary-like glands;Carcinoma of male breast (C50_)

Typ histologiczny (opis)

Kod stopnia histologicznej złośliwości

G2

Skala ECOG

Skala sprawności

Wznova / Przerzut

Stan choroby

Case study

Wszytkie Zgłoszenia Karty cząstkowe

Szukaj [] X Q PESEL ✓

ICD-10 Morfologia ICD-Q-3

Zródło REGON HIS

Nieprzetworzone karty

S...	D...	I...	Mo...	Stro...	Żr...	R...	S...	HIS
Do weryf...	2021-02-26	C50.3	85003	Strona prawa	Modul HIS	000288366	Nieudost...	TAK

Liczba elementów na stronie 100 1- 1 z 1

Notatki

Rodzaj notatki	Data notatki	ID notatki
WYPIS	2021-03-08	4502700
WYWIAD	2021-03-08	4502701

Zaznaczony fragment:

prawe

Treść notatki [EDYCJA TAGU]

Rozpoznanie zasadnicze: Rak piersi **prawe** - stan po leczeniu systemowym.
Poli neuropatia indukowana chemioterapią.
Choroby współistniejące:
Rozpoznanie histopatologiczne:
Epikryza: Pacjentka przyjęta do Oddziału Medycyny Paliatywnej (OMP) z powodu dolegliwości bólowych o charakterze polineuropatii związanych z leczeniem choroby nowotworowej. W wywiadzie rak piersi prawej, stan po mastektomii prawostronnej (X.2007), stan po chemioterapii uzupełniającej (3 X FEC -3 X TAX, od 10.2007), stan po radioterapii pooperacyjnej (zakńczona) 05.2008 do dawki 45 Gy na obszar bliżny po mastektomii i w. c. nadobojczykowo-pachowych prawych), stan po leczeniu herceptyną (06.2008 - 06.2009), stan po hormonoterapii (Anastrozol, Tamoxifen - leczenie zakończone 7.01.2019 r.
Dodatkowo w wywiadzie: choroba niedokrwienna serca, neuralgia n. V - stan po leczeniu zabiegowym .
W chwili przyjęcia do OMP chora zgłaszała dolegliwości bólowe o charakterze neuropatycznym w odcinkach dystalnych kończyn górnych i dolnych (współistniejące zaburzenia czuciowe i ruchowe, nasilone szczególnie w obrębie kończyn górnych) oraz bóle mięśniowe ze współistniejącymi skurczami mięśni w obrębie tylnej części ud i podudzi, szczególnie w nocy, ustępujące po kilkuminutowym poruszaniu się.
W OMP wdrożono leczenie pod postacią dożylnych wlewów kwasu a - lipoonowego (terapia 10 dniowa), kontynuowano stosowane wcześniej koanalgetyk (duloksetyna, karbamazepina), kontynuowano leki intermisyjne. Stosowano fizjoterapię oraz wsparcie psychologiczne. Uzyskano częściową regresję w zakresie dolegliwości polineuropatycznych w dystalnych odcinkach kończyn; skurcze w obrębie mięśni ud i podudzi (przy prawidłowych poziomach potasu, magnezu we krwi) utrzymywały się pomimo stosowania suplementacji magnezu. W nasilonych dolegliwościach bólowych stosowano paracetamol - z dobrym efektem terapeutycznym. Ze względu na tendencję do

Edytor przypadku

zgłoszenie karta cząstkowa karta wynikowa

Numer dokumentu 10365785 Status karty Do weryfikacji 0.1.15 Wersja modelu NLP 0.1.15

Oznaczenie biomarkerów

Charakterystyka przypadku

Strona ciała Strona prawa

Wieloogniskowość

Wersja TNM

T N M

pT pN pM

ypT ypN ypM

Stoień zaawansowania Stadium zaawansowania

Zaakceptuj Odzuć zmiany Zapisz

Case study

Wszytkie Zgłoszenia Karty cząstkowe

Szukaj

PESEL ✓

ICD-10 Morfologia ICD-O-3

Zródło REGON HIS

Nieprzetworzone karty

S...	D...	I...	Mo...	Stro...	Žr...	R...	S...	HIS
Do weryfi...	2021-02-26	C50.3	85003	Strona prawa	Moduł HIS	000288366	Nieudost...	TAK

