

Joanna Grabska-Chrzastowska*, Jan Kulpa**, Urszula Rychlik**

Zastosowanie sieci neuronowych do predykcji przeżycia w przypadku raka jajnika***

1. Wprowadzenie

Rak jajnika zajmuje drugie miejsce co do częstości występowania wśród nowotworów narządu rodowego (po raku szyjki macicy), a powoduje najwięcej zgonów spośród nowotworów złośliwych narządu płciowego.

Rozpoznanie raka jajnika w jego możliwie najwcześniejszym, niezaawansowanym stopniu rozwoju oraz optymalne leczenie w stopniu tzw. wczesnego i późnego zaawansowania to jeden z najważniejszych problemów współczesnej ginekologii. U ponad 70% chorych rozpoznawany jest w II-III-IV stopniach klinicznego zaawansowania. Stąd zaledwie ok. co trzecia chora z zaawansowanym rakiem jajnika przeżywa pięć lat od chwili rozpoznania choroby. Kiedy guz zostaje wykryty w I stopniu klinicznego zaawansowania i odpowiednio leczony – 80% kobiet ma szansę przeżycia 5 lat i więcej.

Celem niniejszej publikacji jest stwierdzenie, w jakim stopniu na podstawie wstępnych badań wykonanych po interwencji chirurgicznej można przewidzieć, które pacjentki mogą nie przeżyć dwóch lat (24 miesięcy), licząc od daty pierwszej operacji. Taka wiedza jest niezbędna do podjęcia decyzji o dodatkowym leczeniu po obowiązkowym I rzucie chemioterapii (6 seriach). Z kolei pacjentki mające pewność przeżycia wystarczy monitorować i nie obciążać dodatkowym bardzo inwazyjnym leczeniem.

Badaniami objęto 156 pacjentek Centrum Onkologii w Krakowie. Brano pod uwagę parametry krwi przed i po chemioterapii. Do wstępnych badań krwi (HB, RBC, HCT i MVC) dochodził parametr uzyskany w klasyfikacji FIGO, a do końcowych wyników krwi po leczeniu kodowany parametr oceny klinicznej. Obliczenia wykonano za pomocą pakietu Statistica Neural Networks.

Uzyskane wyniki w znaczącym stopniu poprawiły jakość prognozy w stosunku do prostej statystyki oceniającej przeżycie albo na podstawie parametru FIGO (International

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie; asior@agh.edu.pl

** Zakład Analityki i Biochemii Klinicznej, Centrum Onkologii, Instytut im. M. Skłodowskiej Curie Oddział w Krakowie

*** Pracę sfinansowano z funduszu badań własnych (umowa AGH nr 10.10.120.39)

Federation of Gynecology and Obstetrics), albo na podstawie oceny klinicznej po leczeniu. Dodatkowo, otrzymane rezultaty dają szansę wykorzystania dwóch otrzymanych sieci do stworzenia systemu oceniającego przeżycie. Połączenie w jedną całość informacji z wstępnego i końcowego badania mogłoby przynieść jeszcze większą pewność predykcji przeżycia pacjentek [1].

2. Opis przeprowadzonych badań medycznych

W przypadku stwierdzenia raka jajnika w pierwszej kolejności wykonuje się zabieg chirurgiczny, mający na celu usunięcie guza lub jego części oraz dokonanie oceny śródoperacyjnej. W celu oceny klinicznego stopnia zaawansowania nowotworów złośliwych opracowano klasyfikację kliniczną (FIGO w 1986 r.) [2].

Federacja FIGO wprowadziła 4 główne stopnie zaawansowania nowotworów:

- I – guz ograniczony do jajników;
- II – guz jednego lub obu jajników z zajęciem miednicy mniejszej;
- III – guz jednego lub obu jajników z przerzutami;
- IV – guz jednego lub obu jajników z odległymi przerzutami.

