

Szacowanie prawdopodobieństwa złośliwości guza na podstawie kombinacji testów we wstępnej diagnostyce różnicowej guzów jajnika

The estimation of the probability of tumor malignancy on the basis of test combination in the primary diagnosis of adnexal tumors

Smoleń Agata¹, Stachowicz Norbert², Czekierowski Artur², Kotarski Jan²

¹ Zakład Matematyki i Biostatystyki Medycznej, Uniwersytet Medyczny w Lublinie,

² I Katedra i Klinika Ginekologii Onkologicznej i Ginekologii, Uniwersytet Medyczny w Lublinie

Streszczenie

Cel pracy: Ocena prawdopodobieństwa złośliwości guza jajnika w oparciu o analizę regresji logistycznej i probitowej oraz konstrukcja praktycznego narzędzia, które można wykorzystać przy łóżku chorej.

Materiał i metody: Badania zostały przeprowadzone w Klinice Ginekologii Onkologicznej i Ginekologii Akademii Medycznej w Lublinie w latach 2002-2007. Do analizy zakwalifikowano 412 kobiet. Analizowano wyniki badania ultrasonograficznego z zastosowaniem skali szarości (2D), oceny przepływu z zastosowaniem kolorowego i spektralnego Dopplera oraz badania z obrazowaniem 3D sonoangiografii i stężenie antygenu CA-125. Wśród zakwalifikowanych chorych było 170 (41,3%) kobiet z nowotworami złośliwymi oraz 242 (58,7%) z niezłośliwymi guzami jajnika.

W celu podwyższenia czułości i specyficzności morfologii ultrasonograficznej, statusu menopauzalnego oraz stężenia CA-125 przed leczeniem operacyjnym obliczono indeks złośliwości guza jajnika RMI i porównano do wyników analizy regresji logistycznej. Ponadto skonstruowano nomogramy, służące do obliczania prawdopodobieństwa istnienia procesu złośliwego w badanym guzie jajnika, które można wykorzystać przy łóżku chorej.

Wyniki: Najwyższe wartości pola powierzchni pod krzywą ROC w jednoczynnikowej analizie uzyskano dla stężenia antygenu CA-125 (0,87) i cechy „Kolor” (0,83). Dla indeksu RMI wartość pola powierzchni wynosiła 0,91. Wykorzystanie analizy regresji logistycznej pozwoliło stwierdzić, że w wieloczynnikowym modelu istotnymi parametrami okazały się: wyrośla endofityczne, cecha „Kolor”, wiek, indeks PI i stężenie markera CA-125. Uzyskana dokładność klasyfikacji dla modelu MLRA w zbiorze uczącym wynosiła 88%, a w zbiorze testowym 92%. Wartość pola pod krzywą ROC wynosiła 0,94 i różniła się istotnie od pola dla CA-125 ($p < 0,05$).

Wniosek: Zastosowanie kombinacji parametrów klinicznych ultrasonograficznych i stężenia antygenu CA-125 w modelu prognostycznym pozwala na wspomaganie procesów decyzyjnych we wstępnym różnicowaniu guzów jajnika.

Słowa kluczowe: rak jajnika / neoangiogeneza / sonoangiografia /
/ regresja logistyczna /

Adres do korespondencji:

Agata Smoleń
Zakład Matematyki i Biostatystyki Medycznej, Uniwersytet Medyczny w Lublinie,
20-950 Lublin, ul. Jaczewskiego 4
tel. +4881 7423646
e-mail: absmolen@wp.pl

Otrzymało: 17.07.2009
Zaakceptowano do druku: 01.03.2010

Summary

Purpose: The aim was to evaluate the risk of ovarian tumor malignancy based on logistic regression analysis and to construct a practical tool which might be used at the bedside.

Material and methods: 379 women with adnexal masses were enrolled to the study. The results of the two-dimensional (2D) gray-scale ultrasound examination, color and spectral Doppler blood flow measurement, three-dimensional (3D) sonoangiography examination and serum levels of CA-125 measurement were analyzed. 160 out of 379 women with abnormal findings, (42.2%), were diagnosed with a malignant ovarian tumor, and 219 (57.8%) were found to have a non-malignant adnexal mass. To improve the sensitivity and specificity of ultrasound scan findings, postmenopausal status and preoperative serum CA-125 levels, the risk of malignancy index (RMI) was calculated and compared to the result of logistic regression analysis. Furthermore, the nomograms applicable at the bedside for estimation of the probability of malignancy for the examined adnexal tumor were derived.

Results: The highest values of the area under the receiver operating characteristic (ROC) curves in univariate analysis were 0.87 for serum levels of CA-125 and 0.83 for the color score. For the RMI, the value of the area under ROC curve was 0.91. The results of logistic regression revealed that papillary growth, color score, age, pulsatility index (PI) and level of serum CA 125 were significant factors in the multivariate model. The obtained classification accuracies for MLRA model in the training set and the test set were 88% and 92%, respectively. The value of the area under the ROC curve was 0.94 and significantly differed from the value of area for CA-125 ($p < 0.05$).

Conclusions: The results of the study suggest that the combined use of ultrasound parameters and serum level of CA-125 in the prognostic model improves the diagnostic precision in the primary discrimination of benign and malignant ovarian tumors in women.

Key words: **ovarian carcinoma / angiogenesis / sonography / logistic regression /**

Wstęp

Guzy złośliwe przydatków stanowią jeden z najtrudniejszych problemów onkologii ginekologicznej zarówno w odniesieniu do diagnostyki jak i leczenia chorych. Brak możliwości pewnego przedoperacyjnego różnicowania każdego guza w miednicy mniejszej powoduje podejmowanie prób oceny przydatności badania klinicznego, oznaczeń markerów i nowoczesnych badań ultradźwiękowych i zastosowania ich w kombinacji [1, 2].

Cel pracy

Ocena prawdopodobieństwa istnienia w guzie jajnika nowotworu złośliwego ma na celu podjęcie dla danej chorej optymalnej decyzji dotyczącej sposobu terapii i wyboru leczenia zachowawczego lub operacyjnego. W związku z tym celem pracy była ocena prawdopodobieństwa złośliwości guza jajnika w oparciu o analizę regresji logistycznej i probitowej oraz konstrukcja praktycznego narzędzia, które można wykorzystać przy łóżku chorej.

Materiał i metody

Badania zostały przeprowadzone w Klinice Ginekologii Onkologicznej i Ginekologii Akademii Medycznej w Lublinie w latach 2002-2007.

