

(19)



URZĄD
PATENTOWY
RZECZYPOSPOLITEJ
POLSKIEJ

(10) **PL 243444 B1**

(12)

Opis patentowy

(21) Numer zgłoszenia: **435967**

(22) Data zgłoszenia: **2020.11.12**

(43) Data publikacji o zgłoszeniu: **2022.05.16 BUP 20/2022**

(45) Data publikacji o udzieleniu patentu: **2023.08.28 WUP 35/2023**

(51) MKP:

C12Q 1/6886 (2018.01)

- | |
|--|
| (73) Uprawniony z patentu:
UNIWERSYTET MEDYCZNY W BIAŁYMSTOKU,
Białystok, PL |
| (72) Twórca(-y) wynalazku:
MAGDALENA NIEMIRA, Białystok, PL
ANNA EROL, Białystok, PL
ADAM JACEK KRĘTOWSKI, Białystok, PL |
| (74) Pełnomocnik:
Marta Kawczyńska, Warszawa, PL |

(54) Tytuł:

Diagnostyczny zestaw biomarkerów miRNA do diagnostyki raka jajnika, sposób diagnostyki *in vitro* raka jajnika, zastosowania takiego zestawu biomarkerów miRNA oraz zestaw testowy do diagnostyki *in vitro* raka jajnika

PL 243444 B1

Opis wynalazku

Przedmiotem wynalazku jest diagnostyczny zestaw biomarkerów miRNA, diagnostyczny zestaw biomarkerów miRNA do zastosowania do diagnostyki raka jajnika, sposób diagnostyki *in vitro* raka jajnika, zastosowania zestawu biomarkerów miRNA w diagnostyce *in vitro* raka jajnika oraz zestaw testowy do diagnostyki *in vitro* raka jajnika.

Dziedzina wynalazku

Niniejszy wynalazek dotyczy ogólnie molekularnej diagnostyki klinicznej, dokładniej molekularnej diagnostyki raka jajnika, w tym nieinwazyjnego testu diagnostycznego, tzw. płynnej biopsji, o wysokiej czułości i swoistości, zwłaszcza do wczesnej diagnostyki raka jajnika, poprzez pomiar, analizę i/lub monitorowanie ekspresji mikro RNA (określanych tu także jako miRNA) w próbkach biologicznych, takich jak surowica krwi, w szczególności z wykorzystaniem klasyfikującego modelu diagnostycznego.

Stan techniki

Rak jajnika jest jednym z najczęstszych nowotworów złośliwych żeńskich narządów płciowych oraz główną przyczyną umieralności z powodu tego rodzaju nowotworów w krajach rozwiniętych. Obecnie diagnostyka raka jajnika oparta jest na wykonaniu badania dwuręcznego miednicy, oznaczaniu stężenia antygenu CA 125 i przezpochwowym badaniu ultrasonograficznym. Jednakże oszacowano, że badanie przedmiotowe przeprowadzone u kobiet bez objawów klinicznych pozwala na wykrycie zaledwie 1 na 10000 raków jajnika. Badanie radioimmunologiczne w kierunku swoistego dla nowotworu antygenu CA125 ujawnia zwiększenie jego stężenia u 80% chorych na raka jajnika, jednakże stężenie to może być większe również w przebiegu chorób nienowotworowych, co ogranicza jego swoistość. Badanie ultrasonograficzne jest nie tylko kosztowne, ale również cechuje się ograniczoną swoistością i czułością. Poza tym, podstawowym problemem w diagnostyce raka jajnika jest najczęściej zupełny brak objawów we wczesnym stadium, a w późniejszych stadiach, gdy pojawiają się już przerzuty, objawy często nie są charakterystyczne i związane z układem pokarmowym, przez co diagnostyczne badania przesiewowe są niezwykle istotne. Pomimo wielu badań wciąż brakuje wiarygodnych biomarkerów diagnostycznych oraz metod wczesnego wykrywania tej choroby, jak również środków do monitorowania jej leczenia i/lub progresji, w tym wczesnego wykrywania ewentualnej wznowy. Wartość diagnostyczna szeroko stosowanej ultrasonografii przezpochwowej oraz oznaczenie w surowicy antygenu CA125 okazała się niewystarczająca z powodu zbyt niskiej czułości i swoistości.

Obecnie Amerykańska Agencja Żywności i Leków (FDA) nie zaleca badań przesiewowych do wczesnego wykrywania raka jajnika samych w sobie, i ostrzega przed istniejącymi dotychczas testami, ponieważ w jej opinii nie są one wiarygodne i mogą one wprowadzać w błąd pacjentów i lekarzy poprzez duży odsetek wyników fałszywie negatywnych, bądź fałszywie pozytywnych. FDA podkreśla jednocześnie, że w przypadku innych nowotworów istnieją skuteczne testy przesiewowe, jednak nie istnieją jeszcze takie testy w przypadku raka jajnika. Co najważniejsze, w podanym oświadczeniu FDA wyraźnie podkreśla, że dobry test przesiewowy w przypadku raka jajnika jest bardzo potrzebny, ze względu na zwykle zbyt późne diagnozowanie. Uruchomiono nawet łatwy dostęp naukowców do biobanku z próbkami krwi od pacjentek z rakiem jajnika, aby przyspieszyć badania w poszukiwaniu skutecznego biomarkera diagnostycznego i sposobu diagnozowania raka jajnika, w szczególności we wczesnym stadium. Badania w kierunku stworzenia czułego i swoistego testu opartego na nieinwazyjnych biomarkarach jest zatem nagłą potrzebą w diagnostyce raka jajnika. Ze względu na to, że jajniki są narządami położonymi całkowicie w jamie otrzewnej, obecnie niemożliwe jest zdiagnozowanie raka jajnika bez chirurgicznej resekcji. Dodatkowo, ze względu na możliwość łatwego rozsiewu komórek nowotworowych, należy również unikać wykonywania biopsji cienkoigłowej. W związku z tym, pilnie potrzebne są nieinwazyjne biomarkery, które mogłyby wesprzeć dotychczasowe metody, takie jak USG przezpochwowe oraz badanie poziomu CA125 (markera o czułości zaledwie 40%).

Wiadomo, że ekspresja mikro RNA (miRNA) może występować w tkance nowotworowej na innym poziomie niż w tkance prawidłowej. Znane są różne miRNA i znane jest zastosowanie analizy ekspresji wielu różnych mikro RNA do diagnostyki różnego rodzaju nowotworów. miR-1246 jest stosowany do diagnostyki nowotworów płuc, jamy ustnej, szyjki macicy, czy prostaty (zobacz np. Liao i wsp., Expression and Clinical Significance of microRNA-1246 in Human Oral Squamous Cell Carcinoma (2015) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4371709/>, Zhang i wsp., Tumour-initiating cell-specific miR-1246 and miR-1290 expression converge to promote non-small cell lung cancer progression, Nature Communications (2016)). Sekwencja miR-1246 dostępna jest w bazach danych, np.

www.mirbase.org pod numerem dostępu: MIMAT0005898. Wiadomo również, że ekspresja tego miRNA jest zmieniona w surowicy pacjentów z nowotworem żołądka. Jednakże, chociaż ekspresja miR-1246 może być zmieniona w surowicy u pacjentów z różnymi typami nowotworów, nie wykazano takiego zjawiska w przypadku raka jajnika. Należy mieć przy tym również na względzie to, że fakt, iż dany typ cząsteczki miRNA występuje na zmienionym poziomie u pacjentów z jednym typem nowotworu, nie wyklucza możliwości, że tego rodzaju zależność nie będzie występowała w przypadku innego typu nowotworów, ani tego, że dopiero w odpowiedniej kombinacji z innym biomarkerem lub biomarkerami, jego ekspresja będzie ewentualnie dostarczać wskaźnika diagnostycznego, a jej pomiar będzie wchodził w skład skutecznego nieinwazyjnego testu dla innego, specyficznego nowotworu, np. raka jajnika.

Mikro RNA 150-5p (miR-150-5p) jest również znany. Sekwencja miRNA dostępna jest w bazach danych, np. www.mirbase.org pod numerem dostępu: MIMAT0000451. Wiadomo, że miR-150-5p jest zaangażowany w powstawanie wielu nowotworów. Przykładowo, obniżony poziom ekspresji miR-150-5p został wykryty w próbkach tkanek (ale nie we krwi czy surowicy) od pacjentów z rakiem trzustki w porównaniu do tkanek zdrowych (Zhonghua Bing Li Xue Za Zhi., lipiec 2013;42(7):460-4. doi:10.3760/cma.j.issn.0529-5807.2013.07.007.). Jednakże, jak dotąd nie wykazano związku pomiędzy ekspresją miR-150-5p a rakiem jajnika, ani samodzielnie, ani w kombinacji z jakimkolwiek innym biomarkerem lub biomarkerami.

Celem niniejszego wynalazku jest zatem dostarczenie czułego i swoistego biomarkera diagnostycznego i prognostycznego do diagnostyki raka jajnika i monitorowania jego leczenia, jak również do przewidywania jego wznowy po zakończonym leczeniu. Celem niniejszego wynalazku jest także dostarczenie nieinwazyjnego sposobu do diagnostyki raka jajnika, charakteryzującego się dużą swoistością i czułością oraz testu diagnostycznego nadającego się do diagnostycznych badań, w tym przesiewowych pod kątem tego nowotworu.

Krótki opis wynalazku

Przedmiotem wynalazku jest diagnostyczny zestaw biomarkerów miRNA składający się z miR-1246 oraz miR-150-5p. Taki zestaw biomarkerów miRNA przeznaczony jest do zastosowania w diagnostowaniu raka jajnika.

Zgodnie z wynalazkiem opisany został także diagnostyczny zestaw biomarkerów miRNA obejmujący 2 biomarkery miRNA: miR-1246 oraz miR-150-5p.

Przedmiotem wynalazku jest także diagnostyczny zestaw biomarkerów miRNA składający się z miR-1246 oraz miR-150-5p do zastosowania do diagnostyki raka jajnika.

Przedmiotem wynalazku jest sposób diagnozowania *in vitro* raka jajnika u osobnika, znamieny tym, że obejmuje następujące etapy:

- i) oznaczania poziomów ekspresji zestawu dwóch biomarkerów miRNA: miR-1246 i miR-150-5p w próbce od osobnika;
- ii) porównania poziomów ekspresji oznaczonych w etapie i) z poziomami ekspresji miR-1246 i miR-150-5p u osobnika bez raka jajnika, przy czym porównanie dostarcza wskaźnika diagnostycznego określającego czy osobnik ma raka jajnika, przy czym wzrost poziomu ekspresji miR-1246 względem poziomu ekspresji miR-1246 u osobnika bez raka jajnika i spadek poziomu ekspresji miR-150-5p względem poziomu ekspresji miR-150-5p u osobnika bez raka jajnika wskazują na raka jajnika u osobnika. Innymi słowy, wskaźnikiem diagnostycznym raka jajnika jest wzrost poziomu ekspresji miR-1246 w badanej próbce względem poziomu ekspresji miR-1246 u osobnika bez raka jajnika i spadek poziomu ekspresji miR-150-5p w badanej próbce względem poziomu ekspresji miR-150-5p u osobnika bez raka jajnika.

Korzystnie w sposobie według wynalazku porównanie poziomów ekspresji przeprowadza się z wykorzystaniem zestawu danych zawierających dane o poziomie ekspresji zestawu markerów miRNA obejmującego miR-1246 i miR-150-5p.

Korzystnie w sposobie według wynalazku poziom ekspresji zestawu biomarkerów miRNA oznacza się z zastosowaniem metody pomiaru ekspresji wybranej spośród ilościowej reakcji odwrotnej transkrypcji z łańcuchową reakcją polimerazy (RT-qPCR), metody NanoString lub mikromacierzy.

Korzystnie w sposobie według wynalazku porównuje się poziomy ekspresji po znormalizowaniu.

