



PRZEBIEG ZAPALENIA AUTOIMMUNOLOGICZNEGO W CZASIE 12-MIESIECZNEJ OBSERWACJI NA PRZYKŁADZIE OFTALMOPATII GRAVESA

Agata Baldys-Waligórska, Uniwersytet Jagielloński Collegium Medicum

*Andrzej Sokołowski, Centrum Onkologii – Oddział w Krakowie,
Uniwersytet Ekonomiczny w Krakowie, StatSoft Polska*

Oftalmopatia naciekowa (GO) jest jednym z objawów choroby Graves-Basedowa (MGB), sprawiającym największe trudności terapeutyczne. Podobnie jak nadczynność tarczycy w MGB, tak i oftalmopatia jest chorobą o podłożu autoimmunologicznym z pewną nie do końca jeszcze wyjaśnioną predyspozycją genetyczną. Głównym antygenem wydaje się być receptor TSH, przeciwko któremu uczulone komórki immunokompetentne produkują przeciwciała (TRAb). Stężenie TRAb jest miarą nasilenia procesu zapalnego w oczodole i w praktyce klinicznej koreluje ze stopniem aktywności choroby CAS (*Clinical Activity Score*, maximum 7 punktów) i stopniem zaawansowania choroby NOSPECS, wyrażanym jako indeks oftalmopatii (IO, maksimum 15 punktów). Leczenie nadczynności tarczycy radiojodem ¹³¹I prowadzi u około 5% pacjentów do zaostrzenia istniejącej łagodnej GO lub powstania GO de novo.

Celem pracy jest:

- ♦ porównanie przebiegu oftalmopatii (ciężkości GO – CAS, NOSPECS) oraz markerów zapalenia (TRAb, cytokiny) pomiędzy grupą chorych z nadczynnością tarczycy leczonych doustnymi lekami tyreostatycznymi (ATD) a grupą leczoną radiojodem ¹³¹I,
- ♦ określenie czynników, które wpływają na stopień zaawansowania GO,
- ♦ pomiar zależności indeksu oftalmopatii IO od stężenia TRAb,
- ♦ określenie częstości występowania oporności na leczenie glikokortykoidami (non-responders) w grupie pacjentów z oftalmopatią w przebiegu choroby Graves-Basedowa w badaniu prospektywnym oraz określenie czynników predykcyjnych.

W latach 2000-2008 do Kliniki Endokrynologii UJCM zgłaszało się około 50 pacjentów rocznie z powodu zaostrzenia objawów oftalmopatii w przebiegu choroby Graves-Basedowa (GO). 168 z nich wyraziło zgodę na udział w prospektywnym badaniu obserwacyjnym z ankietą, polegającym na 12-miesięcznym monitorowaniu po zakończonym leczeniu z powodu GO. Grupa składała się ze 119 kobiet (70,8% pacjentów) i 49 mężczyzn (29,2% pacjentów), średnia wieku 52,2± 11,2 lat. Nadczynność tarczycy rozpoznano średnio



31,5±52,4 miesięcy przed zgłoszeniem się pacjentów do Kliniki z powodu objawów oftalmopatii, które średnio trwały 12,5±20,3 miesięcy.

W latach 2004-2007 w Klinice Endokrynologii UJCM leczono radiojodem 131-I 1200 pacjentów z powodu nadczynności tarczycy, w tym 696 (58%) pacjentów z chorobą Graves-Basedowa. Leczenie radiojodem proponowano tylko pacjentom z indeksem oftalmopatii IO <3 punkty i z indeksem aktywności zapalenia (*Clinical Activity Score*) CAS <3 punkty. Stężenie TSH w surowicy krwi przed leczeniem 131-I wynosiło w tej grupie chorych 0,24±0,58 $\mu\text{U/ml}$, średni 24 hrs wychwyty 131-I nad tarczycą wynosił 54,0±14,6%, średnia aktywność podana 496±141 MBq.

46 pacjentów, 37 kobiet (80,4%) i 9 (19,6%) mężczyzn, średnia wieku 52,1±13,3 lat w kolejnych miesiącach po leczeniu, średnio 11,9±17,8 miesięcy, zgłosiło się do Ambulatorium z powodu pojawienia się de novo lub zaostrzenia objawów oftalmopatii. Nadczynność tarczycy w tej grupie pacjentów rozpoznano średnio 72,3±67,8 miesięcy przed zgłoszeniem się pacjentów do Kliniki w celu leczenia radiojodem. U 76%, którzy zgłosili się z objawami GO, stwierdzono niedoczynność tarczycy, średnie stężenie TSH wynosiło 23,9±24,5 $\mu\text{U/ml}$ (mediana 15,4; max=91,7; min=1,7 $\mu\text{U/ml}$). Pacjentów leczono odpowiednią dawką L-tyroksyny przez cały czas obserwacji w celu przywrócenia i utrzymania eutyreozy.

Pacjenci obu grup zostali zakwalifikowani do leczenia glikokortykoidami metodą pulsacyjną (2,0 g SoluMedrolu i.v. na tydzień do łącznej dawki 8,0g) oraz radioterapii oczodołów (20 Gy w 10 frakcjach) w Zakładzie Radioterapii Szpitala Dziecięcego w Krakowie-Prokocimiu. Po zakończonym leczeniu pacjenci zgłaszali się do kontroli i oceny wyników leczenia po 1, 6, i 12 miesiącach w Poradni Endokrynologicznej i Okulistycznej.