Do analizy zakwalifikowano 412 kobiet. Analizowano wyniki badania ultrasonograficznego z zastosowaniem skali szarości (2D), oceny przepływu z zastosowaniem kolorowego i spektralnego Dopplera oraz badania z obrazowaniem 3D sonoangiografii i stężenie antygenu CA-125. Wśród zakwalifikowanych chorych było 170 (41,3%) kobiet z nowotworami złośliwymi oraz 242 (58,7%) z niezłośliwymi guzami jajnika. Weryfikację rodzaju guza stanowiły wyniki pooperacyjnego badania histopatologicznego. Guzy o granicznej złośliwości i guzy przerzutowe do gonad zakwalifikowano do grupy nowotworów złośliwych jajnika. Wśród pacjentek z rakiem jajnika stopień zaawansowania klinicznego wg FIGO przedstawiał się następująco: 38 chorych było w I stopniu, 11 w stopniu II, pozostałe były w III (113 chorych) lub w IV stopniu zaawansowania (8 chorych).

Ocena stopnia zróżnicowania histologicznego nowotworu wg WHO wykazała, że 14 guzów było w stopniu G1, 72 w stopniu G2 oraz 73 w stopniu G3, dla 11 przypadków stopień nie został określony. Dla każdej pacjentki oceniono cechy kliniczne i ultrasonograficzne wymienione w tabeli I.

Tabela I. Charakterystyka wybranych cech klinicznych i ultrasonograficznych badanej grupy kobiet.

Oceniane cechy
Wiek
Status menopauzalny
Wskaźnik masy ciała BMI (kg/m ²)
Umiejscowienie guza (jedno/obustronne)
Typ histologiczny guza wg WHO
Stopień zaawansowania guza złośliwego wg FIGO
Przegrody w guzie (>3mm lub brak)
Elementy lite w guzie (obecne/brak)
Wyrośla endofityczne w guzie (>3mm lub brak)
Największy wymiar guza (mm)
Objętość guza (ml)
Lokalizacja naczyń (centralna/obwodowa/brak naczyń)
Kolor wg IOTA
Indeks pulsacji PI
Indeks oporu RI
Najwyższa prędkość skurczowa PSV (cm/s)
Średnia maksymalna szybkość przepływu TAMXV (cm/s)
Indeks waskularyzacji VI
Indeks przepływu FI
Indeks naczyniowo-przepływowy VFI
Stężenie antygenu CA-125 (U/ml)

Szacowanie prawdopodobieństwa złośliwości guza na postawie kombinacji testów...

Oznaczenia stężenia antygenu CA-125 wykonano metodą immunoenzymatyczną, za wynik nieprawidłowy dla tak oznaczonego poziomu antygenu CA 125 w surowicy krwi przyjęto wartość powyżej 35U/ml. W celu podwyższenia czułości i specyficzności morfologii ultrasonograficznej (U), statusu menopauzalnego (M) oraz stężenia CA-125 przed leczeniem operacyjnym zastosowano indeksy złośliwości guza jajnika RMI (*Risk of Malignancy Index*) [3, 4]. Przyjmuje się, że wynik RMI powyżej 200 sugeruje złośliwy charakter zmiany.

Dokładność testu została określona przez porównanie pomiędzy wynikami testu a pooperacyjnym wynikiem histopatologicznym, który stwierdza rzeczywistą sytuację chorobową. Sposoby grupowania danych i oceny wartości diagnostycznej podstawowych parametrów przedstawia tabela II.

Tabela II. Tablica kontyngencji przedstawiająca związek badanego czynnika ryzyka z pooperacyjną weryfikacją histopatologiczną.

Wynik testu	Rozpoznanie kliniczne		Razem
	Guz złośliwy	Guz niezłośliwy	
Pozytywny	TP	FP	TP+FP
Negatywny	FN	TN	FN+TN
Razem	TP+FN	FP+TN	TP+FP+FN+TN

Skróty w tabeli oznaczają:

- TP – wyniki prawdziwie pozytywne,
- FP – wyniki fałszywie pozytywne,
- FN – wyniki fałszywie negatywne,
- TN – wyniki prawdziwie negatywne.

Zastosowane podstawowe miary dokładności testu to:

- Czułość (*sensitivity* – SENS):

$$SENS = \frac{TP}{TP + FN} \cdot 100\%$$

- Specyficzność (*specifity* – SPEC):

$$SPEC = \frac{TN}{FP + TN} \cdot 100\%$$

- Dodatnią wartość predykcyjną (*positive predictive value* – PPV):

$$PPV = \frac{TP}{TP + FP} \cdot 100\%$$

- Ujemną wartość predykcyjną (*negative predictive value* – NPV):

$$NPV = \frac{TN}{FN + TN} \cdot 100\%$$

- Dokładność (*accuracy* – ACC):

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \cdot 100\%$$

- Wskaźnik wiarygodności dodatniego wyniku testu (LR+):

$$LR+ = \frac{SENS}{1 - SPEC}$$

- Wskaźnik wiarygodności ujemnego wyniku testu (LR-):

$$LR- = \frac{1 - SENS}{SPEC}$$

Przyjmuje się, że test ma rzeczywistą wartość diagnostyczną jeżeli LR wynosi około 10 i więcej lub około 0,1 i mniej. Wyniki między 5 a 10 oraz między 0,1 a 0,2 wskazują, że test jest przydatny. LR między 0,5 a 2 oznacza, że test nie wpływa w sposób istotny na ocenę prawdopodobieństwa złośliwości guza [5].

W sytuacjach, kiedy wartości testu diagnostycznego były ciągle lub miały wiele kategorii obliczono wartości predykcyjne dla najlepiej różnicujących charakter guza jajnika wartości granicznych (*cutoff*). W przeprowadzonych obliczeniach zmiany czułości i specyficzności testów przy przesuwaniu wartości granicznej przedstawiono w formie krzywych ROC (*Receiver Operating Characteristic Curves*) oraz wykorzystano dodatkowe informacje porównując pola pod krzywą ROC [6, 7, 8].

Najbardziej użyteczną klinicznie metodą przedstawiania właściwości testu jest opisanie jego wyników za pomocą LR i/ lub pola powierzchni pod krzywą ROC [5, 9, 10].

W dalszej części pracy dokonano oceny czy i które z wybranych parametrów (zmienne niezależne) mają istotny statystycznie wpływ na prawdopodobieństwo istnienia nowotworu złośliwego lub niezłośliwego (zmienna zależna). Użyto klasyfikacji wzorcowej, której zadaniem było przypisanie każdego badanego przypadku do jednej z dwóch klas (guz złośliwy bądź niezłośliwy).