Korzystnie w sposobie według wynalazku stosuje się klasyfikujący model diagnostyczny, który na podstawie danych o poziomie ekspresji miR-1246 i miR-150-3p, klasyfikuje próbkę jako próbkę od osobnika z rakiem jajnika lub jako próbkę od osobnika bez raka jajnika. Taki klasyfikujący model diagno-

styczny wykorzystuje optymalne punkty odcięcia (ang. optimal cut-off points) w zależności od zastosowanej metody oznaczenia ekspresji zestawu biomarkerów miRNA, tj. NanoString, mikromacierze, RT-qPCR, z naciskiem na RT-qPCR. W korzystnym sposobie według wynalazku wykorzystuje się zawsze optymalny punkt odcięcia dla zastosowanej metody oznaczenia poziomu ekspresji zestawu biomarkerów miRNA oraz zakres z informacją o swoistości (Sp) i czułości (S). Klasyfikujący model diagnostyczny stosuje w odniesieniu do wyników pomiaru ekspresji zestawu biomarkerów miRNA według wynalazku przeprowadzonego z zastosowaniem metody RT-qPCR optymalny punkt odcięcia w zakresie 0,1–0,8; w przypadku metody NanoString punkt odcięcia w zakresie 0,3–0,9, zaś w przypadku mikromacierzy wartość optymalnego punktu odcięcia wynosi 0,5.

Korzystniej do określania poziomu ekspresji miRNA stosuje się metodę RT-qPCR. Metoda ta nadaje się do wykorzystania na dużą skalę w laboratoriach diagnostycznych, gdyż jest dużo tańsza niż metoda NanoString czy mikromacierze, a ponadto wymaga znacznie mniej materiału do badania i pozwala uzyskać wyniki znacznie szybciej niż w przypadku metod badań wielkoskalowych jak NanoString czy mikromacierze.

Jeszcze korzystniej, w przypadku pomiaru ekspresji z zastosowaniem metody RT-qPCR do normalizacji wyników stosuje się poziom ekspresji referencyjnego miRNA, korzystnie miR-103-3p i/lub miR-199b-5p.

Jeszcze bardziej korzystnie normalizację wyników uzyskuje się stosując wzór:

$\text{deltaCt} = (\text{Ct miR-1246/miR-150-5p} - \text{Ct miR-103-3p}),$

w którym

delta Ct oznacza zmianę wartości cyklu progowego

Ct oznacza wartość cyklu progowego.

Szczególnie korzystnie uzyskane dane w postaci wartości deltaCt podstawia się do klasyfikującego modelu diagnostycznego, a uzyskany wynik porównuje się z punktem odcięcia wybranym z zakresu 0,1–0,8. Wartość tego punktu jest określona za pomocą trenowania klasyfikującego modelu diagnostycznego na danych o znanym statusie próbek.

Co istotne, w tym korzystnym sposobie według wynalazku nie porównuje się wartości deltaCt, a wartości deltaCt podstawia się do powyższego równania i porównuje z punktem odcięcia ustalonym przez klasyfikujący model diagnostyczny.

Korzystnie w sposobie według wynalazku jako próbkę od osobnika stosuje się próbkę surowicy.

Korzystnie zestaw biomarkerów składający się z miR-1246 i miR-150-5p stosuje się do diagnozowania, zwłaszcza niskozróżnicowanego surowiczego raka jajnika.

Przedmiotem wynalazku jest także zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono powyżej, tj. składającego się z miR-1246 i miR-150-5p, do diagnozowania *in vitro* raka jajnika.

Przedmiotem wynalazku jest ponadto zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono powyżej w diagnostycznym badaniu przesiewowym *in vitro* pod kątem występowania raka jajnika.

Przedmiotem wynalazku jest również zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono powyżej do oceny *in vitro* skuteczności leczenia raka jajnika.

Przedmiotem wynalazku jest także zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono powyżej do monitorowania *in vitro* odpowiedzi na leczenie raka jajnika.

Przedmiotem wynalazku jest ponadto zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono powyżej do przewidywania wznowy raka jajnika po ukończonym leczeniu raka jajnika.

Przedmiotem wynalazku jest ponadto zestaw testowy do diagnostyki *in vitro* raka jajnika, charakteryzujący się tym, że zawiera środki do ilościowego oznaczania poziomu ekspresji zestawu dwóch biomarkerów miRNA: miR-1246 i miR-150-5p oraz instrukcje do przeprowadzenia sposobu według wynalazku jak określono powyżej.

Korzystnie zestaw testowy do diagnostyki według wynalazku jako środki do ilościowego oznaczania ekspresji zestawu biomarkerów miRNA: miR-1246 i miR-150-5p zawiera odczynniki i startery do powielania w reakcji RT-qPCR miR-1246 i miR-150-5p.

Korzystniej zestaw testowy do diagnostyki według wynalazku zawiera ponadto środki do ilościowego oznaczania ekspresji referencyjnego miRNA, korzystnie zestawu miR-103-3p i/lub miR-199b-5p, takie jak startery i odczynniki odpowiednie do powielania metodą RT-qPCR.

Szczegółowy opis wynalazku

Wynalazki według niniejszego zgłoszenia bazują na wyselekcjonowaniu przez twórców wynalazku zestawu biomarkerów mikroRNA, tj. miR-1246 oraz miR-150-5p, który nadaje się do diagnozowania raka jajnika, w szczególności niskozróżnicowanego surowiczego raka jajnika, z wysoką czułością i swoistością.

Wynalazki według niniejszego zgłoszenia rozwiązują problem dotychczasowego braku wiarygodnych, czułych i swoistych biomarkerów oraz sposobów diagnostycznych je wykorzystujących do diagnozowania raka jajnika, w szczególności niskozróżnicowanego raka jajnika, i w konsekwencji pozwalają na diagnozę tej choroby nowotworowej, korzystnie nawet wczesną diagnozę raka jajnika, tj. przed wystąpieniem objawów klinicznych tej choroby. Dodatkowo, wyniki klasyfikacji diagnostycznej uzyskane zgodnie z niniejszymi wynalazkami są stabilne również względem zaawansowania choroby, co zostało potwierdzone na zewnętrznych danych. Ta cecha wynalazków według niniejszego zgłoszenia pozwala na ich szersze zastosowanie diagnostyczne, w szczególności w badaniu przesiewowym pod kątem występowania raka jajnika.

Diagnostyczny zestaw biomarkerów miRNA według wynalazku składa się z miR-1246 oraz miR-150-5p pozwala na swoistą i czułą diagnostykę raka jajnika, zwłaszcza niskozróżnicowanego surowiczego raka jajnika.

Diagnostyczny zestaw biomarkerów miRNA według wynalazku znajduje ponadto zastosowanie zgodnie z wynalazkiem w diagnostycznym badaniu przesiewowym *in vitro* pod kątem występowania raka jajnika.

Diagnostyczny zestaw biomarkerów miRNA według wynalazku znajduje również zastosowanie zgodnie z wynalazkiem w sposobie oceny skuteczności *in vitro* leczenia raka jajnika.

Diagnostyczny zestaw biomarkerów miRNA według wynalazku znajduje także zastosowanie zgodnie z wynalazkiem do monitorowania odpowiedzi na leczenie raka jajnika.

Diagnostyczny zestaw biomarkerów miRNA według wynalazku znajduje ponad to zastosowanie zgodnie z wynalazkiem do przewidywania wznowy raka jajnika po ukończonym leczeniu raka jajnika.

W takich zastosowaniach diagnostycznego zestawu biomarkerów miRNA według wynalazku jak określono powyżej zestaw składa się z biomarkerów miR-1246 i miR-150-5p.

Zatem przedstawione tu wynalazki mogą być stosowane nie tylko do diagnostyki raka jajnika, ale także do monitorowania skuteczności leczenia raka jajnika, zarówno w trakcie jego trwania, jak również po zakończeniu leczenia. W takim przypadku próbka do badania pochodzi od osobnika poddanego leczeniu i/lub po zakończonym leczeniu i w określonych odstępach czasu oznacza się poziom ekspresji diagnostycznego zestawu biomarkerów miRNA w próbce od pacjenta i porównuje się z odpowiednimi danymi dla tego pacjenta uzyskanymi wcześniej, tzn. na etapie rozpoznania raka jajnika i/lub na wcześniejszym etapie leczenia. Zmiana poziomu ekspresji diagnostycznego zestawu biomarkerów miRNA według wynalazku względem poziomu oznaczonego wcześniej dostarczy wskaźnika diagnostycznego pozwalającego na stwierdzenie czy zastosowane leczenie jest skuteczne. Wskaźnikiem diagnostycznym raka jajnika korzystnie jest wzrost poziomu ekspresji miR-1246 względem poziomu ekspresji miR-1246 u osobnika bez raka jajnika i spadek poziomu ekspresji miR-150-5p względem poziomu ekspresji miR-150-5p u osobnika bez raka jajnika. Zmiana poziomów ekspresji diagnostycznego zestawu biomarkerów miRNA według wynalazku w trakcie leczenia w kierunku poziomów ekspresji tego zestawu biomarkerów u osobników bez raka jajnika wskazuje na skuteczność leczenia raka jajnika. Po zakończonym pomyślnie leczeniu poziom ekspresji diagnostycznego zestawu biomarkerów miRNA według wynalazku powinien być zasadniczo taki jak u osobnika bez raka jajnika. Stwierdzenie zatem wzrostu poziomu ekspresji miR-1246 względem poziomu ekspresji miR-1246 u osobnika bez raka jajnika i spadku poziomu ekspresji miRNA-150-5p względem poziomu ekspresji miR-150-5p u osobnika bez raka jajnika po zakończonym leczeniu pozwalają stwierdzić występowanie wznowy raka jajnika poprzez zaklasyfikowanie próbki badanej jako próbkę od osobnika z rakiem jajnika.

Sposób diagnozowania *in vitro* raka jajnika według wynalazku jest nie tylko czuły i swoisty, ale także szybki i nieinwazyjny, ponieważ wymaga jedynie zwykłego pobrania próbki krwi w celu izolacji surowicy.

Do realizacji sposobu według wynalazku potrzebna jest próbka od osobnika. Taką próbką jest korzystnie próbka surowicy. Próbkę krwi pełnej pobiera się od osobnika do badania w standardowy sposób do probówek bez antykoagulantu (na tzw. skrzep), i następnie izoluje się z niej surowicę do badania zgodnie z wynalazkiem. Dokładniej, po pobraniu, próbkę krwi odstawia się na 30–60 minut,

a następnie wiruje przez 5 min przy 4000 obr/min. Uzyskaną tak próbkę surowicy, przenosi się do czystej od RNAaz próbki celem realizacji sposobu i/lub testu według wynalazku.

Z uzyskanej tak próbki od osobnika izoluje się następnie miRNA z użyciem komercyjnie dostępnych zestawów do izolacji miRNA, postępując zgodnie z zaleceniami ich wytwórców. W przypadku izolacji RNA z płynów ustrojowych, takich jak surowica, trudno jest ocenić wydajność izolacji na podstawie pomiaru spektrofotometrycznego, gdyż ilość materiału jest mała. Dlatego jako metodę kontroli jakości i ilości wyizolowanego materiału można stosować metodę fluorometryczną. Zatem, przy realizacji sposobu według niniejszego wynalazku po izolacji miRNA korzystnie oznacza się stężenie miRNA stosując metodę fluorometryczną, np. z zastosowaniem zestawu Qubit™ microRNA Assay Kit i fluorometru Qubit™ (ThermoFisher Scientific, USA).

Bazując na zmienionych ekspresyjnie cząsteczkach miRNA w grupie badanej z rakiem jajnika w odniesieniu do grupy bez raka jajnika, wyselekcjonowanych metodą NanoString, zastosowano metodę selekcji atrybutów i wytypowano najsilniejszych kandydatów do tworzenia klasyfikującego modelu diagnostycznego i opartego na nim testu diagnostycznego. Z różnych kombinacji wytypowanych miRNA zostały stworzone modele klasyfikujące, z których najlepszą, tj. najbardziej czułą i swoistą kombinacją okazał się właśnie zestaw biomarkerów miRNA według wynalazku obejmujący miRNA: miR-1246 i miR-150-5p. Zastosowanie diagnostycznego zestawu składającego się z tych dwóch biomarkerów miRNA pozwala na czułą i swoistą diagnostykę raka jajnika, nawet we wczesnym jego stadium rozwoju.

W poniższej tabeli 1 podane są parametry dla diagnostycznego zestawu biomarkerów według wynalazku w porównaniu z inną kombinacją biomarkerów miRNA.