Ponieważ rozkłady wielu badanych cech zdecydowanie odbiegały od rozkładu normalnego, zatem zachodziła potrzeba zastosowania nieparametrycznych metod statystycznych. Porównania grup mają dwojaki charakter. Po pierwsze można w jednej grupie chorych śledzić zmiany konkretnego parametru w czasie, w związku z pomiarami dokonywanymi w miesiącach 0, 1, 6 oraz 12. Stosowane testy muszą uwzględniać fakt, że takie obserwacje są obserwacjami zależnymi. Zazwyczaj lekarze nie zadowolają się wynikiem testowania głoszącym istotność różnic pomiarów wykonywanych w różnych momentach czasu. Chcą odpowiedzi na pytanie, w których interwałach czasowych te zmiany były istotne. Wymaga to stosowania porównań parami, co z kolei stawia problem testowania wielokrotnego. Relatywnie prostszym zagadnieniem jest porównywanie dwóch rozłącznych grup w zadanym momencie czasu.

Statystyczna charakterystyka materiału

Informacje statystyczne o pacjentach zostały zgromadzone w dwóch bazach danych. W tej pracy korzystamy głównie z bazy, która zawierała podstawowe informacje o parametrach chorych z grupy, która w skrócie będzie określana 131-I oraz grupy ATD.



Dane: AW 110126 (149 zm. * 214 prz.)

	1 Grupa	4 płeć	5 wiek	6 nadczymność rozp. msc	7 ototalgalia rozp. msc	8 otalmop. po 131-I msc	9 131-I ile razy leczona	10 TSH 0'	11 fT4 0'	12 TRAb 0'	13 TSH 1msc	14 fT4 1msc	15 TRAb 1msc	16 TSH 6 msc	17 fT4 6 msc	18 hTRAb 6 msc	19 TSH 12 msc	20 fT4 12 msc	21 hTRAb 12 msc	
1	131-I	k	33	120	72	60	1	3,20		1,0	1,74		0,4	1,80		0,3	0,80		2,2	
2	131-I	k	75	36	6	4	3	2,00		8,1	5,54		12,70	5,50						
3	131-I	k	55	48	3	12	1	1,41	14,5	7,9	0,87	23,3		0,06		15,5	0,32	22,1	14,4	
4	131-I	k	51	0	24	17	1	0,14		19,0	0,16	22,0								
5	131-I	k	54	48	3	24	2	0,90	24,0		0,90	20,0								
6	131-I	k	44	18	4	14	1	0,30	21,0	90,0	0,90			45,0	0,30	16,0	42,0	0,20	21,8	35,4
7	131-I	k	56	24	6	18	1	0,15	21,0	15,4	1,00			0,15	28,0		1,3	1,90	22,0	1,9
8	131-I	k	31	48	3	1	1	0,24	21,0	45,0				9,5		3,5	4,00	18,0	0,2	
9	131-I	k	64	48	5	6	1	0,05	15,0	23,0				24,6						
10	131-I	k	38	120	60	24	1	1,98	24,1	11,0	0,05	36,9	14,0	0,37		14,0				
11	131-I	k	60	24	6	6	1	6,50	18,0	20,0	4,00	17,0	2,5	3,00	24,0	2,0	3,20	20,0	1,7	
12	131-I	k	48	60	6	5	1	0,03	25,6	47,9	9,03	0,8	8,5	0,89	16,0	7,3	22,00		5,4	
13	131-I	k	56	156	84	9	1	2,10	16,5	1,4	0,85	18,5	1,4				0,94			
14	131-I	k	25	24	2	1	1	24,90	3,1	12,0	0,09	21,3	6,5	7,20		4,2	13,00		30,8	
15	131-I	k	62	12	3	6	1	0,50	20,0	48,7	2,11		23,9							
16	131-I	k	24	60	7	5	2	91,70	6,0	11,1	0,15	40,0	6,1							
17	131-I	k	63	180	3	81	1	0,80	15,3	70,0	0,29	16,2	63,2	0,77	16,2	66,8	1,60		24,7	
18	131-I	k	50	9	2	1	1	1,10	18,8	50,0	0,57	18,9	80,0	0,60	20,8	96,0	8,10		14,0	
19	131-I	k	44	20	5	1	1	0,09	9,5	2,1			0,05	14,0	1,9			4,5		
20	131-I	k	51	96	6	18	1	0,60	18,9	9,3	1,50	18,2	7,3		4,2	2,60		2,8		
21	131-I	m	65	60	9	15	1	0,70			0,11		14,0	1,10		15,9		12,9		
22	131-I	m	66	120	18	78	1	0,51		14,3	4,47	20,0	7,7	6,00		11,5				
23	131-I	k	75	120	4	14	2	0,20		33,5	3,7	0,40	7,4	2,60				2,86		2,8
24	131-I	k	63	180	6	114	1	1,40	18,7	2,9	1,70		2,5	1,70		3,7	1,10	2,0		
25	131-I	k	57	72	6	6	1	2,07		32,0	0,08		24,1					2,25		11,7

Baza zawiera ponad 140 zmiennych. Część z nich to oryginalne wartości uzyskane w trakcie badań, a część to zmienne zero-jedynkowe oznaczające przekroczenie pewnych wartości granicznych. Te zmienne zostały już utworzone w programie STATISTICA. Typową sytuacją w bazach danych jest niekompletność informacji. Jest to w pewnym sensie stan naturalny i nie stosuje się tu procedury tzw. imputacji danych, popularnej np. w ekonomii.