W celu znalezienia najlepszej kombinacji cech istotnie wpływających na istnienie procesu złośliwego i obliczenie indywidualnego prawdopodobieństwa istnienia nowotworu jajnika zastosowano regresję logistyczną i probitową.

Dostępny zbiór danych podzielono w sposób losowy na dwa zbiory: uczący – 355 chorych i walidacyjny – 24 chore. Trzeci zbiór testowy stanowiła prospektywna grupa 33 pacjentek. Metoda regresji logistycznej pozwoliła na obliczenie tzw. logarytmu wiarygodności, służącego do oceny dopasowania modelu wraz z wartością statystyki χ^2 i poziomem prawdopodobieństwa p . Gdy $p < 0,05$ analizowany model różnił się istotnie od modelu tylko z wyrazem wolnym. W celu oceny mocy predykcji uzyskanego modelu obliczono współczynniki $R^2_{McFaddena}$ i $R^2_{Nagelkerke}$ określających jak duża część zmienności analizowanej cechy zależnej powiązana jest ze zmiennością badanych parametrów niezależnych.

Ponadto obliczone zostały: wartości estymatorów parametrów modelu z 95% przedziałami ufności i asymptotycznymi błędami standardowymi. Istotność estymatorów parametrów modelu sprawdzono przy pomocy wartości testu t-Studenta lub wartości statystyki χ^2 Walda i poziomu prawdopodobieństwa p .

Obliczone zostały również: iloraz szans dla zmiany jednostkowej poszczególnych parametrów oraz iloraz szans dla zmiany równej zakresowi analizowanych zmiennych z 95% przedziałami ufności. W tzw. powiązaniu hierarchicznym zastosowano test ilorazu wiarygodności LR funkcji straty (największej wiarygodności) między poprzednim modelem a modelem po dodaniu kolejnej zmiennej dołączonej do modelu. Do oceny czy dołączenie zmiennej w sposób istotny poprawia dopasowanie modelu do danych użyto testu χ^2 i poziomu p dla precyzji dopasowania.

Ponadto zastosowano wstępującą analizę regresji probitowej z estymacją quasi-Newtona. Obliczono wartości χ^2 przy porównywaniu skonstruowanego modelu i modelu tylko z wyrazem wolnym. Obliczone zostały: wartości estymatorów parametrów modelu z asymptotycznymi błędami standardowymi. Istotność estymatorów parametrów modelu sprawdzono przy pomocy wartości testu t-Studenta i poziomu prawdopodobieństwa p .

Model regresji logistycznej przedstawiał się następująco:

$$p = \frac{1}{1 + e^{-z}}$$

gdzie e jest stałą matematyczną, granicą ciągu

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e = 2,71828\dots,$$

a z jest kombinacją liniową zmiennych x_i i ich estymatorów b_i zawartych w modelu: .

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n.$$

Wartości przewidywane należące do przedziału $<0,1>$ można traktować jako prawdopodobieństwo istnienia procesu złośliwego bądź niezłośliwego. W zastosowanym modelu wartości prawdopodobieństwa $\leq 0,5$ klasyfikowały nowotwór jako niezłośliwy, a przypadki z wartościami predykcyjnymi $>0,5$ jako nowotwór złośliwy. Przy pomocy klasyfikacji przypadków i ilorazu szans >1 oceniano czy klasyfikacja jest lepsza niż można by oczekiwać przez przypadek oraz obliczano wartości predykcyjne uzyskanego modelu [11].

Podsumowaniem analizy regresji logistycznej było skonstruowanie nomogramów w oparciu o zasadę suwaka logarytmicznego, służących do przekształcenia wyników analizy w praktyczne narzędzie, możliwe do wykorzystania przy łóżku chorego. Podziałka w dolnej części została skalibrowana w jednostkach prawdopodobieństwa, co pozwala na bezpośrednie odczytanie prawdopodobieństwa istnienia procesu złośliwego.

Konstrukcję nomogramów przeprowadzono zgodnie z zasadami podanymi przez Lubstena i wsp. w oparciu o pakiet matematyczny R, udostępniony na stronie internetowej: <http://novum.am.lublin.pl/CRAN/> [12].

Podsumowaniem uzyskanych wyników analiz była ocena poprawnie i niepoprawnie zaklasyfikowanych przypadków i obliczone wartości predykcyjne dla zbioru uczącego walidacyjnego i testowego oraz ogółem dla całej grupy badanej i obu zbiorów testujących skuteczność. Analizy statystyczne przeprowadzono z użyciem oprogramowania komputerowego Data Miner (StatSoft, Polska).

Wyniki

We wstępnej analizie wszystkich badanych parametrów najwyższe wartości dokładności diagnostycznej uzyskano dla CA-125 i indeksu złośliwości guza jajnika RMI. Wartości predykcyjne CA-125 i RMI dla najczęściej przyjmowanych wartości progowych w różnicowaniu złośliwych i niezłośliwych guzów przydatkowych przedstawia tabela III.

Dla wartości progowej indeksu RMI wynoszącej 200 uzyskano dokładność diagnostyczną 84%, dla antygenu CA-125 przy wartości 35U/ml lub 65U/ml uzyskano wynik 80% i 79%.

Najwyższą wartość pola pod krzywą ROC uzyskano dla stężenia antygenu CA-125 wynoszącą 0,87, a dla cechy „Kolor”-0,83. Połączenie wyniku badania ultrasonograficznego, statusu menopauzalnego i CA-125 w indeksie RMI zaproponowanym zarówno przez Jacobsa i wsp. oraz Tingustada i wsp. wykazało wartość pola powierzchni równą 0,91 [3, 4].

Porównanie najwyższych wartości pól pod krzywymi ROC: indeksu RMI vs stężenie CA-125 wykazało nieznaczną poprawę dokładności diagnostycznej ($p=0,08$). (Rycina 2).

Wyniki analizy regresji logistycznej i probitowej (MLRA i MPRA)

Wykorzystanie wstępującej analizy regresji logistycznej z estymacją quasi-Newtona z użyciem wszystkich parametrów diagnostycznych w badanej grupie kobiet pozwoliło stwierdzić, że spośród analizowanych kilkunastu zmiennych dołączenie 5 zmiennych tj. wyrosła endofityczne (PAP), cecha „Kolor”, wiek, indeks PI i stężenie markera CA-125 w sposób istotny poprawiało dopasowanie modelu do danych. Potwierdza to końcowa wartość funkcji straty oraz wartość przyrostowego testu $\chi^2=261,24$; $p<0,0001$ dla precyzji dopasowania. W związku z tym 5 wymienionych wyżej parametrów zostało uwzględnionych w równaniu regresji.