Tabela 1. Zmienne w zbudowanych modelach klasyfikujących z różnymi zestawami miRNA.

Współczynniki	Stała a_0	Współczynnik predyktora x_1 (a_1)	Współczynnik predyktora x_2 (a_2)
x_1 = miR-1246 x_2 = miR-150-5p	4,47117	0,07091	-0,31985
x_1 = miR-1246 x_2 = miR-144-3p	-0,94138	0,03202	-0,02179

W poniższej tabeli 2 podane jest porównanie podstawowych parametrów jakości klasyfikacji dla diagnostycznego zestawu biomarkerów według wynalazku w porównaniu z inną kombinacją biomarkerów miRNA. Jak wyraźnie widać parametry jakości, tj. pole pod krzywą, swoistość i czułość, dla zestawu biomarkerów według wynalazku są wyraźnie lepsze niż w przypadku innego zestawu biomarkerów.

Tabela 2. Parametry jakości dla klasyfikujących modeli diagnostycznych na zbiorach treningowym i testowym dla pomiarów ekspresji metodą Nanostring.

Nazwa	miR-1246, miR-150-5p	miR-1246, miR-150-5p	miR-1246, miR-144-3p	miR-1246, miR-144-3p
Zbiór	Treningowy	Testowy	Treningowy	Testowy
Pole pod krzywą (AUC)	98,6%	100%	93,9%	95,2%

Nazwa	miR-1246, miR-150-5p	miR-1246, miR-150-5p	miR-1246, miR-144-3p	miR-1246, miR-144-3p
Zbiór	Treningowy	Testowy	Treningowy	Testowy
Przedział uwności (CI) dolna granica	96,4%	-	85,8%	85,2%
CI górna granica	100%	-	100%	100%
Punkt odcięcia (wskaźnik Youdena)	0,44	0,44	0,56	0,56
Czułość	96,4%	100%	92,9%	92,3%
Swoistość	95,2%	92,3%	95,2%	87,5%
Współczynnik determinacji (R^2)	0,74	-	0,66	-
Średnia kwadratowa błędu (RMSE)	0,23	-	0,29	-

Sposób według wynalazku można przeprowadzić z zastosowaniem dowolnych metod porównywania poziomów ekspresji genów, ale korzystnie przeprowadza się go z zastosowaniem klasyfikujących modeli diagnostycznych stworzonych przez twórców niniejszego wynalazku poprzez wprowadzenie do nich danych, w tym przypadku o poziomach ekspresji zestawu biomarkerów mikroRNA według wynalazku. Zgodnie z niniejszym wynalazkiem opracowano zatem klasyfikujące modele diagnostyczne, których główną rolą jest rozróżnienie przypadków raka jajnika od przypadków bez raka jajnika z wykorzystaniem oznaczenia i porównania poziomu ekspresji zestawu wybranych biomarkerów miRNA, tj. miRNA-1246 i miRNA-150-5p. Te klasyfikujące modele diagnostyczne zostały dopasowane, z wykorzystaniem techniki uczenia maszynowego oraz poprzez wyznaczenie odpowiednich punktów odcięcia, do znanych i ogólnodostępnych metod oceny poziomu ekspresji genów jak opisano powyżej. Zatem, sposób według wynalazku można zrealizować z zastosowaniem znanych w dziedzinie metod pomiaru ekspresji miRNA, korzystnie takich jak mikromacierze lub platforma NanoString, a najkorzystniej reakcja odwrotnej transkrypcji z łańcuchową reakcją polimerazy z analizą ilości produktu w czasie rzeczywistym (RT-qPCR). W celu skutecznego zastosowania sposobu według wynalazku należy jedynie wyznaczyć poziom ekspresji zestawu biomarkerów miRNA według wynalazku w próbce od osobnika, korzystnie w próbce surowicy, oraz znormalizować wynik odpowiednio do użytej metody pomiarowej jak to wiadomo w dziedzinie i jak opisano powyżej. Następnie zaś w sposobie według wynalazku przeprowadza się porównanie poziomów ekspresji, korzystnie poprzez klasyfikację próbki jako próbkę od osobnika z rakiem jajnika lub jako próbkę od osobnika bez raka jajnika z zastosowaniem klasyfikującego modelu diagnostycznego, wykorzystując model regresji logistycznej, podstawiając wynik poziomu ekspresji wyselekcjonowanych miRNA od osobnika i porównując z punktem odcięcia odpowiednim dla wykorzystanej metody pomiaru ekspresji biomarkerów miRNA według wynalazku. Według wynalazku wynik klasyfikującego modelu diagnostycznego zostaje obliczony z wykorzystaniem zestawu danych zawierającego dane o poziomie ekspresji zestawu markerów miRNA obejmującego miR-1246 i miR-150-5p.

Dzięki wyznaczonemu punktowi odcięcia na podstawie wyników klasyfikacji na zestawie treningowym, nie wymaga się następnie przeprowadzania żadnego porównania z danymi od osobników bez raka jajnika, ponieważ klasyfikacja jest oparta o porównanie wyniku klasyfikującego modelu diagnostycznego z optymalnym punktem odcięcia ustalonym na podstawie wytrenowanego klasyfikującego modelu diagnostycznego. Wartość tego optymalnego punktu odcięcia jest wyznaczana w punkcie najlepszych wyników dla swoistości i czułości podczas trenowania modelu.

Oznacza to, że klasyfikujący model diagnostyczny na podstawie wprowadzonych do niego danych ze zbioru treningowego podlega wytrenowaniu i dokonuje klasyfikacji próbki jako próbkę pochodzącą od osobnika z rakiem jajnika lub jako próbkę od osobników bez raka jajnika.

Co istotne w sposobie według wynalazku korzystnie nie porównuje się wartości ΔCt , ale te wartości ΔCt wprowadza się do klasyfikującego modelu diagnostycznego i porównuje z wyznaczonym wcześniej optymalnym punktem odcięcia odpowiednim dla wykorzystanej metody pomiaru ekspresji biomarkerów miRNA według wynalazku.

Zgodnie z wynalazkiem klasyfikujący model diagnostyczny utworzony na podstawie znormalizowanych danych ekspresji miRNA w surowicy na platformie NanoString wygląda następująco:

$$P(Y = 1 | x_1, x_2, \dots, x_k) = \frac{e^{4,47+0,071 \cdot \text{miR}-1246-0,320 \cdot \text{miR}-150-5p}}{1 + e^{4,47+0,071 \cdot \text{miR}-1246-0,320 \cdot \text{miR}-150-5p}}$$

Punkt odcięcia w zakresie $0,3 < x < 0,9$ daje w tym przypadku wynik dla swoistości (Sp) oraz czułości (S) $> 80\%$ (zobacz Tabela 7 poniżej).

Zgodnie z wynalazkiem klasyfikujący model diagnostyczny utworzony na podstawie znormalizowanych danych ekspresji miRNA w surowicy na platformie mikromacierzy (np. Affymetrix) wygląda następująco:

$$P(Y = 1 | x_1, x_2, \dots, x_k) = \frac{e^{-2,57+0,054 \cdot \text{miR}-1246+0,49 \cdot \text{miR}-150-5p}}{1 + e^{-2,57+0,054 \cdot \text{miR}-1246+0,49 \cdot \text{miR}-150-5p}}$$

Punkt odcięcia równy 0,5 daje w tym przypadku wynik dla swoistości (Sp) oraz czułości (S) $> 80\%$ (zobacz Tabela 11 poniżej).

Zgodnie z wynalazkiem klasyfikujący model diagnostyczny utworzony na podstawie znormalizowanych danych ekspresji miRNA w surowicy metodą RT-qPCR. wygląda następująco:

$$P(Y = 1 | x_1, x_2, \dots, x_k) = \frac{e^{55,160-1,616 \cdot \text{miR}-1246+4,277 \cdot \text{miR}-150-5p}}{1 + e^{55,160-1,616 \cdot \text{miR}-1246+4,277 \cdot \text{miR}-150-5p}}$$

Punkt odcięcia w tym przypadku w zakresie $0,1 < x \leq 0,8$ daje wynik dla czułości (S) oraz swoistości (Sp) na zbiorze testowym $> 85\%$ (zobacz Tabela 16 poniżej).

Zgodnie z wynalazkiem pomiar poziomu ekspresji zestawu biomarkerów miRNA według wynalazku najkorzystniej przeprowadza się metodą RT-qPCR. Taka metoda uznawana jest za dokładną, czułą i swoistą w kontekście dojrzałych miRNA. Pozwala na określenie poziomu ekspresji miRNA nawet o niskich poziomach. Do badania potrzebna jest próbka o niewielkiej objętości. Metodę tę realizuje się korzystnie w dwóch etapach: 1) reakcji odwrotnej transkrypcji (RT) i 2) właściwej reakcji łańcuchowej polimerazy (PCR). Oba etapy przeprowadza się w standardowy sposób znany w dziedzinie z zastosowaniem komercyjnie dostępnych zestawów i starterów do reakcji RT i PCR swoistych dla dwóch wyselekcjonowanych miRNA, stanowiących zestaw biomarkerów miRNA według wynalazku, tj. miR-1246 i miR-150-5p, stosując rekomendowane przez ich wytwórców warunki. W metodzie tej korzystnie dokonuje się także pomiaru poziomu ekspresji genu referencyjnego, korzystnie wybranego miR-103-3p, celem umożliwienia normalizacji wyników. Przed rozpoczęciem właściwej reakcji przygotowuje się serię rozcieńczeń matrycy, co służy do obliczenia efektywności reakcji każdej pary starterów. Każdorazowo właściwa reakcja jest przeprowadzana dla cDNA otrzymanego z udziałem RNA odwrotnej transkryptazy (RT+), jak również dla prób kontrolnych bez dodatku odwrotnej transkryptazy (RT-) oraz dla prób z samą wodą zamiast matrycy. Po otrzymaniu wartości Ct dla każdego miRNA, wyniki są normalizowane do wartości Ct genu referencyjnego i oblicza się wartość według następującego wzoru:

$$\text{deltaCt} = (\text{Ct miR-1246/miR-150-5p} - \text{Ct miR-103-3p}).$$

Uzyskane dane z reakcji RT-qPCR w postaci wartości różnicy ekspresji (deltaCt) korzystnie wprowadzane są zgodnie z wynalazkiem do klasyfikującego modelu diagnostycznego, którego wynik po porównaniu wyniku równania z punktem odcięcia, jak opisano powyżej, umożliwia klasyfikację pacjentki, od której pobrano materiał jako pacjentkę z rakiem jajnika lub bez raka jajnika.

Jak wiadomo, szczególnym wskazaniem przy przeprowadzaniu reakcji PCR w czasie rzeczywistym w przypadku analiz miRNA jest wyliczenie wydajności amplifikacji użytych starterów na podstawie wyników uzyskanych z serii 10-krotnych rozcieńczeń matrycy. W tym celu należy wyznaczyć krzywą regresji na podstawie uzyskanych punktów pomiarowych i współczynnik nachylenia prostej. W sposobach według wynalazku korzystnie stosuje się etap optymalizacji metody PCR w postaci oceny wydajności każdej pary zastosowanych starterów, optymalizacji stężenia matrycy cDNA oraz stężenia użytych starterów, w sposób znany w dziedzinie.