W każdej z trzech pierwszych grup rozróżnia się po trzy podgrupy a, b i c. Grupa IV jest wewnątrznie nierozróżnialna. W dalszej części artykułu stopnie zaawansowania w powyższej klasyfikacji będą wzmiankowane jako parametry FIGO. Standardowo po interwencji chirurgicznej bez względu na uzyskany parametr FIGO przeprowadza się leczenie chemioterapią tzw. I rzutu, który obejmuje sześć serii leczenia (co dwa tygodnie). Po zakończeniu leczenia na podstawie badania obrazowego (CT, USG lub NMR) [2] i badania palpacyjnego dokonuje się tzw. oceny klinicznej w klasyfikacji według WHO, tj. Światowej Organizacji Zdrowia.

Ocena kliniczna ogranicza się do 4 głównych klas:

- CR (*Complete Response*) – całkowita odpowiedź na leczenie;
- PR (*Partial Response*) – częściowa odpowiedź na leczenie;
- NC (*no change*) – stabilizacja choroby
- PD (*Progressive Disease*) – progres choroby

W każdym momencie leczenia (po operacji, a przed leczeniem chemioterapią, po każdym rzucie chemioterapii i po jej zakończeniu) wykonuje się analizy krwi, z czego naistotniejsze według klinicystów są badania:

- HB – ilość hemoglobiny;
- RBC – ilość czerwonych krwinek;
- HCT – hematokryt;
- MCV – średnia objętość krwinki czerwonej.

Wszystkie przedstawione rezultaty eksperymentów przeprowadzono na podstawie wyników 156 pacjentek Centrum Onkologii w Krakowie.

3. Opis możliwości oferowanych przez prostą analizę statystyczną

3.1. Sposób przedstawienia wyników

W zastosowaniach medycznych zazwyczaj wyniki przedstawia się w formie tabeli 2×2 (tab. 1), w której wpisuje się liczbę pacjentów odpowiadającej jednemu z czterech wymienionych w opisie tabeli przypadków.

Tabela 1
Sposób przedstawienia wyników predykcji przeżycia dla wybranej metody

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	a *	c ***
PRZEŻYCIE < 24 miesiące	b **	d ****

- * a – liczba przypadków 24-miesięcznego przeżycia klasyfikowanych poprawnie przez sieć – przypadki prawdziwie ujemne;
- ** b – liczba przypadków 24-miesięcznego przeżycia klasyfikowanych przez sieć nieprawidłowo – fałszywie ujemne;
- *** c – liczba przypadków przeżycia mniejszego niż 24 miesiące klasyfikowanych przez sieć nieprawidłowo – fałszywie dodatnie;
- **** d – liczba przypadków przeżycia mniejszego niż 24 miesiące klasyfikowanych poprawnie przez sieć – prawdziwie dodatnie.

Jakość rozpoznania, czyli liczbę poprawnych przewidywań oblicza się za pomocą następującego wzoru:

$$\text{liczba poprawnych przewidywań} = (a + d) * 100\% \quad (1)$$

Oprócz jakości predykcji wprowadza się także dwa parametry statystyczne charakteryzujące wartość diagnostyczną zaproponowanego testu. Zdolność testu do przewidywania przeżycia przez pacjentów pewnego okresu czasu nazywamy czułością metody, a umiejętność testu wykrycia osób, które nie przeżyją tego okresu czasu, jej swoistością. Wartości parametrów czułości i swoistości oblicza się według następujących wzorów:

$$\text{czułość} = \frac{d}{d + b} * 100\% \quad (2)$$

$$\text{swoistość} = \frac{a}{a + c} * 100\% \quad (3)$$

3.2. Wyniki statystyczne dla dwóch metod

Zakładając, że wstępne badanie kliniczne oceniane w skali FIGO stanowi podstawę prognozy 24-miesięcznego przeżycia i przyjmując *a priori*, że przypadki IA, IB, IC, IIA, IIB oraz IIC oznaczają pewność przeżycia, a pozostałe przypadki IIIA, IIIB, IIIC i IV roku ją niepomyślnie, można stwierdzić, że jakość takiego wstępnego prognozowania (po zabiegu a przed leczeniem) wynosi 71,8%.