W celu oceny mocy predykcji uzyskanego modelu obliczono współczynniki $R^2_{McFaddena}$ równy 0,54 i $R^2_{Nagelkerke'a}$ wynoszący 0,7, które świadczą o przeciętnej możliwości przewidywania prawdopodobieństwa istnienia guza złośliwego.

Otrzymany model (MLRA) istotnie ($p<0,0001$) różnił się od modelu tylko z wyrazem wolnym. (Tabela IV).

Tabela IV przedstawia wartości uzyskanych parametrów modelu tj. wartości estymatorów parametrów modelu z błędami szacunku, wyniki testu t-Studenta i χ^2 Walda oraz ilorazy szans z 95% przedziałami ufności dla skonstruowanego modelu.

Wyliczone wartości estymatorów parametrów modelu, których istotność potwierdzają wyniki testu t-Studenta i Walda χ^2 pozwalają opisać model wraz z błędami szacunku.

Model ten opisany poprzez uwzględniane zmienne przedstawia się następująco:

prawdopodobieństwo istnienia procesu złośliwego

$$P(X) = \frac{1}{1 + e^{-z}}, \quad \text{gdzie } e \text{ jest stałą matematyczną, a}$$

$$\begin{aligned} \log itP = z = & -8,756 + 1,313 \cdot \text{PAP} + 1,169 \cdot \text{KOLOR} \\ & + 0,066 \cdot \text{WIEK} - 0,872 \cdot \text{PI} + 2,331 \cdot \text{CA125} \end{aligned}$$

Tabela III. Wartości prognostyczne CA-125 i RMI dla najczęściej przyjmowanych wartości progowych.

Parametr (optymalna wartość graniczna)	TP	FP	FN	TN	SENS	SPEC	LR+	LR-	PPV	NPV	ACC
CA-125 (35)	128	49	32	170	0,80	0,78	3,58	0,26	0,72	0,84	0,79
CA-125 (65)	114	28	46	191	0,71	0,87	5,57	0,33	0,80	0,81	0,80
RMI (200)	130	31	30	188	0,81	0,86	5,74	0,22	0,81	0,86	0,84

Szacowanie prawdopodobieństwa złośliwości guza na podstawie kombinacji testów..

Tabela IV. Wyniki analizy regresji logistycznej analizowanych parametrów (MLRA).

		Model: Regr. logistyczna (logit) N guz niezłośliwy: 207 guz złośliwy: 148 Zmn. zal.: GUZ Strata: Największe prawd. bł. średnkw. skal. Całkowita strata: 110,52090849 Chi2(5)=261,24 p<0,0001 Warunek uwzględniania: grupy="u"					
N=355		Stała B0	PAP	KOLOR	WIEK	PI	CA125_p
Ocena		-8,756	1,313	1,169	0,066	-0,872	2,331
Błąd standard.		1,047	0,382	0,176	0,012	0,335	0,362
t(349)		-8,364	3,432	6,661	5,443	-2,608	6,445
poziom p		0,000	0,001	0,000	0,000	0,010	0,000
-95%CL		-10,815	0,561	0,824	0,042	-1,530	1,620
+95%CL		-6,697	2,065	1,514	0,090	-0,214	3,043
Chi-kwadrat Walda		69,956	11,781	44,372	29,630	6,800	41,540
poziom p		0,000	0,001	0,000	0,000	0,009	0,000
Iloraz szans z jedn.		0,000	3,717	3,220	1,069	0,418	10,291
-95%CL		0,000	1,752	2,280	1,043	0,216	5,053
+95%CL		0,001	7,886	4,547	1,095	0,807	20,962
Iloraz szans zakr.			3,717	107,442	248,886	0,063	10,291
-95%CL			1,752	27,006	33,906	0,008	5,053
+95%CL			7,886	427,460	1826,939	0,507	20,962

Tabela V. Wyniki analizy regresji probitowej analizowanych parametrów (MPRA).

		Model: regresja probit N guz niezłośliwy: 207 guz złośliwy: 148 Zmn. zal.: GUZ Strata: Największe prawd. bł. średnkw. skal. Całkowita strata: 111,32265944 Chi2(5)=259,64 p=0,0000 Warunek uwzględniania: grupy="u"					
N=355		Stała B0	PAP	KOLOR	WIEK	PI	CA125_p
Ocena		-4,882	0,698	0,646	0,037	-0,473	1,291
Błąd std		0,534	0,205	0,093	0,007	0,179	0,195
t(349)		-9,147	3,404	6,974	5,630	-2,642	6,615
poziom p		0,000	0,001	0,000	0,000	0,009	0,000

Tabela VI. Wartości prognostyczne modelu skonstruowanego z użyciem analizy regresji logistycznej (MLRA).

Model MLRA	TP	FP	FN	TN	SENS	SPEC	LR+	LR-	PPV	NPV	ACC	
Zbiór	U	125	21	23	186	0,84	0,90	8,33	0,17	0,86	0,89	0,88
	W	10	0	2	12	0,83	1,00	-	0,17	1,00	0,86	0,92
	T	9	3	1	20	0,90	0,87	6,90	0,12	0,75	0,95	0,88
	U+W	135	21	25	198	0,84	0,90	8,80	0,17	0,87	0,89	<u>0,88</u>
	W+T	19	3	3	32	0,86	0,91	10,08	0,15	0,86	0,91	<u>0,89</u>

Dodatkowo wartości parametrów z odpowiednimi przedziałami ufności odpowiadające badanym zmiennym oznaczają, że ich wzrost zwiększa prawdopodobieństwo istnienia nowotworu złośliwego w badanym guzie.

Ilorazy szans jednostkowe i ich 95% przedziały ufności wskazują proporcjonalną zmianę prawdopodobieństwa istnienia procesu złośliwego, jeżeli dany, uwzględniony w równaniu parametr wzrośnie o jednostkę, natomiast ilorazy szans z zakresu i ich przedziały ufności mówią o proporcjonalnej zmianie tego prawdopodobieństwa, które spowoduje maksymalną zmianę danego parametru. Klasyfikację badanych guzów w modelu regresji logistycznej MLRA przedstawia tabela VI.