Sposób diagnostyki raka jajnika według wynalazku pozwala na diagnozowanie raka jajnika w sposób nieinwazyjny. Zaletą sposobu według wynalazku jest to, że charakteryzuje się on wysoką czułością i swoistością diagnostyczną i/lub prognostyczną. Należy podkreślić fakt, że sposób według wynalazku okazał się skuteczny w diagnostyce raka jajnika w różnych jego stadiach, również we wczesnych stadiach (FIGO I-II), które jest bardzo rzadko wykrywane z zastosowaniem obecnie dostępnych metod. Taki sposób jest szczególnie korzystny ekonomicznie dla służb zdrowia, bo pozwala zdiagnozować raka jajnika na wczesnym etapie jego zaawansowania, a co za tym wdrożyć skuteczne leczenie na wcześniejszym etapie, a przez to zwiększyć skuteczność leczenia, obniżyć koszty leczenia i poprawić jakość życia pacjenta. Zaletą niniejszego wynalazku jest również to, że do badania potrzebna jest tylko niewielka ilość próbki krwi, z której izoluje się surowicę do badania zgodnie z wynalazkiem przez co nie ma konieczności przeprowadzenia inwazyjnej biopsji cienkoigłowej obciążonej ryzykiem dla pacjenta. Dzięki temu, że w sposobie według wynalazku wykorzystywany jest zestaw biomarkerów miRNA, który jest wysoce precyzyjny, bo wykazuje wyższe AUC, czyli czułość i swoistość niż szeroko używany w diagnostyce nowotworu jajnika marker CA 125 (na podstawie danych z literatury), zastosowanie tego zestawu biomarkerów i korzystnie sposoby wykorzystujące klasyfikujące modele diagnostyczne mogą być rekomendowane przez onkologów w celu szybkiego potwierdzenia wstępnej diagnozy. Sposób i test diagnostyczny według wynalazku mogą być stosowane w laboratoriach diagnostycznych, w badaniach przesiewowych wspierających wczesną diagnostykę raka jajnika, do oceny skuteczności terapii raka jajnika, bądź do kontroli pacjentów po zakończeniu leczenia (tzw. follow-up). Wynalazki według niniejszego zgłoszenia nadają się również do stosowania w diagnostycznych badaniach przesiewowych pod kątem występowania raka jajnika, np. w grupach ryzyka takich jak kobiety po 40 roku życia, w szczególności obciążone wywiadem rodzinnym w kierunku raka jajnika lub raka piersi. Takie diagnostyczne badania przesiewowe są w tego rodzaju przypadkach szczególnie korzystne, ponieważ obecnie nie istnieją żadne skuteczne środki, które pozwoliłyby na diagnozowanie raka jajnika na wczesnych etapach jego zaawansowania, zwłaszcza w takich populacjach.

Wyniki uzyskane zgodnie z wynalazkiem świadczą o wysokiej czułości i swoistości przedstawionych tu wynalazków i przewyższają wyniki uzyskane z zastosowaniem środków diagnostycznych znanych w dziedzinie, w szczególności opartych na pomiarze markera CA125 i USG przezpochwowym.

Wynalazek zostanie teraz zilustrowany za pomocą przykładów wykonania i figur, które nie mają jednak w jakikolwiek sposób ograniczać zakresu ochrony zdefiniowanego w zastrzeżeniach patentowych.

Krótki opis figur

Fig. 1 przedstawia mapę cieplną ze zgrupowaniem hierarchicznym dla zróżnicowanych miRNA.

Fig. 2 przedstawia krzywe ROC i AUC (pole pod krzywą) dla zróżnicowanych miRNA otrzymane na podstawie danych poziomu ekspresji cząsteczek miRNA z użyciem platformy NanoString.

Fig. 3 przedstawia krzywe ROC i AUC (pole pod krzywą) dla zbioru treningowego i testowego dla klasyfikującego modelu diagnostycznego bazującego na danych poziomu ekspresji miR-1246 i miR-150-5p uzyskanych z użyciem platformy NanoString. W obrębie wykresu przedstawiono wartość AUC, punkt odcięcia obliczony metodą statystyki J Youdena oraz odpowiadającą temu punktowi czułość i swoistość.

Fig. 4 przedstawia krzywe ROC oraz AUC dla zbioru treningowego i testowego dla klasyfikującego modelu diagnostycznego bazującego na publicznych danych poziomu ekspresji miR-1246 i miR-

-150-5p otrzymanych z użyciem techniki mikromacierzy, znajdujących się w bazie Gene Expression Omnibus (GEO). W obrębie wykresu przedstawiono wartość AUC, punkt odcięcia oraz odpowiadającą temu punktowi czułość i swoistość.

Fig. 5 przedstawia krzywe ROC oraz AUC dla klasyfikującego modelu diagnostycznego w zależności od stadium raka jajnika (FIGO I–IV). W obrębie wykresu przedstawiono wartość AUC, punkt odcięcia oraz odpowiednio odpowiadającą temu punktowi czułość i swoistość.

Fig. 6 przedstawia krzywe ROC i AUC dla klasyfikującego modelu diagnostycznego w zależności od rodzaju nowotworu, w nawiasach podany został przedział ufności. Analizę przeprowadzono z wykorzystaniem publicznych danych poziomu ekspresji miRNA uzyskanych techniką mikromacierzy dla raka płuca i raka jelita grubego.

Fig. 7 przedstawia względny poziom ekspresji miR-1246 i miR-150-5p w surowicy kobiet z rakiem jajnika w porównaniu do grupy kontrolnej uzyskany z zastosowaniem techniki RT-qPCR.

Fig. 8 przedstawia krzywe ROC oraz AUC dla zbioru treningowego i testowego dla klasyfikującego modelu diagnostycznego otrzymane na podstawie danych uzyskanych z zastosowaniem techniki RT-qPCR. W obrębie wykresu przedstawiono wartość AUC, a w nawiasie podany został przedział ufności.

Przykłady

Przy realizacji wszystkich badań i testów tu opisanych wykorzystano standardowe, ogólnie znane procedury przygotowania materiału biologicznego, izolacji miRNA i pomiaru ekspresji miRNA, a także komercyjnie dostępne zestawy i aparaty do tych celów, w tym znane oprogramowanie do analizy, postępując zgodnie z zaleceniami ich wytwórców oraz znane metody statystyczne i bioinformatyczne, o ile wyraźnie nie wskazano inaczej.

Opracowanie sposobu diagnostyki raka jajnika przeprowadzono w oparciu o ekspresję wybranych cząsteczek miRNA, najpierw przez wybranie odpowiednich cząsteczek przy użyciu platformy wielkoskalowej NanoString, a następnie przez weryfikację uzyskanych rezultatów w badaniach opartych na ilościowej reakcji łańcuchowej polimerazy DNA (RT-qPCR). Wszystkie badania zostały zatwierdzone przez Komisję Bioetyczną Uniwersytetu Medycznego w Białymstoku (nr zgody APK.002.69.2020).

Profilowanie ekspresji miRNA w surowicy z użyciem platformy NanoString

Profilowanie ekspresji miRNA w surowicy z użyciem platformy NanoString zostało przeprowadzone w grupie (n=70) obejmującej pacjentki z niskozróżnicowanym rakiem jajnika (n=36) oraz próbie kontrolnej, czyli grupie osób bez raka jajnika (n=34). Za pomocą testu sumy rang Wilcozona została obliczona wartość p dla wieku i wskaźnika masy ciała (BMI). Między grupami nie stwierdzono istotnych statystycznie różnic w wieku i BMI (wartość $p > 0,05$). Charakterystykę uczestników przedstawiono poniżej w Tabeli 3. Próbkę krwi pobierano przed rozpoczęciem leczenia.

Tabela 3. Charakterystyka pacjentek, których profil miRNA w surowicy został oznaczony z użyciem platformy Nanostring.

	<i>Kontrola</i>	<i>Rak jajnika</i>		
Liczba pacjentek	34	36		
	<i>Średnia ± SD</i>	<i>Min</i>	<i>Max</i>	<i>Wartość p</i>
Wiek w chwili diagnozy (lata)				
Zdrowi	63±14	38	86	0,1
Chorzy	59±6	45	72	
BMI (waga/wzrost²)				
Zdrowi	29±14	38	86	0,4
Chorzy	27±3	21	34	
	<i>FIGO I</i>	<i>FIGO II</i>	<i>FIGO III-IV</i>	<i>N/A</i>
Liczba przypadków	3	1	31	1

Weryfikacja doświadczalna niniejszego wynalazku obejmowała następujące etapy:

I. Profilowanie miRNA w grupie z rakiem jajnika i grupie kontrolnej

Ekstrakcja RNA z surowicy została przeprowadzona za pomocą zestawu miRCURY™ RNA isolation Kit (Exiqon, Dania), zgodnie z protokołem producenta. We wszystkich próbkach oznaczano ilościowo stężenie RNA za pomocą metody fluorometrycznej przy użyciu aparatu Qubit™ (Thermo Fisher Inc., USA). Analiza została wykonana w 6 zbalansowanych eksperymentach przy użyciu platformy NanoString. Pozwala ona na detekcję jednocześnie w jednej próbce poziomu ekspresji 798 miRNA. Wszystkie etapy były wykonywane według protokołu producenta. Dane przeanalizowano z użyciem oprogramowania nSolver wersja 4.0. Normalizacja została przeprowadzona za pomocą średniej geometrycznej dla Top100 miRNA. Krotność zmiany („Fold change”, FC) została policzona poprzez określenie zdrowych osobników jako poziom podstawowy. Korekta dla wielokrotnego testowania hipotez została wprowadzona za pomocą spodziewanego odsetka wyników fałszywie dodatnich (*ang.* False Discovery Rate, FDR) czyli wartości oczekiwanej frakcji fałszywie odrzuconych hipotez zerowych w zbiorze wszystkich odrzuconych hipotez zerowych, przemnożona przez prawdopodobieństwo odrzucenia co najmniej jednej hipotezy według Benjamina-Hochberga. U pacjentów z niskozróżnicowanym surowiczym rakiem jajnika stwierdzono istotnie różne profile ekspresji dwunastu miRNA w surowicy w porównaniu z kontrolami (Tabela 4). Dane w postaci zliczeń po normalizacji zostały zaimportowane i użyte w dalszej analizie statystycznej. W tabeli strzałką w dół zostały zaznaczone miRNA, których ekspresja jest obniżona, strzałką w górę te miRNA, których ekspresja jest podwyższona u pacjentek z rakiem jajnika w porównaniu do grupy bez raka jajnika.

Tabela 4. Parametry krotności zmiany i wartości p po korekcie Benjamina-Hochberga (FDR).

<i>miRNA</i>	<i>FC</i>		<i>FDR</i>
miR-144-3p	-2,31	↓	0,00
miR-142-3p	-1,94	↓	0,00
miR-150-5p	-1,81	↓	0,00
miR-15a-5p	-1,75	↓	0,00
miR-15b-5p	-1,65	↓	0,00
miR-126-3p	-1,61	↓	0,00
miR-4454+miR-7975	2,45	↑	0,00
miR-1246	2,61	↑	0,00
miR-191-5p	-1,51	↓	0,01
miR-4516	1,68	↑	0,01
miR-630	2,16	↑	0,01
miR-106b-5p	-1,36	↓	0,04

Na Fig. 1. przedstawiono mapę cieplną uwzględniającą 12 miRNA różnicująco eksprymowanych u pacjentek z rakiem jajnika i w grupie kontrolnej, czyli bez raka jajnika. Cząsteczki miRNA o zróżnicowanej ekspresji zostały zgrupowane hierarchicznie za pomocą metryki odległości euklidesowej z kompletną metodą połączenia. Uwidoczniono występowanie czterech klastrów miRNA.

II. ROC dla każdego ze zróżnicowanych miRNA

Krzywa ROC pozwala na ocenę poprawności klasyfikatora, który może się okazać potencjalnym markerem diagnostycznym. Pozwala ona również na obliczenie swoistości i czułości w konkretnym punkcie odcięcia. Swoistość (Sp) jest stosunkiem wyników prawdziwie ujemnych, czyli osobników, którzy nie mają nowotworu i tak przyporządkowanych na podstawie klasyfikatora. Czułość (S) natomiast

jest równa stosunkowi wyników prawdziwie dodatnich, czyli osobników, które mają nowotwór i tak porządkowanych na podstawie klasyfikatora. Za pomocą pakietu pROC [X. Robin i wsp., „pROC: An open-source package for R and S+ to analyze and compare ROC curves”, BMC Bioinformatics, vol. 12, nr 1, str. 77, marzec 2011] w R [R. C. T. R. Foundation, „R: A Language and Environment for Statistical Computing”, vol. 2, <https://www.R-project.org>, 2013.] zostały obliczone AUC wraz z przedziałem ufności (95%), punkt odcięcia, czułość i swoistość jak i została stworzona krzywa ROC dla każdego ze zróżnicowanych miRNA. Wyniki są przedstawione w Tabeli 5 oraz na Fig. 2.