Obliczenia dotyczące ogólnej charakterystyki materiału wykonujemy w module *Statystyki podstawowe i tabele*, nie zapominając o skorzystaniu z klawisza *Select cases*, który pozwala wybrać grupę poddawaną właśnie analizie. Poniżej pokazujemy ustawienia na karcie *Statystyki opisowe* dla obliczenia charakterystyk zmiennej TRAb, w grupie 131-I, w czterech momentach czasu.

Statystyki opisowe: AW 110126

Zmienne: 12 15 18 21

Podsumowanie

W. prawdopodob. i rozrzu tu W. skategoryzowane Opcje

Podstawowe Więcej Odporne Normalność

Statystyki W1 W2 Oblicz statystyki:

Pokożenia, N

- N ważnych
- % ważnych
- Średnia
- Suma
- Mediana
- Moda
- Śr. geometr.
- Śr. harmon.

Zmienności, momenty

- Odchylenie standardowe
- PU dla odch. std.
- Przedział: 95,00 %
- Wsp. zmienności
- Wariancja
- Błąd stand. średniej
- Przedz. ufn. średniej
- Przedział: 95,00 %
- Skośność
- Błąd st. skośności
- Kurtozja
- Błąd st. kurtozy

Percentyle, zakresy

- Minimum i maksimum
- Dolny i górny kwartyl
- Zakres percentyli
- Pierwszy: 10,00 %
- Drugi: 90,00 %
- Rozstęp Rozstęp kwartyłowy

Wszystkie Od nowa Zapisz ust. jako domyślne

Anuluj Opcje Grupami...

SELECT CASES W W

Momenty ważone

DF = W-1 N-1

Usuwanie BD

- Przypadkami
- Parami

Wyniki pokazują żądane charakterystyki.

Zmienna	N ważnych	Średnia	Mediana	Minimum	Maksimum	Dolny Kwartyl.	Górny Kwartyl.	Odch. std
TRAb 0'	41	22,97	14,30	0,60	90,00	5,40	34,60	22,29
TRAb 1 msc	35	15,43	7,70	0,10	80,00	2,80	22,00	18,52
hTRAb 6 msc	32	14,06	4,65	0,10	96,00	2,70	13,40	21,49
hTRAb 12 msc	28	9,47	3,65	0,10	41,00	1,45	13,60	11,60

Zauważamy przede wszystkim duże różnice pomiędzy średnimi arytmetycznymi a medianami, co sugeruje mocną asymetrię prawostronną.

Porównanie poziomu cechy w różnych momentach czasu

W poprzedniej tabeli widzimy, że poziom TRAb w grupie 131-I zmniejszał się w czasie. Dobrze jest podjąć próbę sformułowania precyzyjnej odpowiedzi na pytanie, w których momentach czasu, przeciętny poziom badanej cechy różnił się istotnie. W tym celu trzeba zastosować właściwy test statystyczny. Zauważmy przede wszystkim, że obserwujemy w czasie tych samych pacjentów. Mamy więc do czynienia z tzw. próbami zależnymi. Niewielką komplikacją jest fakt, że nie wszyscy pacjenci mają komplet wyników z czterech momentów. Ogólna strategia, którą tu przyjmujemy, to porównywanie każdego okresu z każdym. Dzięki temu wykorzystamy maksimum informacji z prób zależnych. Przy każdym porównaniu brani pod uwagę są ci pacjenci, którzy mają wyniki w dwóch porównywanych momentach czasu. W ten sposób pojedynczy pacjent nie musi brać udziału we wszystkich porównaniach par okresów.

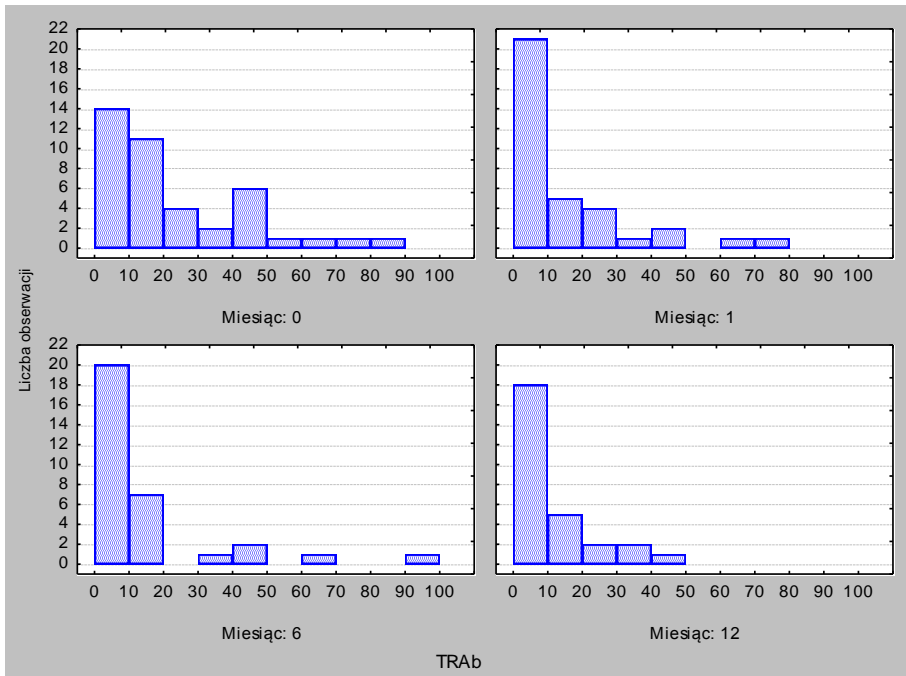
Przy wyborze właściwego testu statystycznego należy rozważyć rozkład badanej cechy. Duże różnice pomiędzy średnimi a medianami w grupie 131-I sugerują rozkłady prawostronnie asymetryczne. Sprawdzone to najpierw, tworząc wykresy histogramów rozkładu. Łatwe umieszczenie wszystkich czterech histogramów na jednym wykresie można uzyskać, tworząc *histogram skategoryzowany*. W tym celu przygotowano plik danych składający się z dwóch kolumn. Jedna zawiera wartości TRAb umieszczone w jednej kolumnie, najpierw te z miesiąca 0, potem z miesiąca 1 itd. Druga kolumna zawiera numer miesiąca.