Potwierdzenie uzyskanego wyniku otrzymano konstruując model probitowy. Podobnie jak w przypadku modelu logitowego wartość funkcji straty bardzo znacznie większa od 0 i wartość testu $\chi^2=259,64$ przy porównaniu tego modelu i modelu tylko z wyrazem wolnym jest statystycznie istotna ($p<0,0001$). Estymatory parametrów modelu wraz z błędami szacunku i poziomami istotności przedstawia tabela V.

Przewidywane prawdopodobieństwa w zbiorze uczącym wykazywały dużą zgodność wyników uzyskanych z użyciem modelu logistycznego i probitowego. Weryfikację uzyskanego modelu i jego przydatności diagnostycznej potwierdziła klasyfikacja przypadków zbioru testowego.

Uzyskano wyniki w 100% zgodne z wynikami otrzymanymi w przypadku modelu logistycznego.

Podsumowanie wyników analizy regresji logistycznej (dla regresji probitowej uzyskano wyniki identyczne) stanowią wartości prognostyczne obliczone dla modelu (MLRA) w zbiorze uczącym (U), walidacyjnym (W) i testowym (T) oraz ogółem dla grupy badanej (U+W) i ogółem w zbiorach testujących skuteczność (W+T). Uzyskane wyniki przedstawia tabela VI.

Uzyskana dokładność klasyfikacji dla modelu MLRA w zbiorze uczącym wynosiła 88%, w zbiorze walidacyjnym 92%, co dało ogólną dokładność wynoszącą 88%. Ocenę trafności prognostycznej z użyciem skonstruowanego modelu przeprowadzono na zbiorze walidacyjnym (W) oraz zbiorze testowym (T). Dokładność klasyfikacji dla zbioru (W+T) dla modelu MLRA wyniosła 89%. Skuteczność diagnostyczną otrzymanego modelu potwierdza analiza krzywych ROC przedstawiona na rycinie 1.

Wartość pola pod krzywą ROC dla skonstruowanego modelu MLRA wynosiła 0,94 i różniła się istotnie od pola dla RMI ($p < 0,05$). Potwierdza to jej przydatność we wstępnym różnicowaniu guzów przydatkowych.

Nomogramy

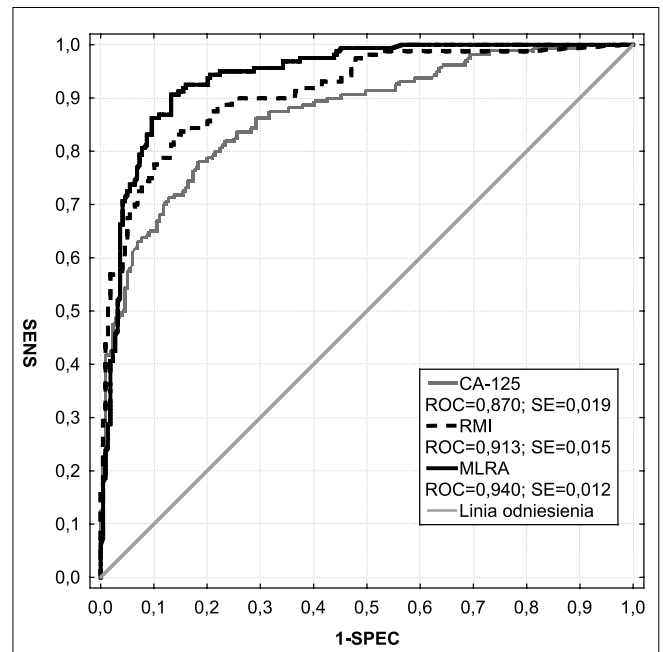
Na podstawie wyników analizy regresji logistycznej modelu MLRA, wykorzystując formułę opartą na wykładniczym przekształceniu współczynników regresji, skonstruowano nomogram, służący do obliczania prawdopodobieństwa istnienia procesu złośliwego. Każdemu z uwzględnionych parametrów w równaniu regresji logistycznej odpowiada określona liczba punktów, którą należy odczytać na górnej skali. Suma punktów, naniesiona na skalę w dolnej części nomogramu pozwala na obliczenie wartości predyktora liniowego w równaniu regresji, którego wartość po przekształceniu logitowym umożliwia obliczenie prawdopodobieństwa złośliwości guza. Praktyczne narzędzie, które można wykorzystać przy łóżku chorej przedstawia rycina 2.

Dyskusja

Badanie ultrasonograficzne i ocena stężenia CA-125 w surowicy krwi są dobrze udokumentowanymi wskaźnikami w przewidywaniu charakteru guzów przydatkowych. Jednakże nawet wśród tych wyselekcjonowanych kobiet z guzami przynajmniej 18% nowotworów złośliwych może nie zostać wykrytych przy pomocy tych metod [13].

W pierwszych próbach podwyższenia czułości i specyficzności testu klinicznego klasyfikującego guzy jajnika Jacobs i wsp., a następnie Tigulstad i wsp. zaproponowali indeksy złośliwości guza jajnika RMI (*Risk of Malignancy Index*) [3, 4].

W celu obliczenia ryzyka istnienia procesu złośliwego zalecono uwzględnić status menopauzalny, stężenie CA-125 w surowicy krwi przed leczeniem operacyjnym i podejrzaną cechy ultrasonograficzne guza. Różnica w wymienionych indeksach polega na przypisaniu innych wartości „wagowych” dla cechy wyniku badania ultrasonograficznego i statusu menopauzalnego. Morgante i wsp. [1999] ocenili działanie tych wskaźników na grupie włoskich kobiet, które poddano operacji chirurgicznej z powodu guza przydatków. Skuteczność obu indeksów RMI okazała się lepsza niż skuteczność samego poziomu CA-125. Autorzy podali, iż współczynnik zaproponowany przez Tigulstad'a i wsp. był bardziej czuły niż współczynnik Jacobsa i wsp., a jego wartość



Rycina 1. Krzywe ROC dla stężenia antygenu CA-125 i wskaźnika złośliwości guza RMI oraz dla modelu MLRA.

korelowała ze stanem klinicznym i wynosiła 80%. Nie przedstawiono jednak testów na istotność statystyczną, co jest niepokojące, zwłaszcza, że podana różnica była bardzo niewielka. Trzy badania kohortowe porównujące wartości tych dwóch indeksów wykazały, że wyliczona wartość powyżej 200 wskazuje na złośliwy charakter zmiany [4, 14, 15].