Tabela 5. Przewidywalność zróżnicowanego miRNA. Przedstawiono parametry krzywej ROC (pole pod krzywą (AUC) i 95% przedział ufności (CI) oraz czułość (S) i swoistość (Sp) odpowiadające optymalnemu punktowi odcięcia wybranemu według metody Youden'a).

Zróżnicowane miRNA	AUC (%)	Dolna granica CI (%)	Górna granica CI (%)	Punkt odcięcia	S	Sp
miR-1246	92,3	86,1	98,6	92,5	80,6	94,1
miR-144-3p	82,2	72,4	92	53,0	55,6	97,1
miR-4454 + miR-7975	85,6	76,3	94,9	74,2	75	88,2
miR-150-5p	87,2	78,8	95,5	31,1	75	85,3
miR-630	78,2	67,3	89,1	149,8	80,6	64,7
miR-142-3p	83,3	73,8	92,9	168,1	86,1	76,5
miR-15a-5p	82,4	72,4	92,5	27,79	86,1	70,6
miR-15b-5p	80,1	70,0	90,1	44,1	83,3	64,7
miR-191-5p	79,6	69,1	90,1	27,2	55,6	97,1
miR-106b-5p	75,0	63,3	86,7	23,4	47,2	97,1
miR- 4516	78,4	67,7	89,2	88,35	55,6	88,2

III. Selekcja atrybutów do zminimalizowania zmiennych i wybrania kombinacji miRNA

W celu selekcji miRNA, które są najlepszymi potencjalnymi biomarkerami, dane zostały podzielone losowo, za pomocą pakietu caret [M. Kuhn, „Building Predictive Models in R Using the caret Package”, *J. Stat. Softw.*, 2008.] w R version 3.6.1 [R. C. T. R. Foundation, „R: A Language and Environment for Statistical Computing”, vol. 2, <https://www.R-project.org>, 2013.], na zbiór treningowy (70%) i zbiór testowy (30%).

Z zastosowaniem zestawu treningowego za pomocą dwóch metod selekcji atrybutów: Information Gain (zysk informacyjny) oraz Correlation-based Feature Subset Selection (wybór podzbioru cech oparty na korelacji) zostały wybrane 3 miRNA, z których zostały stworzone dwa modele regresji logistycznej. Te modele regresji logistycznej są klasyfikującymi modelami diagnostycznymi. W każdym z nich ograniczono liczbę zmiennych zależnych do dwóch. Te klasyfikujące modele diagnostyczne zostały zwalidowane na zbiorze testowym. Selekcja atrybutów została wykonana za pomocą programu WEKA (Waikato Environment for Knowledge Analysis Version 3.8.3). Obie metody zostały wykonane na zbiorze testowym z użyciem sprawdzianu krzyżowego - leave-one-out (LOOV). Została użyta metoda Information Gain z metodą poszukiwań typu Ranker, która bazuje na obliczaniu zmniejszającej się entropii poprzez dodawanie atrybutów. Na tej podstawie zostały wyselekcjonowane najlepsze trzy miRNA, które najsilniej zmniejszają entropię: miR-1246, miR-144-3p and miR-150-5p.

Do wyboru podzbioru cech opartego na korelacji, czyli z ang. Correlation-based Feature Selection, została użyta metoda szukania: BestFirst (algorytm zachłanny). Metoda ta bazuje na wynikach korelacji z klasą oraz między atrybutami.

Najsilniejsze atrybuty według tej metody to:

miR-1246, miR-4454+miR-7975, miR-150-5p, miR-4516 i miR-144-3p.

IV. Utworzenie klasyfikującego modelu diagnostycznego i ocena jego jakości

Do utworzenia klasyfikującego modelu diagnostycznego użyto regresji logistycznej opartej o miR-1246 oraz miR-150-5p. Wielowymiarowe modele umożliwiają badanie zależności wielu zmiennych niezależnych na jedną zmienną zależną. Celem regresji logistycznej jest znalezienie takiej funkcji bazującej na zmiennych, która z najwyższym prawdopodobieństwem właściwie klasyfikuje dane. Dzięki znalezieniu takiej funkcji możliwe jest, na podstawie zmiennych niezależnych, obliczenie prawdopodobieństwa oraz klasyfikacja nowego osobnika. W tym przypadku na podstawie poziomu znormalizowanej ilości zliczeń wytypowanych miRNA z platformy Nanostring, które są zmiennymi niezależnymi, wytrenowany został model regresji logistycznej, który klasyfikuje kobiety chore na niskozróżnicowany surowiczy nowotwór jajnika i kobiety bez takiego nowotworu. Poniżej znajduje się równanie regresji logistycznej, na której podstawie został zbudowany, klasyfikujący model diagnostyczny:

$$P(Y = 1|x_1, x_2, \dots, x_k) = \frac{e^{a_0 + \sum_{i=2}^k a_i x_i}}{1 + e^{a_0 + \sum_{i=2}^k a_i x_i}}$$

w którym

$P(Y = 1|x_1, x_2, \dots, x_k)$ to warunkowe prawdopodobieństwo, że zmienna zależna Y przyjmie wartość równą 1 dla wartości zmiennych niezależnych $x_1, x_2, \dots, x_k$

e to liczba Eulera, $\approx 2,718$

a_0 to stała (punkt przecięcia)

$a_1, a_2, \dots, a_k$ to współczynniki regresji dla poszczególnych zmiennych niezależnych, predyktorów

$x_1, x_2, \dots, x_k$ to zmienne niezależne, predyktory, zmienne wyjaśniające.

Wraz ze wzrostem liczby predyktorów powstaje niebezpieczeństwo zbytniego „dopasowania” modelu („overfitting”) do danych, dlatego twórcy wynalazku zdecydowali się na użycie jedynie dwóch miRNA w klasyfikującym modelu diagnostycznym.

V. Budowanie modelu regresji logistycznej i ocena przydatności

Model regresji logistycznej (rozkład dwumianowy GLM z pakietu caret [M. Kuhn, „Building Predictive Models in R Using the caret Package”, *J. Stat. Softw.*, 2008.]) został wytrenowany na zbiorze treningowym, który zawierał 70% danych, a następnie zwalidowany na zbiorze testowym (30% danych). Aby zbudować stabilny model w trakcie trenowania został użyty sprawdzian krzyżowy, a dokładniej K-krotna walidacja ($K = 10$). Takim sposobem dane zostają podzielone na 10 podzbiorów. Następnie każdy z tych podzbiorów zostaje kolejno użyty jako zbiór testowy, gdy pozostałe są użyte jako zbiór treningowy. W wyniku tego analiza zostaje przeprowadzona 10 razy. W budowaniu modelu ta K-krotna walidacja została przeprowadzona trzy razy, czyli cały zbiór treningowy został podzielony trzykrotnie na 10 podzbiorów. Wyniki analizy zostały następnie uśrednione w celu uzyskania jednego wyniku.

Ocenę właściwej klasyfikacji za pomocą modeli zbudowanych na połączeniach wytypowanych miRNA wykonano za pomocą wykresów ROC, AUC wraz z przedziałem ufności, poziomem czułości i swoistości w punkcie odcięcia oraz R^2 i RMSE. Obliczenia jak i wykresy zostały wykonane z użyciem oprogramowania R wersja 3.6.1 [R. C. T. R. Foundation, „R: A Language and Environment for Statistical Computing”, vol. 2, <https://www.R-project.org>, 2013.] oraz pakietów pROC [X. Robin i wsp., „pROC: An open-source package for R and S+ to analyze and compare ROC curves”, *BMC Bioinformatics*, vol. 12, nr 1, str. 77, marzec 2011], Optimal Cutpoints [M. López-Ratón, M. X. Rodríguez-Álvarez, C. Cadarso-Suárez, and F. Gude-Sampedro, „Optimalcutpoints: An R package for selecting optimal cutpoints in diagnostic tests”, *J. Stat. Softw.*, vol. 61, nr 8, str. 1–36, listopad 2014] oraz programu GraphPad Prism wersja 8. Punkt odcięcia został obliczony za pomocą metody Youdena na zbiorze treningowym. W tym punkcie została obliczana czułość i swoistość na zbiorze testowym. W poniższej tabeli (Tabela 6) przedstawione są parametry wytrenowanego modelu.

Tabela 6. Zmienne w zbudowanym modelu dla platformy NanoString. a_0 – stała, a_1 – współczynnik predyktora x_1 , a_2 – współczynnik predyktora x_2 .

Współczynniki	a_0	a_1	a_2
$x_1 = \text{miR-1246}$	4,47117	0,07091	-0,31985
$x_2 = \text{miR-150-5p}$			

Aby móc ocenić jakość klasyfikacji opracowanego modelu, niezbędne jest obliczenie podstawowych parametrów jakości, które zostały przedstawione w tabeli poniżej (Tabela 7). Model utworzony na podstawie znormalizowanych danych ekspresji miRNA w surowicy na platformie NanoString:

$$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150) = \frac{e^{4,47+0,071 \cdot \text{miR} - 1246 - 0,320 \cdot \text{miR} - 150 - 5p}}{1 + e^{4,47+0,071 \cdot \text{miR} - 1246 - 0,320 \cdot \text{miR} - 150 - 5p}}$$

w którym

$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150)$ to warunkowe prawdopodobieństwo, że zmienna zależna Y przyjmie wartość równą 1 dla wartości zmiennych niezależnych $\text{miR} - 1246$, $\text{miR} - 150$
 e to liczba Eulera, $\approx 2,718$.

Punkt odcięcia w zakresie $0,1 \leq x \leq 0,9$ daje wynik dla swoistości (Sp) oraz czułości (S) $> 70\%$ (Tabela 7).

Tabela 7. Punkty odcięcia w zakresie $0 < x < 1$ oraz wartości czułości i swoistości.

Punkt odcięcia	Swoistość (Sp)	Czułość (S)
0,0	0	100,0
0,1	76,92308	100,0
0,2	76,92308	100,0
0,3	84,61538	100,0
0,4	92,30769	100,0
0,5	92,30769	100,0
0,6	92,30769	100,0
0,7	100	100,0
0,8	100	100,0
0,9	100	87,5
1,0	100	0

Tabela 8. Parametry jakości dla o modelu na zbiorze treningowym i testowym.

Nazwa	miRNA-1246, miRNA-150-5p	miRNA-1246, miRNA-150-5p
Zbiór	Treningowy	Testowy
AUC	98,6%	100%
CI dolna granica	96,4%	-
CI górna granica	100%	-
Optymalny punkt odcięcia (Youden Index)	0,44	0,44
Czułość	96,4%	100%
Swoistość	95,2%	92,3%
R²	0,74	-
RMSE	0,23	-

Krzywa ROC pomaga zwizualizować potencjał diagnostyczny opracowanych modeli. Na Fig. 3. przedstawione są krzywe ROC, jedna dla zbioru treningowego oraz druga dla zbioru testowego. Na każdej z krzywych został zaznaczony punkt odcięcia, obliczony metodą statystyki J Youdena (wskaźnik Youdena) oraz odpowiednio w nawiasie odpowiadająca temu punktowi czułość i swoistość.

Następnie została stworzona Tablica pomyłek (Tabela 9) dla obu modeli na danych, które nie zostały włączone podczas trenowania modelu. Przedstawione na niej (Tabela 9) są informacje o TP (prawdziwie pozytywne), TN (prawdziwie negatywne), FP (fałszywie pozytywne) oraz FN (fałszywie negatywne).

Tabela 9. Tablica pomyłek w zbiorze testowym dla modelu z miRNA-1246 i miRNA-150-5p.

		Stan faktyczny	
		0	1
Predykcja	0	12 (TN)	0 (FN)
	1	1 (FP)	8 (TP)

VI. Walidacja doboru klasyfikujących miRNA na niezależnym zbiorze

Do walidacji potencjału wybranych miRNA jako silnego klasyfikatora raka jajnika na niezależnej kohorcie pacjentów wykorzystano dane z publicznie dostępnej bazy danych Gene Expression Omnibus (GEO) nr GSE1068 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106817>). Badanie zostało przeprowadzone z użyciem mikromacierzy 3D-Gene Human miRNA V21_1.0.0 (Toray Industries, Inc.). Dany zbiór zawiera profile miRNA 4046 pacjentek, w tym 333 ze zdiagnozowanym nowotworem jajnika, 66 kobiet z guzami granicznymi jajnika, 29 z łagodnymi odmianami chorób jajników, 2759 pacjentek bez nowotworu oraz 859 przypadków z innymi nowotworami. Analiza została wykonana na danych zliczeń panelu miRNA w surowicy.