Z menu *Wykresy* wybieramy *Wykresy skategoryzowane*, a potem *Histogramy*.

Jako *Kategorie X* można wybrać *Unikalne wartości* lub *Kody*. W tym drugim przypadku program zapyta nas o wartości kodów i wtedy trzeba wybrać *Wszystkie*. Na otrzymanym wykresie widać wyraźnie prawostronną asymetrię. Większość testów parametrycznych ma założenie normalności rozkładu badanej cechy. Testy te – szczególnie przy małych próbach – mają moc lepszą od testów nieparametrycznych, które nie wymagają spełnienia normalności rozkładu. Nie zawsze sytuacja na wykresie jest tak oczywista jak w przypadku naszej zmiennej TRAb. Formalnej kontroli normalności rozkładu możemy dokonać poprzez zas-

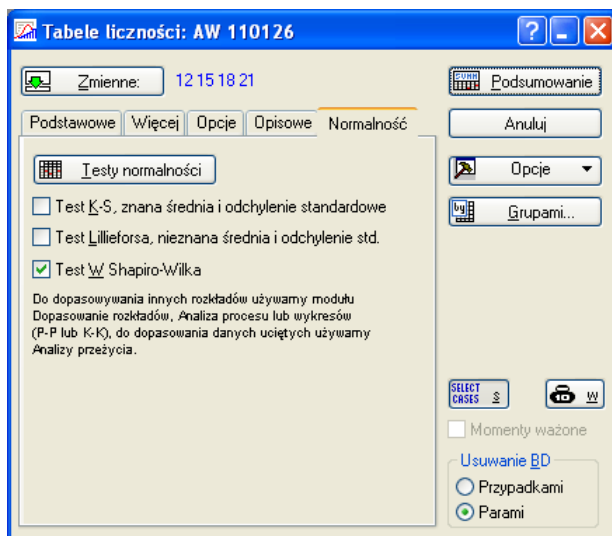


tosowanie testu Shapiro-Wilka. Jest to test powszechnie uważany za najlepszy test normalności i działa już od trzech przypadków.





Test normalności można znaleźć między innymi w module *Statystyki podstawowe i tabele*, na karcie *Tabele liczebności*.



Po naciśnięciu klawisza *Testy normalności* otrzymujemy tabelkę z wartościami statystyki testowej W , a przede wszystkim z wartościami p . Warto pamiętać, że test Shapiro-Wilka – dość nietypowo dla testów zgodności – ma lewostronny zbiór krytyczny. Oznacza to, że im większe odstępstwo od normalności tym mniejsza wartość statystyki testowej W .

Testy normalności (AW 110126)				
Warunek uwzględniania: $v1=131-1'$				
Zmienna	N	W	p	
TRAb 0'	41	0,857370	0,0001	
TRAb 1 msc	35	0,746705	0,0000	
hTRAb 6 msc	32	0,636829	0,0000	
hTRAb 12 msc	28	0,783468	0,0001	

Otrzymane wyniki potwierdzają oczywiście poprzednią obserwację o tym, że rozkłady TRAb nie są rozkładami normalnymi. W takiej sytuacji badacz ma trzy wyjścia:

- ◆ liczyć na odporność testu, co zazwyczaj zachodzi w przypadku dużych prób,
- ◆ próbować przekształcić zmienną oryginalną tak, aby uzyskać rozkład zbliżony do normalnego (najczęściej stosuje się tu logarytmowanie lub przekształcenie Boxa-Coxa),
- ◆ zastosować test nieparametryczny.

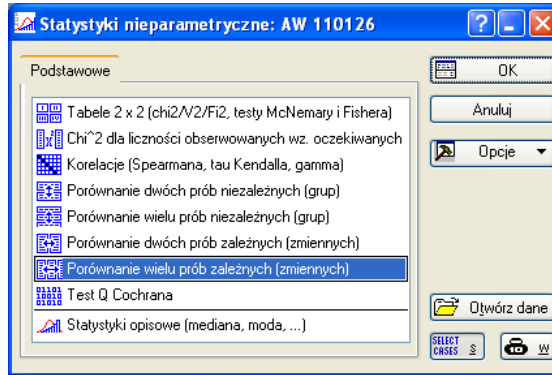
W naszym zagadnieniu liczebności nie są takie, aby gwarantowały odporność testów porównujących średnie. Logarytmowanie też nie przyniosło pożądanych skutków. Dlatego zdecydowano się na zastosowanie testów nieparametrycznych.