Liczba elementów na stronie 100 1 - 1 z 1

NLP

pokaż wszystkie

Szukaj

Tag	Klucz	Wartość	Źródło	Do sprawdzenia
Metody leczenia	gMetLecz	Data rozpoczęcia le...	FEC	NIE
Metody leczenia	gMetLecz	Data rozpoczęcia le...	chemioterapia	NIE
Metody leczenia	gMetLecz	Data rozpoczęcia le...	radioterapii	NIE
Metody leczenia	gMetLecz	Data rozpoczęcia le...	hormonoterapii	NIE

Notatki

Rodzaj notatki	Data notatki	ID notatki
WYPIS	2021-03-08	4502700
WYWIAD	2021-03-08	4502701

Zaznaczony fragment:

FEC

Treść notatki [EDYCJA TAGU]

Epikryza: Pacjentka przyjęta do Oddziału Medycyny Paliatywnej (OMP) z powodu dolegliwości bólowych o charakterze polineuropatii związanych z leczeniem choroby nowotworowej. W wywiadzie rak piersi prawej, stan po mastektomii prawostronnej (X. 2007), stan po chemioterapii uzupełniającej (3 X **FEC**-3 X TAX, od 10.2007), stan po radioterapii pooperacyjnej (zakończona 05.2008 do dawki 45 Gy na obszar bliżny po mastektomii i w. chł. nadobojczykowo-pachowych prawych), stan po leczeniu herceptyną (06.2008 - 06.2009), stan po hormonoterapii (Anastrozol) Tamoxifen - leczenie zakończone 7.01.2019 r.

Dodatkowo w wywiadzie: choroba niedokrwienna serca, neuralgia n. V - stan po leczeniu zabiegowym. W chwili przyjęcia do OMP chora zgłaszała dolegliwości bólowe o charakterze neuropatycznym w oddinkach dystalnych kończyn górnych i dolnych (współistniejące zaburzenia czuciowe i ruchowe, nasilone szczególnie w obrębie kończyn górnych) oraz bóle mięśniowe ze współistniejącymi skurczami mięśni w obrębie tylnej części ud i podudzi, szczególnie w nocy, ustępujące po kilkuminutowym poruszaniu się.

W OMP wdrożono leczenie pod postacią dożylnych wlewów kwasu a - lipoonowego (terapia 10 dniowa), kontynuowano stosowane wcześniej koanalgetyki (duloksetyna, karbamazepina), kontynuowano leki internistyczne. Stosowano fizjoterapię oraz wsparcie psychologiczne. Uzyskano częściową regresję w zakresie dolegliwości polineuropatycznych w dystalnych oddinkach kończyn; skurcze w obrębie mięśni ud i podudzi (przy prawidłowych poziomach potasu, magnezu we krwi) utrzymywały się pomimo stosowania suplementacji magnezu. W nasilonych doległościach bólowych stosowano paracetamol - z dobrym efektem terapeutycznym. Ze względu na tendencję do zaparcí stosowano leczenie wspomagające - z dobrym efektem.

W dniu 08.03.2021r. pacjentkę wypisano do domu w stanie ogólnym dobrym z zaleceniami.

Edytor przypadku

zgłoszenie karta cząstkowa karta wynikowa

Numer dokumentu 10365785 Status karty Do weryfikacji Wersja modelu NLP 0.1.15

Rodzaj i metody leczenia

Rodzaj leczenia Paliatywne

Metody leczenia

Nazwa metody leczenia	Data rozpoczęcia leczenia
Chemioterapia	rrrr-mm-dd
Radioterapia	rrrr-mm-dd
Hormonoterapia	rrrr-mm-dd