Tabela 2
Wyniki statystyczne na podstawie klasyfikacji FIGO

	PRZEŻYCIE > 24 miesiące STATYSTYKA *	PRZEŻYCIE < 24 miesiące STATYSTYKA
PRZEŻYCIE > 24 miesięcy	56	32
PRZEŻYCIE < 24 miesiące	12	56

* Jakość prognozy – 71,8%.

Szczegółowy rozkład poprawnych i niepoprawnych przewidywań przedstawiono w tabeli 2.

Na podstawie wyników w tabeli 3 można stwierdzić, że w grupie ocenianych pacjentek statystyczne wyliczenie zawiodło przy przewidywaniu przeżycia 44 kobiet, z czego 12, mimo dobrej kategorii w skali FIGO, nie przeżyło 2 lat.

Tabela 3
Wyniki statystyczne na podstawie oceny klinicznej

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	67	21
PRZEŻYCIE < 24 miesiące	16	52

* Jakość prognozy – 76,3%.

Po leczeniu za pomocą standardowej chemioterapii (I rzutu) dokonuje się oceny klinicznej (głównie na podstawie badania obrazowego i palpacyjnego). Jeżeli założymy, że ocena CR (całkowita remisja) jest podstawą optymistycznej prognozy, a pozostałe PR (częściowa remisja), NC (stabilizacja) i PD (postęp choroby) oznaczają brak 24-miesięcznego przeżycia, to takie przewidywanie sprawdza się w 76,3% przypadków (tab. 3). W tym przypadku liczba pacjentek niepoprawnie zakwalifikowanych do danej grupy wynosi 37, w tym

aż 16 (więcej niż poprzednio) należy do grupy wyników fałszywie ujemnych tzn. mimo dobrych rokowań pacjentki zmarły przed upływem 2 lat. Natomiast liczba przypadków fałszywie dodatnich znacznie spadła (do 21).

Przy zastosowaniu jakiegokolwiek innego narzędzia prognozowania, w tym przypadku sieci neuronowych, interesujące są tylko wyniki powyżej wartości oferowanych przez prostą statystykę.

4. Wyniki otrzymane za pomocą sieci neuronowych

4.1. Rezultaty na podstawie wstępnych danych o pacjentkach

Sieci neuronowe, jak wskazują liczne publikacje [3], bardzo dobrze sprawdzają się w zadaniach predykcji. Pakiet programu Statistica oferuje szereg narzędzi ułatwiających uczenie wielu sieci o różnych strukturach i oceny ich jakości. Uczenie jest prowadzone w oparciu o zbiór uczący, zatrzymywane w zależności od wyników obliczonych dla zbioru walidacyjnego a weryfikowane w oparciu o dane testowe. W przypadku 156 wektorów uczących podział zbioru na uczący, walidacyjny i testowy wynosi odpowiednio 78, 39 i 39 przypadków. Jako dane do sieci neuronowej wprowadzono parametr FIGO (kodowany liczbami od 0 do 9) oraz początkowe wyniki badania krwi (HB, RBC, HCT, MCV). Na jedynym wyjściu sieci oczekiwano odpowiedzi binarnej: 0 – prognoza pomyślna (przeżycie 24 miesięcy), 1 – prognoza niepomyślna. Najlepsze wyniki otrzymano dla struktury o 13 neuronach w pierwszej warstwie ukrytej i 6 neuronach w drugiej. W trakcie uczenia sieci (dobór liczby wejść sieci przy pomocy algorytmu genetycznego) optymalną okazała się grupa tylko czterech parametrów wejściowych (FIGO, HB, HCT i MCV). Parametr RBC zwiększał rozmiar sieci nie wprowadzając dodatkowej informacji. Wyniki dla sieci $4 \rightarrow 16 \rightarrow 6 \rightarrow 1$ przedstawiono w tabeli 4. Standardowo przy jednym wyjściu sieci wprowadza się konwencję, że liczby mniejsze lub równe 0,5 traktuje się jako wynik równy zero, a liczby większe od 0,5 jako 1. Taki próg interpretacji wyników nie zawsze jest idealny w stosunku do jakości sieci. Zmiana progu zmienia liczby prognoz fałszywie ujemnych kosztem fałszywie dodatnich a ich suma niekoniecznie pozostaje taka sama, zmieniając ogólny wynik jakości predykcji. W kolejnych tabelach (tab. 5 i 6) pokazano zmianę struktury wyników dla dwóch różnych progów – 0,6 i 0,45. Taka manipulacja interpretacją wyników może być korzystna, gdy próbujemy za wszelką cenę zmniejszyć jedną z grup przypadków fałszywie diagnozowanych przez test (np. przypadki fałszywie ujemne). Często oczekujemy polepszenia parametru swoistości sieci nawet kosztem jego czułości.