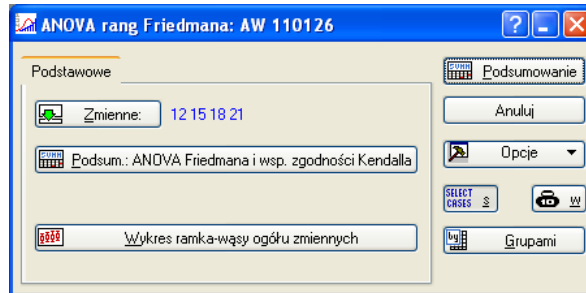
Najpierw zweryfikujemy hipotezę o tym, że przeciętny poziom TRAb w grupie 131-I nie zmieniał się w czasie:

$$H_0: \mu(\text{TRAb } 0) = \mu(\text{TRAb } 1) = \mu(\text{TRAb } 6) = \mu(\text{TRAb } 12)$$

Hipoteza alternatywna głosi, że przynajmniej w dwóch z tych czterech momentów przeciętny poziom TRAb różnił się istotnie. Do weryfikacji hipotezy użyjemy testu zaproponowanego przez sławnego ekonomistę Miltona Friedmana. Znajduje on się w module *Statystyki nieparametryczne*.



Jeżeli chcemy porównywać poziom tej samej cechy w różnych momentach czasu, to obserwacje dla każdego momentu powinny być umieszczone w innej zmiennej, a przyporządkowanie obiektów wynika oczywiście z tego, że wszystkie informacje w danym wierszu dotyczą tego samego pacjenta.

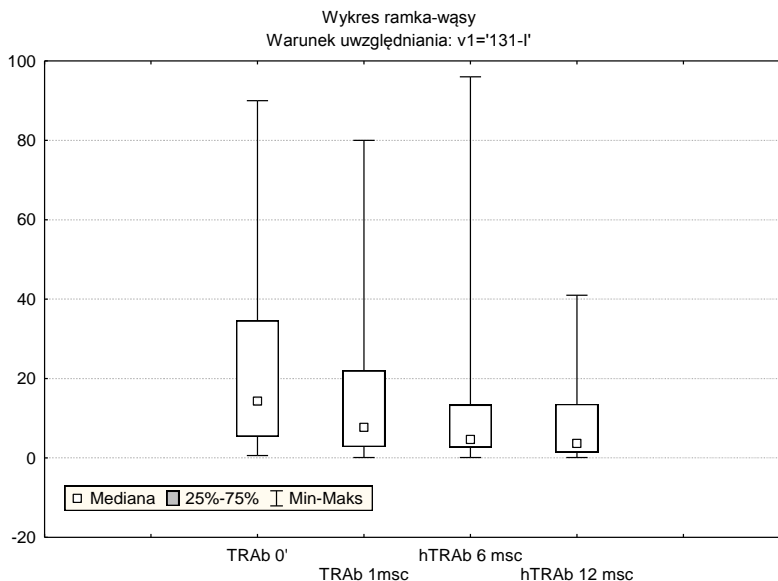
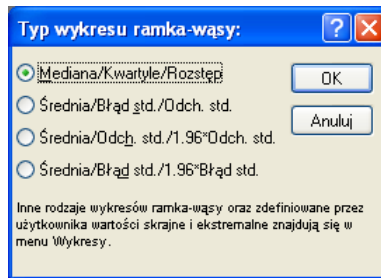


W wyniku otrzymujemy tabelę przedstawioną poniżej.

Dane: ANOVA Friedmana i współczynnik zgodności Kendalla (AW 11...)					
ANOVA Friedmana i współczynnik zgodności Kendalla (AW 110126)					
Chi kwad. ANOVA (N = 20 , df 3) =21,43077 p ,00009					
Współczynnik zgodności= ,35718 r śred. rang = ,32335					
Warunek uwzględniania: v1='131-I'					
Zmienna	Średnia Ranga	Suma Rang	Średnia	Odch. std	
TRAb 0'	3,550000	71,00000	29,14000	27,88600	
TRAb 1 msc	2,475000	49,50000	18,87000	23,12975	
hTRAb 6 msc	2,225000	44,50000	17,93500	26,30534	
hTRAb 12 msc	1,750000	35,00000	10,80000	13,11616	

Wartość p jest równa 0,0001 i ponieważ jest mniejsza od poziomu istotności $\alpha=0,05$, dlatego odrzucamy hipotezę o niezmienności poziomu TRAb w czasie. Zauważmy, że w grupie 131-I mamy tylko 20 pacjentów, którzy mają komplet wyników TRAb z czterech momentów badania. W panelu *ANOVA rang Friedmana* dostępny jest klawisz umożliwiający sporządzenie wykresu ramka-wąsy.

Z uwagi na asymetrię rozkładu TRAb rozsądne wydaje się wybranie opcji wykresu, wykorzystujących medianę i kwantyle.