Szczegóły leczenia nowotworu piersi

Leczenie Mastektomia

Zaakceptuj Odzuć zmiany Zapisz

Case study

Wszystkie Zgłoszenia Karty cząstkowe

Szukaj X Q PESEL: ✓

Status Data rozpoznania
 ICD-10 Morfologia ICD-O-3
 Źródło REGON HIS

Nieprzetworzone karty

S...	D...	I...	Mo...	Stro...	Źr...	R...	S...	HIS
Do weryfi...	2020-12-29	C50.9	85003	Strona lewa	Moduł HIS	000288366	Nieudost...	TAK
Do weryfi...	2020-12-29	C50.9	85413	Strona lewa	Moduł HIS	000288366	Nieudost...	TAK
Do weryfi...	2021-02-09	C50.9	82302	Strona lewa	Moduł HIS	000288366	Nieudost...	TAK
Do weryfi...	2021-02-09	C50.9	85002	Strona lewa	Moduł HIS	000288366	Nieudost...	TAK
Do weryfi...	2021-02-25	C50.9	80003	Strona lewa	Moduł HIS	000288366	Nieudost...	TAK

Liczba elementów na stronę 100 1 - 6 z 6

NLP

Szukaj X Q

Tag	Klucz	Wartość	Źródło	Do sprawdzenia
pT	sPT	T1c	T1c	NIE

Notatki

Rodzaj notatki	Data notatki	ID notatki
INNE	2021-01-22	4459738
BADHIST	2021-01-21	4459737
WYWIAD	2021-01-04	4459733
WYPIS	2021-01-04	4459735
WYWIAD	2021-01-04	4459739
ZABIEG	2020-12-31	4459732

Zaznaczony fragment:
pT1cN(sn)1a

Treść notatki [EDYCJA TAGU]

Brodawka: H20/18324/1/1
 (1) Ok, zabrodawkowa: H20/18324/1/2
 (1) Guz+powieź: H20/18324/1/3, H20/18324/1/4
 (1) Guz+brodawka przekrój: H20/18324/1/5, H20/18324/1/6
 (1) Guz: H20/18324/1/7
 (2) W/wart 1: H20/18324/2/1, H20/18324/2/2
 (2) W/wart 2: H20/18324/2/3
 (2) W/wart 3: H20/18324/2/4
 (2) Tk, tk: H20/18324/2/5

: LW: lek. med. ; P: Materiał po mastektomii z pobraniem węzła wartowniczego

1. Typ histologiczny: Rak przewodowy naciekający - NOS
2. Stopień złośliwości: G1 (2+2+1) wg Nottingham
3. Największy wymiar komponentu naciekającego: 1,8 cm
4. Odsetek utkania in situ: CDIS około 10%
5. Marginesy chirurgiczne mikroskopowo: jak w opisie makroskopowym
6. Radykalność zabiegu: R0
7. Liczba wszystkich węzłów chłonnych z przerzutami: 1 (z naciekaniem tkanki okołowężowej na odległość 0,2 cm)
8. Liczba wszystkich znalezionych węzłów chłonnych: 3
9. Średnica największego przerzutu w węźle chłonnym: 0,8 cm
10. Stopień zaawansowania: **pT1cN(sn)1a**; DIAG: sp. patomorfolog lek. med.

IDZ: H20/18324; DATAREJ: 2020-12-30 13:43; DATAOTRZ: 2020-12-30 13:44; RM: Wynik HP

Edytor przypadku

zgłoszenie karta cząstkowa karta wynikowa

Numer dokumentu 10361268 Status karty Do weryfikacji 0.1.15 Wersja modelu NLP 0.1.15

Charakterystyka przypadku

Strona ciała
Strona lewa

Wielogniskowość

Wersja TNM

T T2	N N0	M M0
pT T1c	pN N1a	pM
ypT	ypN	ypM

Stopień zaawansowania
II

Stadium zaawansowania
In situ

Informacja: Rozpoznanie z zakresu C00-C97 nie dotyczy nowotworów in situ

Conclusion

- This method eliminates the doctor's involvement in filling out the cancer notification form.
- The work lasted 2.5 years, this approach was applied to the centralized database and was also used in the PROH.
- We continue to refine our NLP models to achieve optimal performance, but this remains a challenge due to the inherent complexity of the problem, the limitations of available data, and the scarcity of models tailored to the Polish language.



Thank you!



Contact

Klaudia Barańska
*Data analyst,
PhD candidate in Biomedical Engineering*

klaudia.baranska@nio.gov.pl