Na podstawie wyników w tabeli 6 można stwierdzić, że zmniejszyła się liczba przypadków fałszywie dodatnich kosztem fałszywie ujemnych i kosztem ogólnej jakości sieci.

Analizując wartości liczbowe w tabeli 6 zauważamy, że zmniejszyła się liczba przypadków fałszywie ujemnych (w stosunku do wyników sieci o progu 0,5) niewielkim kosztem przypadków „po przekątnej” tabeli, stąd znakomity wynik jakości predykcji czyli 78,8% przy wysokiej swoistości (89,7%) i przyzwoitej czułości testu (70,5%). Porównanie wszystkich przedstawionych sieci przedstawiono w rozdziale 5 w tabelach 10 i 11.

Tabela 4

Wyniki sieci neuronowej na podstawie wstępnych danych – próg 0,5

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	64	24
PRZEŻYCIE < 24 miesiące	12	56

* Jakość wstępnego prognozowania: 76,9%.

Tabela 5

Wyniki sieci neuronowej na podstawie wstępnych danych – próg 0,6

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	70	18
PRZEŻYCIE < 24 miesiące	26	43

* Jakość wstępnego prognozowania: 71,8%.

Tabela 6

Wyniki sieci neuronowej na podstawie wstępnych danych – próg 0,45

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	62	26
PRZEŻYCIE < 24 miesiące	7	61

* Jakość wstępnego prognozowania: 78,8%.

4.2. Rezultaty sieci na podstawie wyników po leczeniu

Przy przewidywaniu 24-miesięcznego przeżycia pacjentek z rakiem jajnika nie musimy ograniczać się do wstępnych wyników badań, które otrzymujemy przed rozpoczęciem leczenia, gdyż w każdym przypadku stosujemy standardowa chemioterapię czyli sześć serii tzw. I rzutu chemii.

Również wyniki pacjentek zebrane po zakończeniu wstępnego leczenia mogą wpływać na jego dalszy przebieg.

Stąd kolejne próby znalezienia najlepszej sieci neuronowej do predykcji 24-miesięcznego przeżycia na podstawie ocena kliniczny i końcowych parametrów krwi po chemioterapii I rzutu (HB, RBC, HCT, MCV).

Wyniki najlepszej otrzymanej sieci o strukturze $5 \rightarrow 13 \rightarrow 8 \rightarrow 1$ przedstawiono w tabeli 7. Wynik jakości predykcji przekracza rezultat otrzymany przy zastosowaniu prostej statystyki (76,3%).

Tabela 7
Wyniki sieci neuronowych na podstawie wyników
po zakończeniu leczenia – próg 0,5

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	75	13
PRZEŻYCIE < 24 miesiące	18	50

* Jakość wstępnego prognozowania po zakończeniu leczenia (I rzutu chemioterapii): 80,1%.

Próba zmiany progu z 0,5 na wartość 0,4 (tab. 8) nie polepszyła jakości sieci, ale zrównoważyła czułość testu z jego swoistością (zmniejszenie różnic między liczbą przypadków fałszywie dodatnich i fałszywie ujemnych).

Tabela 8
Wyniki sieci neuronowych na podstawie wyników
po zakończeniu leczenia – próg 0,4

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA *	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	71	17
PRZEŻYCIE < 24 miesiące	14	54

* Jakość wstępnego prognozowania po zakończeniu leczenia (I rzutu chemioterapii): 80,1%.