Z całego zbioru danych zostały wybrane dane zdrowych osobników ($n = 2759$) oraz pacjentek z rakiem jajnika ($n = 320$). Jednak, aby zbalansować wielkość grup, zostały losowo wybrane dane osób zdrowych w ilości 320. Z tego względu, iż dane zostały wygenerowane za pomocą innego urządzenia, ponownie został stworzony model bazujący na ekspresji miR-1246, miR-150-5p, w ten sam sposób jak w punkcie V. Potencjał diagnostyczny został określony za pomocą krzywej ROC, punkt odcięcia został ustalony na zbiorze treningowym w grupie badanych. Następnie zwalidowany na zbiorze testowym, również z uwzględnieniem stadium choroby.

Poniżej przedstawione są parametry wytrenowanego modelu.

Tabela 10. Zmienne w zbudowanym modelu dla mikromacierzy.

Współczynniki	Stała a_0	Współczynnik predyktora x_1 (a_1)	Współczynnik predyktora x_2 (a_2)
$x_1 = \text{miR-1246}$	-2,57007	0,05412	0,49059
$x_2 = \text{miR-150-5p}$			

Model regresji logistycznej został utworzony na podstawie znormalizowanych danych ekspresji miRNA w surowicy na platformie Affymetrix (mikromacierz) z użyciem następującego wzoru:

$$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150) = \frac{e^{-2,57+0,054 \cdot \text{miR}-1246+0,49 \cdot \text{miR}-150-5p}}{1 + e^{-2,57+0,054 \cdot \text{miR}-1246+0,49 \cdot \text{miR}-150-5p}}$$

w którym

$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150)$ to warunkowe prawdopodobieństwo, że zmienna zależna Y przyjmie wartość równą 1 dla wartości zmiennych niezależnych $\text{miR} - 1246, \text{miR} - 150$ e to liczba Eulera, $\approx 2,718$.

Punkt odcięcia w zakresie $0,4 \leq x \leq 0,7$ daje wynik dla swoistości (Sp) oraz czułości (S) $> 70\%$ (Tabela 11).

Tabela 11. Punkty odcięcia w zakresie $0 < x < 1$ oraz wartości czułości i swoistości.

Punkt odcięcia	Swoistość (Sp)	Czułość (S)
0,0	0	100,0
0,1	35,41667	92,70833
0,2	47,91667	91,66667
0,3	65,62500	90,62500
0,4	77,08333	87,50000
0,5	85,41667	83,33333
0,6	89,58333	78,12500
0,7	94,79167	70,83333

Na Fig. 4 przedstawiono uzyskane krzywe ROC oraz AUC dla wytrenowanego modelu na zbiorze treningowym oraz testowym. Na krzywej zaznaczony jest punkt odcięcia, podano także AUC, a w nawiasach przedział ufności.

Następnie została sporządzona tablica pomyłek w zbiorze testowym dla klasyfikującego modelu diagnostycznego dla mikromacierzy z miR-1246 i miR-150-5p na podstawie danych ze zbioru testowego. W Tabeli 12 przedstawione są informacje o ilości wyników prawdziwie pozytywnych (TP), prawdziwie negatywnych (TN), fałszywie pozytywnych (FP) oraz fałszywie negatywnych (FN).

Tabela 12. Tablica pomyłek w zbiorze testowym dla modelu dla danych z mikromacierzy z miRNA-1246, miRNA-150-5p.

		Stan faktyczny	
		0	1
Predykcja	0	182 (TP)	37 (FN)
	1	42 (FP)	187 (TF)

Na Fig. 5 przedstawiono krzywe ROC dla wytrenowanego modelu klasyfikującego opartego na regresji logistycznej oraz obliczone AUC wraz z przedziałem ufności dla różnych stadiów raka jajnika według FIGO. Dla każdego ze stadiów została obliczona czułość i swoistość modelu w określonym punkcie odcięcia. Tabela 13 zawiera streszczenie wyników parametrów jakościowych dla opracowanego modelu opartego o dwa wybrane miRNA (miR-1246 oraz miR-150-5p). Ponieważ w danych walidacyjnych była dostępna informacja o zaawansowaniu choroby (FIGO), w tabeli również podana jest informacja o liczności grupy oraz o wyniku klasyfikacji dla poszczególnych stadiów choroby.

Tabela 13. Wyniki oceny jakości modelu na danych zewnętrznych.

Nazwa	miRNA-1246, miRNA-150-5p	miRNA-1246, miRNA-150-5p
Zbiór	Treningowy	Testowy
AUC (95% CI)	89,1%	89,5%
CI dolna granica	86,1%	84,8%
CI górna granica	92,1%	94,1%
Punkt odcięcia	0,56	0,56
Czułość	81,2%	82,3%
Swoistość	87,1%	87,5%
FIGO I (n = 80)		
AUC (95% CI)	-	88,3%
FIGO II (n = 30)		
AUC (95% CI)	-	92,8%
FIGO III (n = 112)		
AUC (95% CI)	-	88,4%
FIGO IV (n = 32)		
AUC (95% CI)	-	88,7%

Dodatkowo, z wykorzystaniem publicznie dostępnych danych VI (dane z: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106817>) zawierających również dane o ekspresji miRNA w surowicy u pacjentów z innymi nowotworami, zostały utworzone krzywe ROC oraz obliczone AUC wraz z przedziałem ufności na podstawie stworzonego klasyfikującego modelu diagnostycznego. Wyniki dla wytrenowanego modelu przedstawione są na Fig. 6. Widoczne jest, że klasyfikujący model diagnostyczny jest w stanie właściwie odróżnić chorych na raka jajnika od chorych na raka płuca i od chorych na raka jelita grubego.

VII. Klasyfikujący model diagnostyczny dla metody RT-qPCR

W celu wytworzenia klasyfikującego modelu diagnostycznego możliwego do stosowania w praktyce, do oceny poziomu ekspresji miR-1246 i miR-150-5p zastosowano reakcję odwrotnej transkrypcji wraz z reakcją łańcuchowej polimerazy z analizą ilości produktu w czasie rzeczywistym (RT-qPCR). Do badania włączono powiększoną grupę pacjentek (n=88), w tym 42 pacjentki z rakiem jajnika oraz 46 pacjentki w grupie kontrolnej (Tabela 14).

Tabela 14. Charakterystyka pacjentek, których próbki zostały włączone do badań metodą RT-qPCR.

	<i>Kontrola</i>	<i>Rak</i>		
		<i>jajnika</i>		
Liczba pacjentek	46	42		
	<i>Średnia ± SD</i>	<i>Min</i>	<i>Max</i>	<i>Wartość p</i>
Wiek w chwili diagnozy (lata)				
Zdrowi	59±7	43	77	0,08
Chorzy	63±13	38	86	
BMI (waga/wzrost²)				
Zdrowi	27±4	21	39	0,7
Chorzy	28±5	19	42	
	<i>FIGO I</i>	<i>FIGO II</i>	<i>FIGO III-IV</i>	<i>N/A</i>
Liczba przypadków	3	1	34	4

Do reakcji odwrotnej transkrypcji zastosowano zestaw odczynników mirCURY LNA RT Kit (Qiagen, Germany), natomiast do właściwej reakcji PCR w czasie rzeczywistym zastosowano zestaw miR-CURY LNA SYBR Green PCR Kit (Qiagen, Germany). Reakcję przeprowadzono z użyciem następujących starterów: dla miR-1246, MIMAT0005898: 5'AAUGGAUUUUUGGAGCAGG ; dla miR-150-5p, MIMAT0000451 : 5'UCUCCCAACCCUUGUACCAGUG. Zastosowano dwa referencyjne miRNA: miR-103-3p, MIMAT0000101: 5'AGCAGCAUUGUACAGGGCUAUGA oraz miR-199b-5p, MIMAT0000263: 5TCCAGUGUUUAGACUAUCUGUUC. Profil temperaturowy RT- qPCR był następujący: 2 min w temp. 95°C i 50 cykli: 10 s w temp. 95°C i 60 s w temp. 56°C. Obliczono wydajność reakcji dla każdej pary starterów poprzez przygotowanie serii rozcieńczeń matrycy. Następnie wyznaczono cykle progowe (C_t) PCR miRNA badanego i referencyjnego dla prób badanych i kalibratora. Względny poziom ekspresji badanego miRNA wyznaczono według wzoru:

$$R = \frac{(E_{\text{gen badany}})^{\Delta C_p \text{ gen badany (kontrola-próba)}}}{(E_{\text{gen kontrolny}})^{\Delta C_p \text{ gen kontrolny (kontrola-próba)}}$$

Wykazano, że miR-1246 wykazuje podwyższoną ekspresję w surowicy kobiet z rakiem jajnika, w odróżnieniu od miR-150-5p, którego ekspresja u kobiet chorych na raka jajnika jest obniżona. Wyniki przedstawiono na Fig. 7.

Bazując na znormalizowanych danych RT-qPCR został utworzony klasyfikujący model diagnostyczny (rozkład dwumianowy GLM z pakietu caret [M. Kuhn, „Building Predictive Models in R Using the caret Package”, *J Stat. Softw.*, 2008.]), który został wytrenowany na zbiorze treningowym zawierającym 70% danych oraz zwalidowany na zbiorze testowym (30% danych). W celu zbudowania stabilnego klasyfikującego modelu diagnostycznego w trakcie trenowania został użyty sprawdzian krzyżowy – metoda leave-one-out crossvalidation. Zostało utworzonych tyle podzbiorów, aby każdy pacjent raz znalazł się w grupie testowej, reszta w zbiorze treningowym, jak również został utworzony model regresji logistycznej na podstawie każdego z podzbiorów. Wyniki tych analiz zostały następnie uśrednione w celu uzyskania jednego, najbardziej optymalnego wyniku.

Ocenę właściwej klasyfikacji wykonano za pomocą wykresów ROC, wyznaczając AUC wraz z przedziałem ufności, poziomem czułości i swoistości w punkcie odcięcia oraz R^2 i RMSE. Obliczenia jak i wykresy zostały wykonane z użyciem oprogramowania R wersja 3.6.1 [R. C. T. R. Foundation, „R: A Language and Environment for Statistical Computing”, vol. 2, <https://www.R-project.org>, 2013.] oraz pakiety pROC [X. Robin i wsp., „pROC: An open-source package for R and S+ to analyze and compare ROC curves”, *BMC Bioinformatics*, vol. 12, nr 1, str. 77, marzec 2011], Optimal Cutpoints [M. López-Ratón, M. X. Rodríguez-Álvarez, C. Cadarso-Suárez, and F. Gude-Sampedro, „Optimalcutpoints: An R package for selecting optimal cutpoints in diagnostic tests”, *J. Stat. Softw.*, vol. 61, nr 8, str. 1–36, listopad 2014.]. Punkt odcięcia został obliczony za pomocą metody Youdena na zbiorze treningowym. W tym punkcie została obliczana czułość i swoistość na zbiorze testowym (Tabela 15).

Tabela 15. Zmienne w zbudowanym modelu dla danych z RT-qPCR. a_0 – stała, a_1 – współczynnik predyktora x_1 , a_2 – współczynnik predyktora x_2 .

Współczynniki	a_0	a_1	a_2
$x_1 = \text{miR-1246}$	55,160	-1,616	4,277
$x_2 = \text{miR-150-5p}$			

Klasyfikujący model diagnostyczny został utworzony na podstawie znormalizowanych danych poziomu ekspresji miRNA w surowicy uzyskanych metodą RT-qPCR, stosując wzór.

$$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150) = \frac{e^{55,160 - 1,616 \cdot \text{miR} - 1246 + 4,277 \cdot \text{miR} - 150 - 5p}}{1 + e^{55,160 - 1,616 \cdot \text{miR} - 1246 + 4,277 \cdot \text{miR} - 150 - 5p}}$$

w którym

$P(Y = 1 | \text{miR} - 1246, \text{miR} - 150)$ to warunkowe prawdopodobieństwo, że zmienna zależna Y przyjmie wartość równą 1 dla wartości zmiennych niezależnych $\text{miR} - 1246$, $\text{miR} - 150$ e to liczba Eulera, $\approx 2,718$.