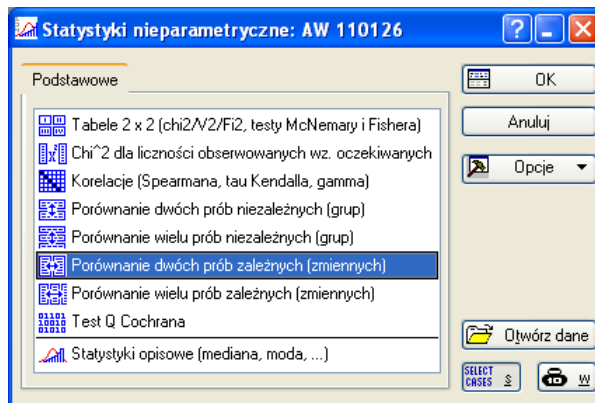


Na powyższym wykresie widać tendencje zmian poziomu TRAb, jednak nie można tu wysnuwać głębszej interpretacji, gdyż czas na osi pionowej nie jest traktowany jako zmienna ciągła, tylko jako porządkowa.

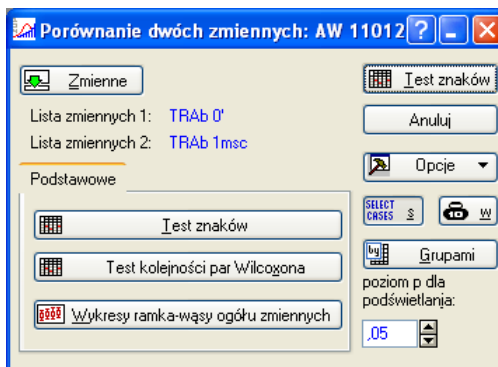
Przy odrzuceniu hipotezy zerowej w analizie wariacji niemal zawsze pojawia się następnie pytanie o to, które grupy (u nas – momenty badania) różnią się od siebie istotnie. Przy porównywaniu prób niezależnych program *STATISTICA* proponuje poważną baterię tzw. testów *post-hoc*, zarówno w wersji parametrycznej, jak i nieparametrycznej.



W tej analizie zastosowano porównywanie parami przy pomocy nieparametrycznego testu dla dwóch prób zależnych. Wybieramy *Statystyki nieparametryczne*.



Następnie wybieramy *Porównanie dwóch prób zależnych* i w oknie, które się ukáže, wskażemy dwie zmienne do porównania.



Stosujemy *Test kolejności par Wilcoxona*, mocniejszy niż dość prymitywny test znaków.

Test kolejności par Wilcoxona (AW 110126)				
Zaznaczone wyniki są istotne z $p < ,05000$				
Warunek uwzględniania: $v1='131-1'$				
Para zmiennych	N	T	Z	p
TRAb 0' & TRAb 1msc	33	1,66,00000	3,832649	0,000127

W porównaniu poziomym TRAb w pierwszym miesiącu z miesiącem zerowym brało udział 33 pacjentów. Różnica okazała się istotna statystycznie. Jeżeli porównamy wszystkie pary momentów czasowych pomiaru TRAb, to otrzymujemy następujące wyniki:

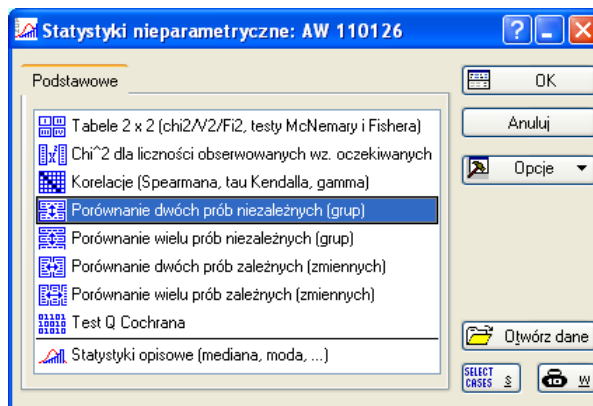


	Miesiąc 0	Miesiąc 1	Miesiąc 6	Miesiąc 12
Miesiąc 0	.	0,00013	0,00042	0,00013
Miesiąc 1		.	0,38352	0,01558
Miesiąc 6			.	0,08913
Miesiąc 12				.

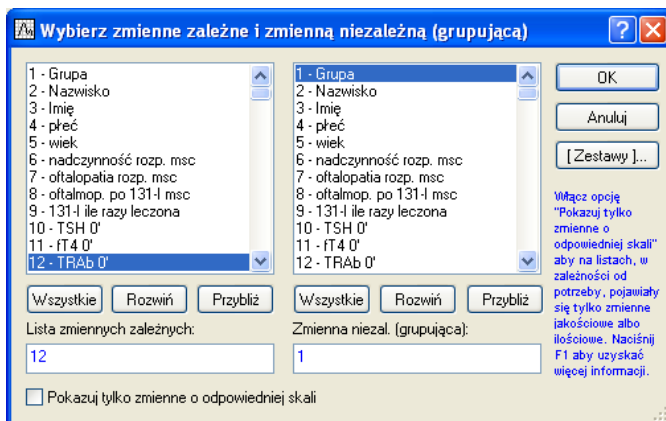
Wyniki przedstawione w powyższej tabeli pokazują, że test kolejności par Wilcozona zastosowaliśmy sześć razy. Porównywane były cztery zbiory danych – każdy z każdym. W takiej sytuacji mamy do czynienia z tzw. testowaniem wielokrotnym. W literaturze statystycznej zaproponowano wiele metod, które umożliwiają zadanie poziomu istotności nie dla pojedynczego testu, ale dla rodziny testowej. Stosunkowo popularna jest poprawka na testowanie wielokrotne Bonferroniego-Dunna. Polega ona na podzieleniu żądanego poziomu istotności przez liczbę testowań i w ten sposób otrzymuje się poziom istotności dla pojedynczego porównania i do niego trzeba odnosić wartość p. W naszym przypadku mamy $0,05/6=0,0083$. Po zastosowaniu poprawki na testowanie wielokrotne nieistotna okazuje się różnica w poziomie TRAb pomiędzy miesiącem pierwszym a dwunastym. Trzeba tu jednak zaznaczyć, że konieczność stosowania poprawki na testowanie wielokrotne nie jest powszechnie akceptowana przez statystyków.