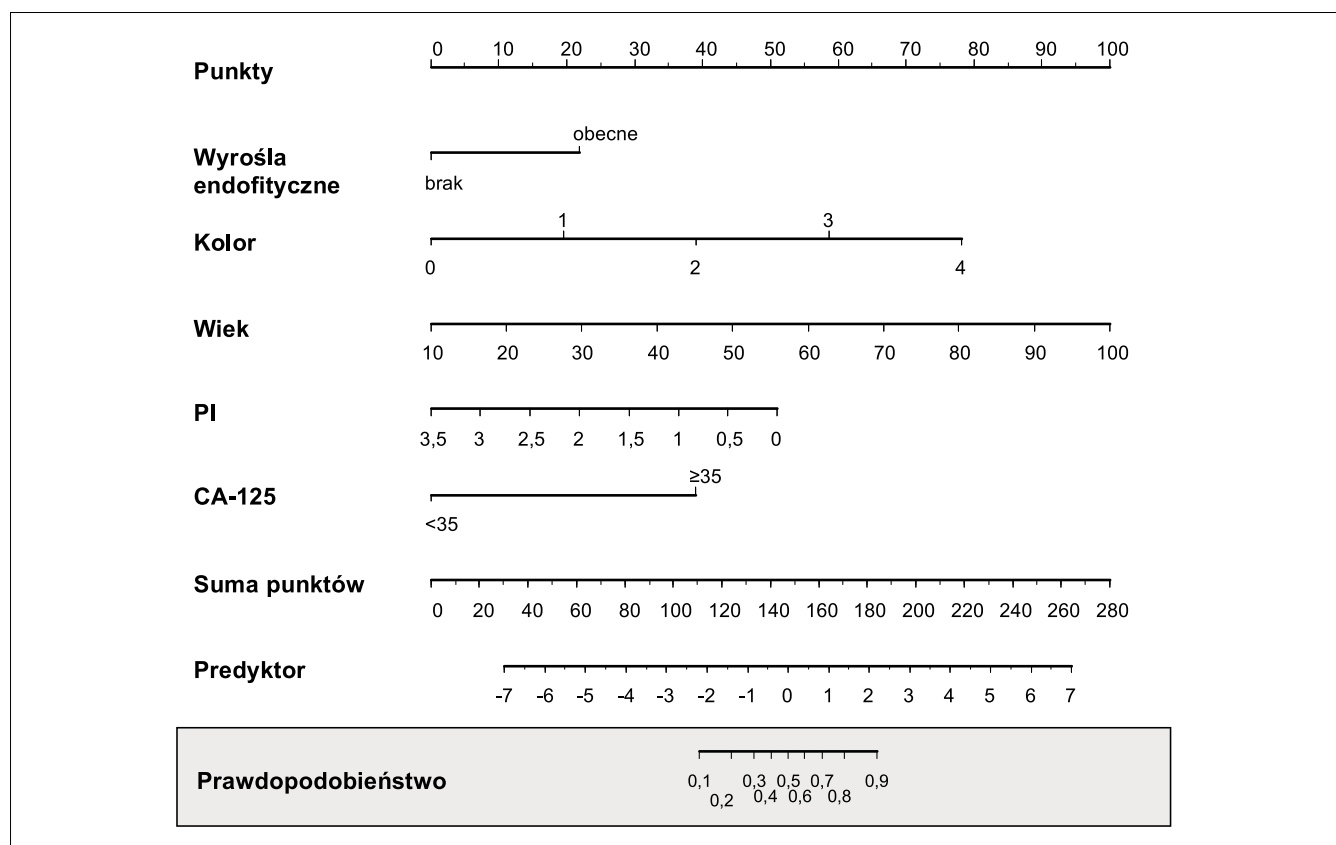
W 2007 roku Ulusoy i wsp. opublikowali wyniki analizy własnego indeksu RMI i różnych modyfikacji indeksu złośliwości guza przy granicy odcięcia 200 [16]. Czułość zawierała się w przedziale od 58% do 90%, specyficzność od 80,5% do 96,9%. Uważa się, że takie indeksy poprawiają przedoperacyjną ocenę guza przydatków, brak jest im jednak akceptowanej dokładności diagnostycznej [17, 18].

W badanej grupie 412 kobiet obliczono również wartości RMI poprzez połączenie wyniku badania ultrasonograficznego, statusu menopauzalnego i CA-125 wg Jacobsa i wsp. oraz Tingulstada i wsp. [3,4] W obu przypadkach uzyskano wartość pola powierzchni pod krzywą ROC wynoszącą 0,91. Przy wartości odcięcia równej 200 punktów uzyskano dokładność diagnostyczną wynoszącą 84%, czułość 81% i specyficzność 86%, co potwierdza, że jest to wynik niezadowolający. Porównanie najwyższych wartości pól pod krzywymi ROC: indeksu RMI i stężenia antygenu CA-125 wykazało nieznaczną poprawę dokładności diagnostycznej ($p=0,08$).

Innym sposobem analizy danych klinicznych i ultrasonograficznych są modele statystyczne tworzone w oparciu o analizę regresji logistycznej [19, 20, 21, 22, 23].

Wartości predykcyjne szeroko stosowanych metod wczesnego rozpoznawania raka jajnika są niezadowolające, ponieważ nie dają żadnej informacji dotyczącej aktualnego prawdopodobieństwa istnienia guza złośliwego u każdej pacjentki. Ryzyko istnienia nowotworu złośliwego obliczane jest w oparciu o szereg wskaźników uzyskanych na podstawie cech klinicznych i sonograficznych i stężenia antygenu CA-125.

Szacowanie prawdopodobieństwa złośliwości guza na podstawie kombinacji testów..



Rycina 2. Nomogram dla modelu MLRA do szacowania prawdopodobieństwa istnienia procesu złośliwego w badanym guzie jajnika.

W przypadku analizy logistycznej przewiduje się na podstawie predyktorów ilościowych lub jakościowych zmienną ciągłą, która zawiera się w przedziale od 0 do 1. Podczas gdy czułość i specyficzność nie podają wprost tej informacji, wskaźnik prawdopodobieństwa określa o ile dany wynik testu zwiększy lub obniży prawdopodobieństwo zaistnienia zdarzenia, które ma przewidzieć ten test. Obliczenie prawdopodobieństwa istnienia nowotworu złośliwego może mieć potencjalnie dużą wartość kliniczną, gdyż pozwala zaplanować najlepszą metodę leczenia dla każdej pacjentki z wykrytym guzem jajnika [20].