Punkt odcięcia w zakresie $0,1 < x \leq 0,8$ daje wynik dla czułości (S) oraz swoistości (Sp) na zbiorze testowym $> 85\%$ (Tabela 16).

Tabela 16. Punkty odcięcia w zakresie $0 < x \leq 0,9$ oraz wartości czułości i swoistości.

Punkt odcięcia	Swoistość	Czułość
0,0	0	100,0
0,1	91,66667	100,0
0,2	91,66667	100,0
0,3	91,66667	100,0
0,4	91,66667	92,85714
0,5	91,66667	92,85714
0,6	91,66667	92,85714
0,7	91,66667	92,85714
0,8	91,66667	85,71429
0,9	91,66667	78,57143

Obliczono również podstawowe parametry jakości w celu oceny jakości klasyfikacji modelu diagnostycznego, które przedstawiono w poniższej Tabeli 17.

Tabela 17. Parametry jakości dla modeli na zbiorach treningowym i testowym.

Nazwa	miR-1246, miR-150-5p	miR-1246, miR-150-5p
Zbiór	Treningowy	Testowy
AUC	99,7 %	94,6 %
CI dolna granica	99,0 %	83,9 %
CI górna granica	100%	100%
Optymalny punkt odcięcia (Youden Index)	0.2	-
Czułość	96,4 %	100 %
Swoistość	94,1 %	91,7 %

Krzywa ROC pomaga zwizualizować potencjał diagnostyczny. Na Fig. 8 przedstawione są krzywe ROC dla wytrenowanego modelu, jedna dla zbioru treningowego oraz druga dla zbioru testowego. Na każdej z krzywych został zaznaczony punkt odcięcia obliczony metodą statystyki i Youdena oraz odpowiadająca temu punktowi czułość i swoistość.

Następnie została sporządzona tablica pomyłek (Tabela 18) na danych, które nie zostały włączone podczas trenowania modelu. W Tabeli 16 przedstawione są informacje o ilości wyników prawdziwie pozytywnych (TP), prawdziwie negatywnych (TN), fałszywie pozytywnych (FP) oraz fałszywie negatywnych (FN) po klasyfikacji.

Tabela 18. Tablica pomyłek w zbiorze testowym dla modelu z miRNA-1246 i miRNA-150-5p.

		Stan faktyczny	
		0	1
Predykcja	0	11 (TP)	0 (FN)
	1	1 (FP)	14 (TN)

Zastrzeżenia patentowe

1. Diagnostyczny zestaw biomarkerów miRNA składający się z biomarkerów miRNA: miR-1246 oraz miR-150-5p.
2. Diagnostyczny zestaw biomarkerów miRNA składający się z biomarkerów miRNA: miR-1246 oraz miR-150-5p do zastosowania do diagnostyki raka jajnika.
3. Sposób diagnozowania *in vitro* raka jajnika u osobnika, **znamienny tym**, że obejmuje następujące etapy:
 - i) oznaczania poziomów ekspresji diagnostycznego zestawu składającego się z dwóch biomarkerów miRNA: miR-1246 i miR-150-5p w próbce od osobnika i
 - ii) porównania poziomów ekspresji oznaczonych w etapie i) z poziomami ekspresji miR-1246 i miR-150-5p u osobnika bez raka jajnika, przy czym porównanie dostarcza wskaźnika diagnostycznego określającego czy osobnik ma raka jajnika, przy czym wzrost poziomu ekspresji miR-1246 względem poziomu ekspresji miR-1246 u osobnika bez raka jajnika i spadek poziomu ekspresji miR-150-5p względem poziomu ekspresji miR-150-5p u osobnika bez raka jajnika wskazują na raka jajnika u osobnika.
4. Sposób według zastrz. 3, **znamienny tym**, że poziomy ekspresji to poziomy znormalizowane.
5. Sposób według zastrz. 3 albo 4, **znamienny tym**, że porównanie poziomów ekspresji przeprowadza się z zastosowaniem klasyfikującego modelu diagnostycznego, który klasyfikuje próbkę jako próbkę od osobnika z rakiem jajnika lub jako próbkę od osobnika bez raka jajnika.
6. Sposób według zastrz. 5, **znamienny tym**, że klasyfikujący model diagnostyczny przeprowadza porównanie poziomów ekspresji z wykorzystaniem zestawu danych zawierającego dane o poziomie ekspresji zestawu markerów miRNA obejmującego miR-1246 i miR-150-5p.
7. Sposób według dowolnego z zastrz. 1 do 6, **znamienny tym**, że poziom ekspresji zestawu biomarkerów miRNA oznacza się z zastosowaniem metody pomiaru ekspresji wybranej spośród ilościowej reakcji odwrotnej transkrypcji z łańcuchową reakcją polimerazy (qPCR), korzystnie z analizą ilości produktu w czasie rzeczywistym (RT-qPCR), metody NanoString i mikromacierzy.
8. Sposób według zastrz. 7, **znamienny tym**, że klasyfikujący model diagnostyczny wykorzystuje do klasyfikacji optymalny punkt odcięcia wybrany spośród: 0,1–0,8 w przypadku pomiaru ekspresji metodą RT-qPCR; 0,3–0,9 w przypadku pomiaru ekspresji metodą NanoString oraz wartość 0,5 w przypadku pomiaru ekspresji z wykorzystaniem mikromacierzy.
9. Sposób według zastrz. 8, **znamienny tym**, że do określania poziomu ekspresji miRNA stosuje się metodę RT-qPCR.
10. Sposób według zastrz. 9, **znamienny tym**, że do normalizacji wyników stosuje się poziom ekspresji referencyjnego miRNA, korzystnie miR-103-3p i/lub miR-199b-5p.
11. Sposób według zastrz. 10, **znamienny tym**, że normalizację wyników uzyskuje się stosując wzór:

$$\text{deltaCt} = (\text{Ct miR-1246/miR-150-5p} - \text{Ct miR-103-3p}),$$
 w którym
 delta Ct oznacza zmianę wartości cyklu progowego
 Ct oznacza wartość cyklu progowego.
12. Sposób według zastrz. 11, **znamienny tym**, że uzyskana znormalizowana wartość deltaCt po porównaniu z punktem odcięcia klasyfikuje osobnika, od którego pobrano próbkę jako osobnika z rakiem jajnika lub bez raka jajnika.

13. Sposób według dowolnego z zastrz. 3 do 12, **znamienny tym**, że próbkę stanowi próbka surowicy.
14. Sposób według dowolnego z zastrz. 3 do 13, **znamienny tym**, że raka jajnika stanowi nisko-zróżnicowany surowiczy rak jajnika.
15. Zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono w zastrz. 1 do diagnozowania *in vitro* raka jajnika.
16. Zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono w zastrz. 1 w diagnostycznym badaniu przesiewowym *in vitro* pod kątem występowania raka jajnika.
17. Zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono w zastrz. 1 do oceny skuteczności leczenia raka jajnika.
18. Zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono w zastrz. 1 do monitorowania odpowiedzi na leczenie raka jajnika.
19. Zastosowanie diagnostycznego zestawu biomarkerów miRNA jak określono w zastrz. 1 do przewidywania wznowy raka jajnika po ukończonym leczeniu raka jajnika.
20. Zestaw testowy do diagnostyki *in vitro* raka jajnika, **znamienny tym**, że zawiera środki do ilościowego oznaczania poziomu ekspresji diagnostycznego zestawu składającego się z dwóch biomarkerów miRNA: miR-1246 i miR-150-5p oraz instrukcje do przeprowadzenia sposobu według dowolnego z zastrz. 3 do 14, przy czym jako środki do ilościowego oznaczania ekspresji zestawu biomarkerów miRNA: miR-1246 i miR-150-5p zawiera odczynniki i startery do powielania w reakcji RT-qPCR.
21. Zestaw testowy do diagnostyki według zastrz. 20, **znamienny tym**, że ponadto zawiera środki do ilościowego oznaczania ekspresji referencyjnego miRNA, korzystnie zestawu miR-103-3p i/lub miR-199b-5p.

Rysunki

Fig. 1.