Ciekawy wynik otrzymano dla progu 0,7 (tab. 9). Do 82,1% zwiększyła się jakość prognozy, a czułość wzrosła gwałtownie, niestety kosztem swoistości (tylko 6 przypadków źle zakwalifikowanych w grupie pierwszej, ale aż 22 błędne prognozy w grupie drugiej).

Tabela 9
Wyniki sieci neuronowych na podstawie wyników
po zakończeniu leczenia – próg 0,7

	PRZEŻYCIE > 24 miesiące SIEĆ NEURONOWA*	PRZEŻYCIE < 24 miesiące SIEĆ NEURONOWA
PRZEŻYCIE > 24 miesięcy	82	6
PRZEŻYCIE < 24 miesiące	22	46

* Jakość prognozowania po zakończeniu leczenia (I rzutu chemioterapii):
82,1%.

5. Porównanie wyników

Aby móc porównać wszystkie sieci i zanalizować ich wyniki, wszystkie ważne wskaźniki jakości przedstawiono w tabelach 10 i 11. Dla wstępnych wyników badań za pomocą sieci neuronowych otrzymano znacznie lepsze wyniki niż wynikające z prostego testu statystycznego – dodatkowo aż 11 ze 156 pacjentek zostało poprawnie sklasyfikowanych do jednej z dwóch grup. Oznacza to, oprócz niewątpliwego sukcesu, również i to, że wyniki krwi w pewnym stopniu determinują możliwość przeżycia 24 miesięcy od operacji.

Tabela 10
Wyniki jakości poszczególnych metod na podstawie wstępnych badań

	Liczba prawidłowych prognoz [%]	Czułość metody [%]	Swoistość metody [%]
Statystyka na podstawie FIGO	71,8	63,6	82,4
Sieć z progiem 0,5	76,9	72,7	82,4
Sieć z progiem 0,45	78,8	70,5	89,7
Sieć z progiem 0,4	76,9	63,6	94,1

Tabela 11
Wyniki jakości poszczególnych metod na podstawie badań po zakończeniu leczenia

	Liczba prawidłowych prognoz [%]	Czułość metody [%]	Swoistość metody [%]
Statystyka na podstawie oceny klinicznej	73,8	76,1	76,5
Sieć z progiem 0,5	80,1	85,2	73,5
Sieć z progiem 0,4	80,1	80,7	79,4
Sieć z progiem 0,7	82,1	93,2	67,6

W przypadku analizy sieci neuronowych dla końcowych wyników po zakończeniu chemioterapii to dla sieci o najlepszej jakości rozpoznania (próg 0,7) otrzymujemy poprawę przewidywań w stosunku do 12 kobiet, a dla sieci bardziej zrównoważonej w stosunku do wiarygodności testu, o 10 pacjentek.

6. Wnioski

W pracy przedstawiono udaną próbę zastosowania sieci neuronowej do predykcji 24-miesięcznego przeżycia wśród pacjentek objętych leczeniem onkologicznym po przebyciu operacji związanej z rozpoznaniem raka jajnika. W obu omówionych przypadkach wykorzystania wstępnych wyników badań i wyników po zakończeniu leczenia chemioterapią I rzutu otrzymano lepsze wyniki niż wynikające z prostej statystyki czy to łączącej parametr FIGO czy ocenę kliniczną z 2-letnim przeżyciem. Dodatkowo z wyników dla poszczególnych sieci można mieć nadzieję na uzyskanie jeszcze lepszej prognozy na podstawie zespołu sieci wykorzystującego oba rodzaje dostępnej informacji równocześnie. Manipulacja wartością progu interpretacji wyników dla każdej z sieci wchodzącej w skład takiego zespołu mogłaby prowadzić do zwiększonej jakości predykcji.

Literatura

- [1] Grabska-Chrzastowska J.W., Tomalak W.: *Zastosowanie neuronowego systemu wnioskującego do klasyfikacji wyników badania spirometrycznego*. Materiały konferencyjne „Sztuczna Inteligencja w Inżynierii Biomedycznej”, Kraków 2004
- [2] Markowska J.: *Rak jajnika*. Warszawa, Springer PWN 1997
- [3] Tadeusiewicz R.: *Sieci neuronowe*. Warszawa, PWN 1993