Porównanie zmian w czasie, w grupach 131-I z ATD

Jako przykład rozpatrzmy znów poziom TRAb mierzony w czterech momentach. Tym razem jednak chodzi o porównanie poziomu TRAb w dwóch grupach – 131-I oraz ATD. Ponieważ oczywiście w tych grupach są różni pacjenci, mamy więc do czynienia z próbami niezależnymi. W związku z mocną asymetrią prawostronną badanej cechy wybieramy też test nieparametryczny.



Testujemy różnice w każdym momencie osobno. Trzeba skorzystać z pierwotnej bazy, w której wartość TRAb jest w jednej kolumnie, a oprócz tego mamy kolumnę identyfikującą przynależność do grupy.



Wybieramy *Test U Manna-Whitneya* jako nieparametryczny odpowiednik testu Studenta.

Dane: Test U Manna-Whitneya (AW 110126) (10 zm. * 1 prz.)

Test U Manna-Whitneya (AW 110126)
Względem zmiennej: Grupa
Zaznaczone wyniki są istotne z $p < ,05000$

Zmienna	Sum.rang 131-I	Sum.rang ATD	U	Z	p	Z popraw.	p	N ważn. 131-I	N ważn. ATD	2*1 str. dokł. p
TRAb 0'	4122,500	8757,500	1617,500	3,210849	0,001324	3,211013	0,001323	41	119	0,001171

Interpretacja wyników tego testu jest zasadniczo taka sama jak w przypadku testu t dla prób niezależnych. Różnica polega na tym, że obliczenia w teście U są wykonywane w oparciu o sumę rang, a nie o średnie. Test U jest najmocniejszą nieparametryczną alternatywą dla testu t dla prób niezależnych. W przypadku prób, których licznosc przekracza 20, rozkład z próby statystyki U szybko zmierza do rozkładu normalnego. Stąd też obok wartości statystyki U (skorygowanej ze względu na rangi wiązane) podawana jest wartość z (wartość zmiennej o rozkładzie normalnym) oraz odpowiadająca jej wartość p.



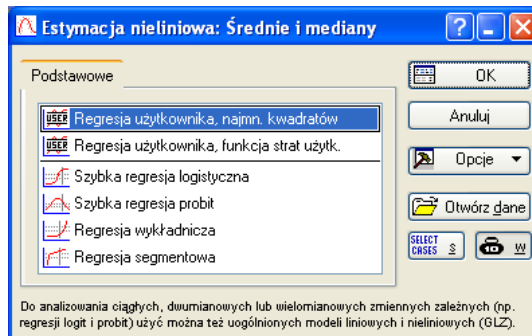
Z porównania TRAb w dwóch grupach w czterech momentach czasu otrzymujemy następujące wyniki.

Miesiąc	Liczba pacjentów 131-I/ATD	Średnia 131-I	Średnia ATD	Mediana 131-I	Mediana ATD	Wartość p z testu U
0	41/119	22,97	13,20	14,30	5,60	0,00132
1	35/90	15,43	5,56	7,70	2,20	0,00006
6	32/99	14,06	5,26	4,65	1,60	0,00088
12	28/79	9,47	4,75	3,65	1,40	0,01276

Jeżeli teraz chcemy zastosować poprawkę na testowanie wielokrotne, to poziom istotności dla pojedynczego testu wynosi $0,05/4=0,0125$. Zastosowanie tej poprawki nie zmienia więc merytorycznych wyników testowania, ponieważ wszystkie wartości p są mniejsze od tego skorygowanego poziomu istotności. Przeciętny poziom TRAb był więc w całym okresie obserwacji wyższy w grupie 131-I niż w ATD.

Interesującym zadaniem jest próba opisanie zmian w poziomie TRAb przy pomocy funkcji trendu. Zastosowanie oryginalnych obserwacji nie jest zbyt uzasadnione z powodu – wielokrotnie wspomianej tu – znacznej asymetrii prawostronnej. Nie udałoby się w żaden sposób obronić zasadności założenia normalności składnika losowego trendu. Ponadto klasyczna metoda najmniejszych kwadratów minimalizuje sumę kwadratów odległości wartości rzeczywistych od wartości teoretycznych trendu w ten sam sposób, w jaki średnia arytmetyczna minimalizuje odległość obserwacji od centrum w pojedynczej próbie. Przy wykorzystaniu wszystkich obserwacji oryginalnych metoda najmniejszych kwadratów forsowałaby trend podążający mniej więcej wzdłuż średnich arytmetycznych.