Najbardziej znaczącym badaniem dotyczącym przedoperacyjnej diagnostyki raka jajnika jest badanie IOTA (Międzynarodowa Analiza Guzów Jajnika). Protokół badania IOTA opisano szczegółowo w pracy [24].

Wyniki opublikowane na dużej grupie 1066 guzów, w tym 266 złośliwych (25%) mogą być porównywane z wynikami uzyskanymi w przedstawionej pracy. Spośród ponad 50 parametrów wyselekcjonowano 12 zmiennych, obejmujących parametry tj. wiek pacjentki, maksymalny wymiar jajnika, maksymalny wymiar elementów litych, obciążony wywiad rodzinny, stosowana terapia hormonalna, nasilony przepływ krwi w guzie (cecha „Kolor”), obecność bólu podczas badania, wodobrzusze, występowanie przegród i elementów litych w guzie, przepływ krwi w obrębie wyrośli endofitycznych, nieregularna budowa wewnętrznej ściany torbieli, cienie akustyczne [25].

Analizowane parametry, które zostały uznane w projekcji IOTA za najważniejsze w diagnostyce różnicowej uwzględniono w prezentowanej pracy.

Oceniono je zgodnie z zaleceniami tej grupy badaczy i dodatkowo rozszerzono badanie o parametry trójwymiarowej ultrasonografii. Powodem były poprzednie badania przeprowadzone w badanej w naszym ośrodku grupie kobiet z guzami jajnika, w których wykorzystano ultrasonografię dwu- i trójwymiarową i modelowanie statystyczne danych [21, 22, 26].

Wyniki te wskazywały, że korzystne może okazać się w przyszłości dołączenie do modelu predykcyjnego innych parametrów np. ultrasonografii trójwymiarowej [21, 22].

Kolejne wstępne wyniki z użyciem ultrasonografii 3D zasugerowały, że model prognostyczny oparty na ocenie danych ultrasonograficznych i pomiarach indeksu przepływu FI w trójwymiarowej sonoangiografii umożliwia obliczenie indywidualnego prawdopodobieństwa istnienia nowotworu złośliwego w badanym guzie przydatkowym [26, 27].

Przy założonym progowym prawdopodobieństwie istnienia guza złośliwego większym niż 50% najwyższa obliczona czułość i specyficzność dla modelu regresji logistycznej wynosiły odpowiednio 80% i 98%. W analizowanej grupie kobiet skonstruowany model predykcyjny pozwalał na uzyskanie istotnie wyższej czułości i specyficzności testu niż stosowane pojedynczo inne metody diagnostyczne.

Wykorzystanie regresji logistycznej w ocenie przydatności szeregu często używanych parametrów klinicznych i ultrasonograficznych w przedoperacyjnej diagnostyce guzów przydatkowych u kobiet było do tej pory przedmiotem wstępnych doniesień [19, 20, 21, 22, 23, 28, 29, 30, 31].

Pola pod krzywymi ROC dla modeli regresji logistycznej zawierały się w przedziale 0,74-0,98. Wyniki części z tych prac muszą być oceniane ostrożnie, gdyż liczba chorych i szczególnie nowotworów złośliwych była relatywnie niewielka. Opracowania które dotyczyły większych grup nie uwzględniały natomiast wyników ultrasonografii trójwymiarowej. Wnioskiem z przedstawionych badań jest fakt, że uwzględnienie kombinacji cech klinicznych, ultrasonograficznych i stężenia antygenu CA-125 w modelach regresji logistycznej pozwala na uzyskanie wyższej dokładności diagnostycznej niż pojedyncze parametry oraz proponowane proste indeksy morfologiczne.

W przedstawionej pracy oceniano skonstruowane modele matematyczne przy wykorzystaniu zbioru uczącego (355 przypadków) i zbioru walidacyjnego (24 przypadki), a zweryfikowano prospektywnie stosując niezależne dane 33 kolejnych nowych pacjentek. Wynik histopatologiczny tego niezależnego zbioru testowego, był dostępny dopiero po uzyskaniu wyników skonstruowanych modeli (zebranych prospektywnie).

W badaniu IOTA w jednej z prac dotyczącej regresji logistycznej w różnicowaniu guzów przydatkowych przedstawiono wyniki skonstruowanych dwóch modeli dla podgrupy 236 przypadków z uwzględnieniem stężenia antygenu CA-125: jeden oparty na 12 parametrach (M1) i drugi uproszczony (M2) oparty na 6 najważniejszych parametrach (wiek pacjentki, maksymalny wymiar elementów litych, wodobrzusze, nasilony przepływ krwi w obrębie wyrosła endofitycznych i elementów litych, nieregularna budowa wewnętrznej ściany torbieli, cienie akustyczne) [25]. Uzyskano dla modelu M1 pole pod krzywą ROC wynoszące 0,94, a dla modelu M2 – 0,92. Modele te wykazywały wyższe wartości niż przetestowany wskaźnik RMI i zaproponowane wcześniej modele zaproponowane przez Tailora i wsp. oraz Timmermana i wsp. [19, 20]. Uzyskane własne wyniki w przedstawionej pracy były porównywalne, pole pod krzywą ROC wynosiło 0,94 jak dla modelu M1.

W 2007 roku Holsbeke i wsp. przedstawili wyniki zewnętrznej walidacji modeli skonstruowanych w oparciu o bazę IOTA w celu różnicowania złośliwych i niezłośliwych guzów przydatkowych [32]. Dla indeksu Jacobsoa uzyskano pole pod krzywą ROC – RMI, które wynosiło 0,88. Odpowiadający tym indeksom model regresji logistycznej wykazał pole równe 0,76 i 0,9. Dla indeksu RMI uzyskano wyniki podobne do regresji logistycznej.

Wniosek

Zastosowanie kombinacji parametrów klinicznych ultrasonograficznych i stężenia antygenu CA-125 w modelu prognostycznym pozwala na wspomaganie procesów decyzyjnych we wstępnym różnicowaniu guzów jajnika.

Praca zgłoszona na XXX Jubileuszowy Kongres Polskiego Towarzystwa Ginekologicznego „Jakość życia kobiety - Salus feminae suprema lex esto” - w dniach 16-19 września 2009 roku w Lublinie.

Piśmiennictwo

1. Kurjak A, Kupesic S, Anic T, [et al.]. Three-dimensional ultrasound and power doppler improve the diagnosis of ovarian lesions. *Gynecol Oncol.* 2000, 76, 28-32.
2. Bordes K, Bory A, Benchaib M, [et al.]. Reproducibility of transvaginal three-dimensional endometrial volume measurements with virtual organ computer-aided analysis (VOCAL) during ovarian stimulation. *Ultrasound Obstet Gynecol.* 2002, 19, 76-80.