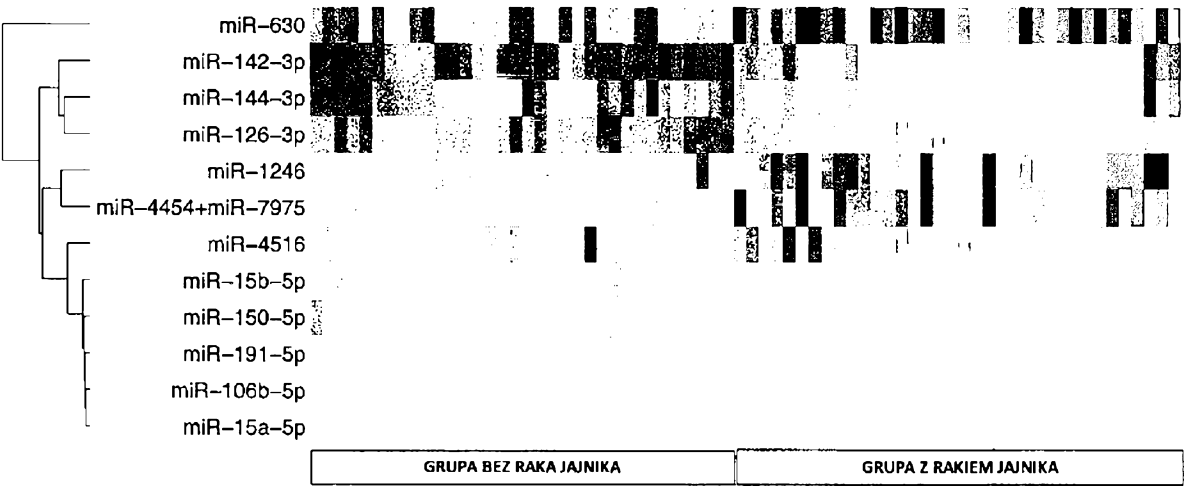


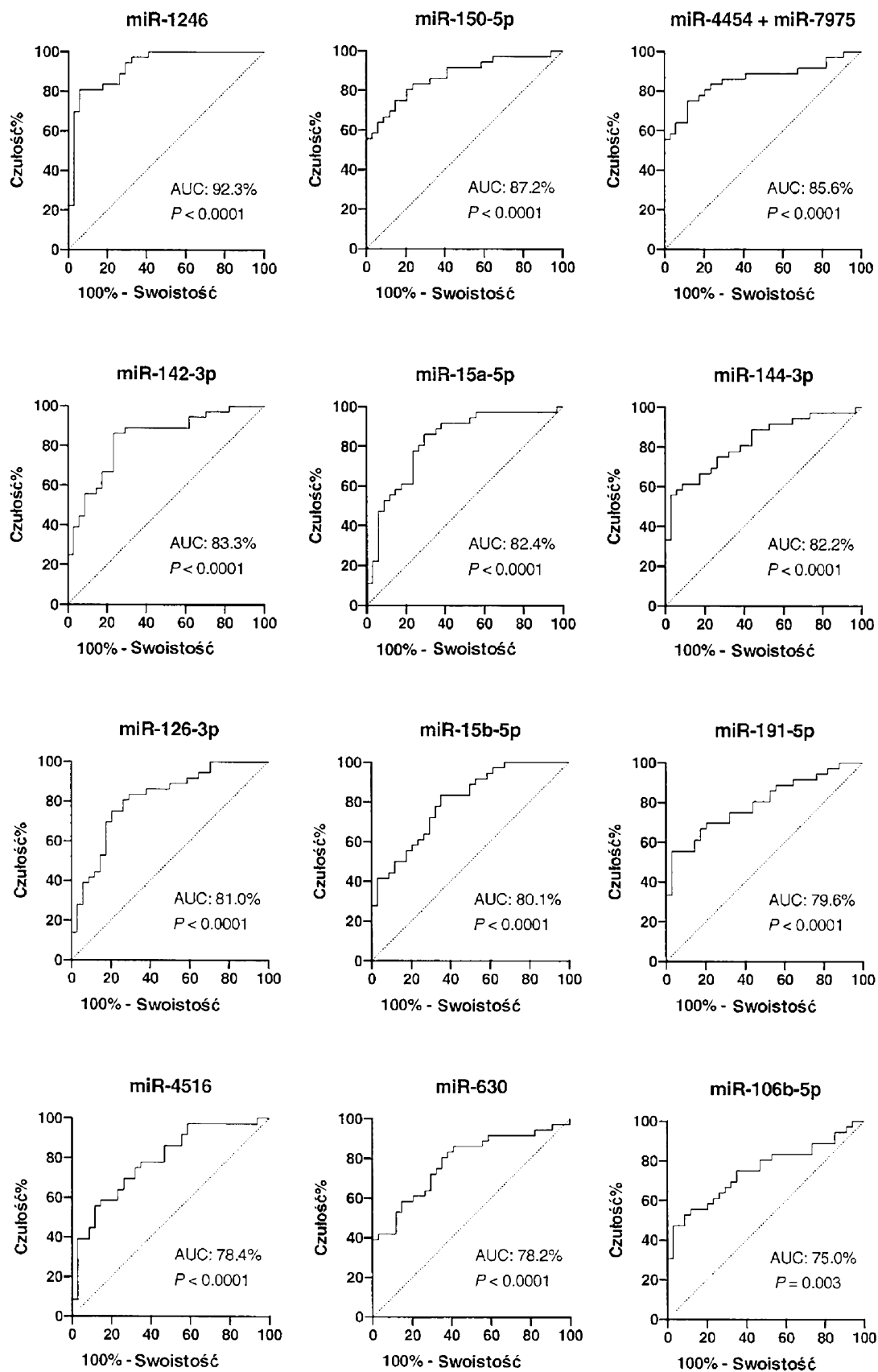
Fig. 2

Fig. 3

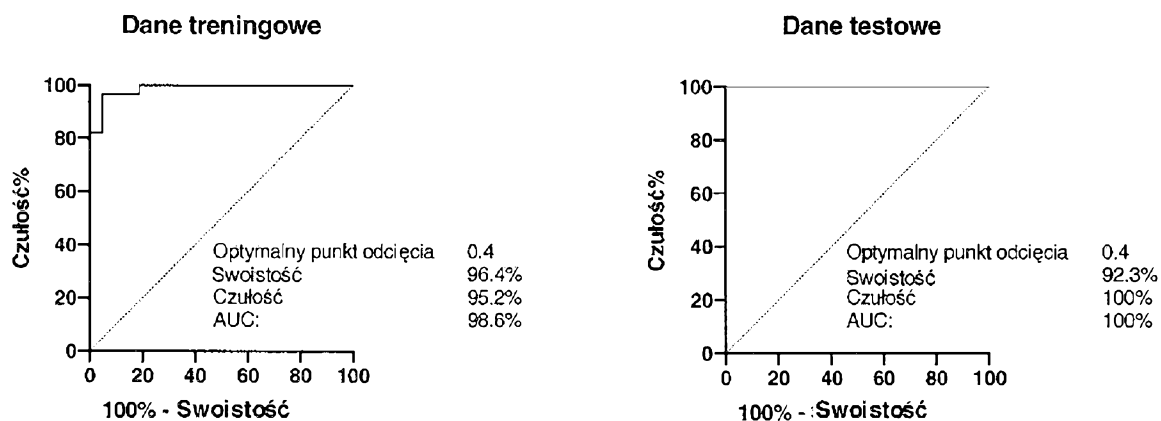


Fig. 4

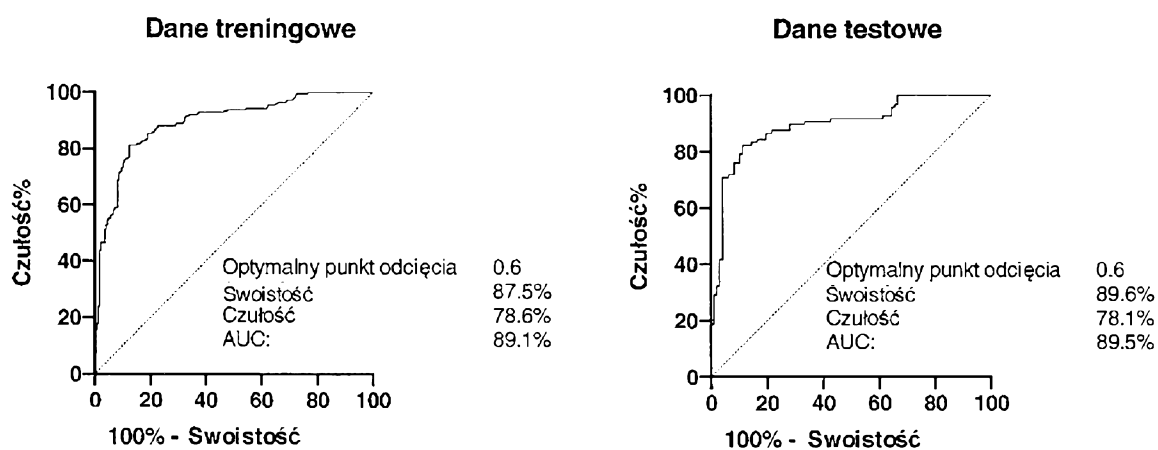


Fig. 5

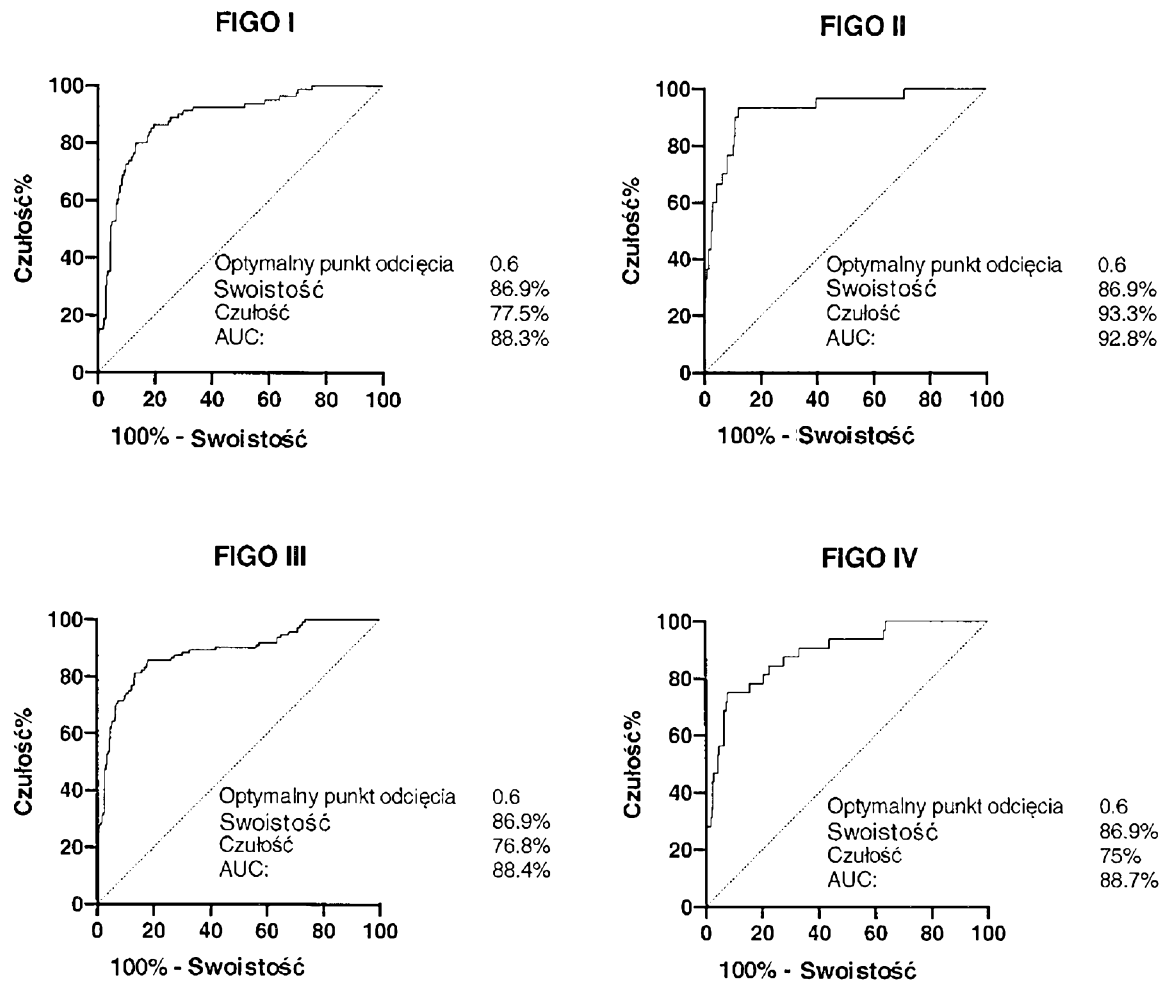


Fig. 6

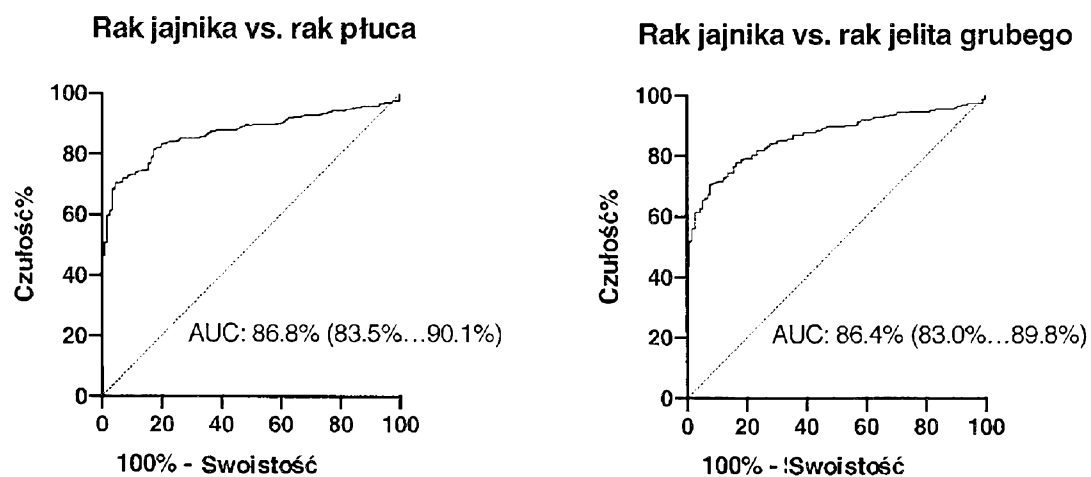


Fig. 7

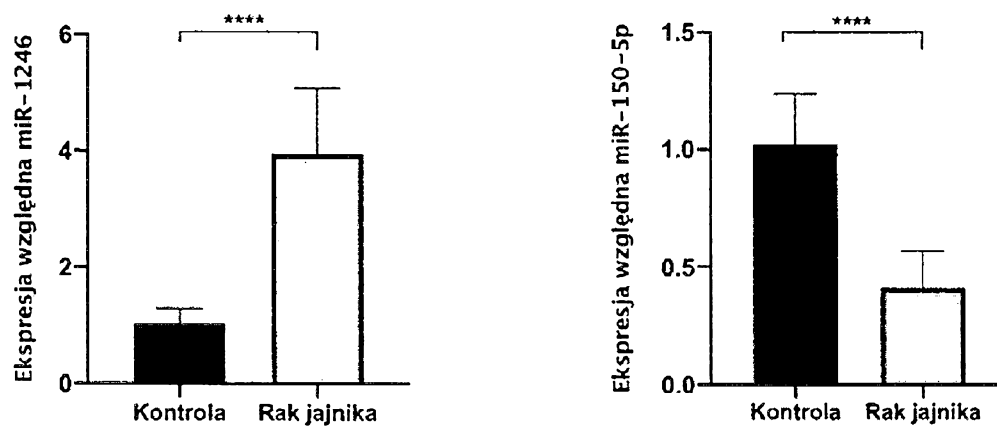


Fig. 8