3. Jacobs I, Oram D, Fairbanks J, [et al.]. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol.* 1990, 97, 922-929.
4. Tingulstad S, Hagen B, Skjeldestad F, [et al.]. Evaluation of a risk of malignancy index based on serum CA-125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. *Br J Obstet Gynaecol.* 1996, 103, 826-831.
5. Jaeschke R, Cook D, Guyatt G. Ocena artykułów na temat testów diagnostycznych. Cz. II - metody określania przydatności testu. *Medycyna Praktyczna.* 1998, 11, 184-191.
6. Hanley J, McNeil B. The meaning and use of the area under the receiver operating characteristic (ROC) curve. *Radiology.* 1982, 143, 29-36.
7. Hanley J, McNeil B. A method of comparing of the areas under the receiver operating characteristic curve derived from the same cases. *Radiology.* 1983, 148, 839-843.
8. Zweig M, Campbell G. Receiver-operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993, 39, 561-577.
9. Panzer R, Black E, Griner P. Diagnostic strategies for common medical problems. Philadelphia, PA: *American College of Physicians.* 1991.
10. Sackett D, Rennie D. The science and art of the clinical examination. *JAMA.* 1992, 267, 2650-2652.
11. Stanisław A. Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom II. Modele liniowe i nieliniowe. Kraków: *Wyd. StatSoft.* 2007.
12. Lubsen J, Pool J, van der Does E. A practical device for the application of diagnostic or prognostic function. *Methods Inf Med.* 1978, 17, 127-129.
13. Spaczynski M. Onkologia ginekologiczna. Wrocław: *Wydawnictwo Medyczne Urban & Partner.* 1997, 156.
14. Morgante G, la Marca A, Ditto A, [et al.]. Comparison of two malignancy risk indices based on serum CA125, ultrasound score and menopausal status in the diagnosis of ovarian masses. *Br J Obstet Gynaecol.* 1999, 106, 524-527.
15. Asiam N, Tailor A, Lawton F, [et al.]. Prospective evaluation of three different models of pre-operative diagnosis of ovarian cancer. *Br J Obstet Gynaecol.* 2000, 107, 1347-1353.
16. Ulusoy S, Akbayir O, Numanoglu C, [et al.]. The risk of malignancy index in discrimination of adnexal masses. *Int J Gynaecol Obstet.* 2007, 96, 186-191.
17. Mol B, Boll D, DeKanter M, [et al.]. Distinguishing the benign and malignant adnexal mass: an external validation of prognostic models. *Gynecol Oncol.* 2001, 80, 162-167.
18. Geomini P, Bremer G, Kruitwagen R, [et al.]. Diagnostic accuracy of frozen section diagnosis of the adnexal mass: a metaanalysis. *Gynecol Oncol.* 2005, 96, 1-9.
19. Tailor A, Jurkovic D, Bourne T, [et al.]. Sonographic prediction of malignancy in adnexal masses using multivariate logistic regression analysis. *Ultrasound Obstet Gynecol.* 1997, 10, 41-47.
20. Timmerman D, Bourne T, Tailor A, [et al.]. A comparison of methods for preoperative discrimination between malignant and benign adnexal masses: The development of a new logistic regression model. *Am J Obstet Gynecol.* 1999, 181, 57-65.
21. Smoleń A, Czekierdowski A, Daniłó J, [i wsp.]. Sonoangiografia i zastosowanie logistycznej analizy regresji w przedoperacyjnej charakterystyce guzów jajnika. *Ginekol Pol.* 2002, 73, 1053-1060.
22. Smoleń A, Szperek D, Czekierdowski A, [i wsp.]. Ultrasonografia z kolorowym Dopplerem w charakterystyce guzów jajnika-porównanie modeli predykcyjnych z dwóch ośrodków. *Ginekol Pol.* 2003, 74, 863-871.
23. Szperek D, Moszynski R, Smolen A, [et al.]. Using logistic regression analysis in preliminary differential diagnosis of adnexal masses. *Int J Gynecol Cancer.* 2005, 15, 817-823.
24. Timmerman D, Valentin L, Bourne T, [et al.]. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) group. *Ultrasound Obstet Gynecol.* 2000, 16, 500-505.
25. Timmerman D, Tesat A, Bourne T, [et al.]. Logistic regression model to distinguish between the benign and malignant adnexal mass before the surgery: a multicenter study by the international ovarian tumor analysis group. *J Clin Oncol.* 2005, 23, 8794-8801.
26. Czekierdowski A. Zastosowanie ultrasonografii trójwymiarowej w różnicowaniu nowotworów jajnika. W: *Ultrasonografia dopplerowska w położnictwie i ginekologii.* Red. Brązert J. Poznań: *OWN.* 2004, 185.
27. Czekierdowski A, Stachowicz N, Smoleń A, [i wsp.]. Zastosowanie ultrasonografii dopplerowskiej i sonoangiografii trójwymiarowej w badaniu unaczynienia i przepływu krwi w guzach przydatków macicy u kobiet. *Ginekol Pol.* 2006, 77, 296-306.
28. Yamashita Y, Hatanaka Y, Torashima M, [et al.]. Characterization of sonographically indeterminate ovarian tumors with MR imaging. A logistic regression analysis. *Acta Radiol.* 1997, 38, 572-577.
29. Alcázar J, Jurado M. Using logistic model to predict malignancy of adnexal masses based on menopausal status, ultrasound morphology and color Doppler findings. *Gynecol Oncol.* 1998, 69, 146-150.
30. Czekierdowski A. Badania nad angiogenezą w łagodnych i złośliwych guzach jajnika przy zastosowaniu ultrasonografii z kolorowym Dopplerem oraz oceny stężeń śródbłonkowego czynnika wzrostu i wybranych markerów nowotworowych CA 125, CA72.4 i CA19.9. *Rozprawa habilitacyjna.* Lublin. Folium. 2000.
31. Alcázar J, Errasti T, Laparte C, [et al.]. Assessment of a new logistic model in the preoperative evaluation of adnexal masses. *J Ultrasound Med.* 2001, 20, 841-848.
32. Van Holsbeke C, Van Calster B, Valentin L, [et al.]. External Validation of Mathematical Models to Distinguish Between Benign and Malignant Adnexal Tumors: A Multicenter Study by the International Ovarian Tumor Analysis Group. *Clin Cancer Res.* 2007, 13, 4440-4447.