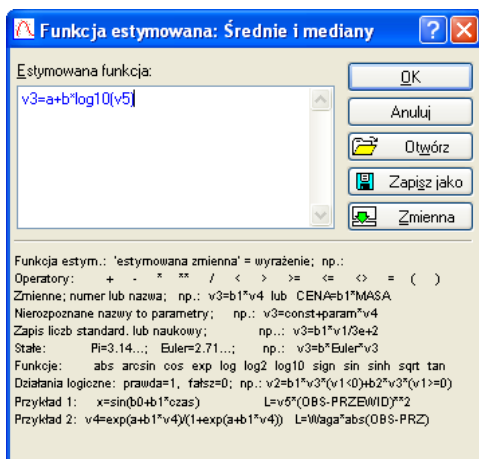
Zdecydowano się zatem zbudować trend w oparciu o mediany. Tych median jest tylko cztery w każdej grupie, ale za każdą z nich stoi pokaźna liczba pacjentów. Robocza tabela danych do trendu ważonego ma pięć wierszy odpowiadających kolejnym momentom czasu. Zmiany w medianach sugerują trend logarytmiczny. Szacujemy go w module *Zaawansowane modele liniowe i nieliniowe*, wybierając *Estymację nieliniową*.



Opcja *Regresja użytkownika* pozwala na szacowanie praktycznie dowolnych funkcji. Nasza funkcja trendu ma ogólną postać:

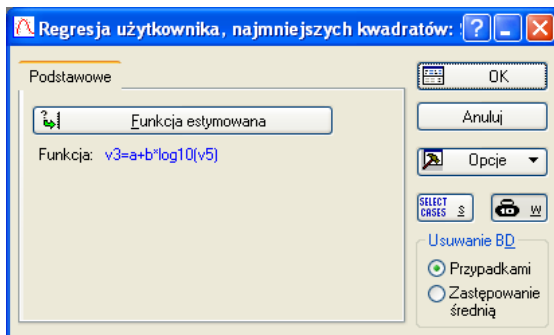


$$\hat{y}_t = a + b \log t$$

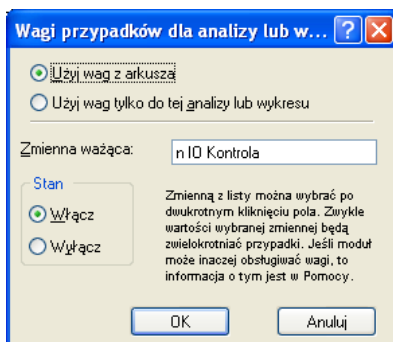


W naszym pliku danych w zmiennej v3 podano wartości median TRAb w grupie 131-I, a w zmiennej v5 – numer miesiąca. Ponieważ pierwszy moment miał numer zero, więc dla umożliwienia wykorzystania logarytmowania zamieniono go na 0,01.

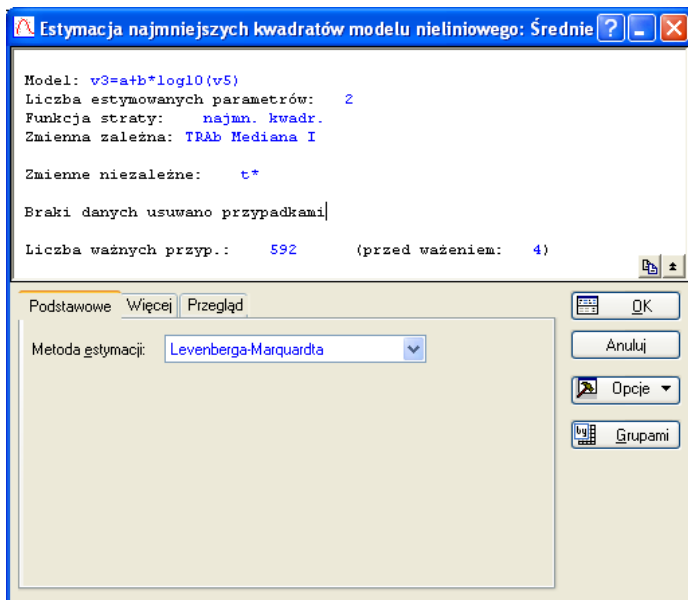
Po naciśnięciu klawisza OK ukazuje się panel, na którym trzeba wybrać możliwość uwzględnienia ważenia cech.



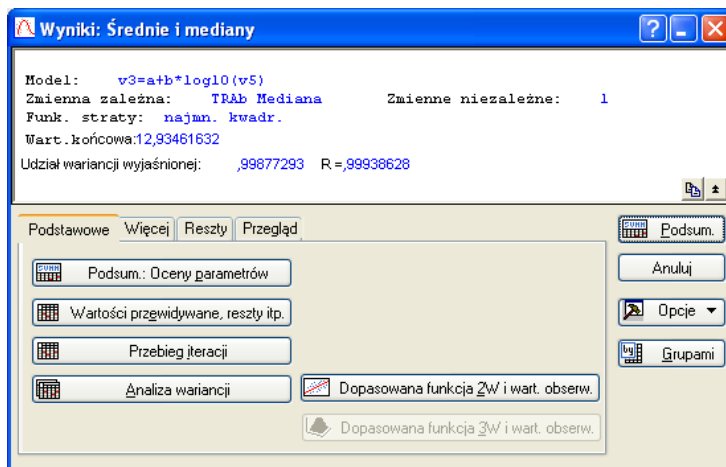
Wagami są liczby obserwacji.



Jako metodę estymacji wybieramy proponowaną przez program metodę Levenberga-Marquarta.



Na kolejnym panelu wybieramy *Podsum.: Oceny parametrów*.



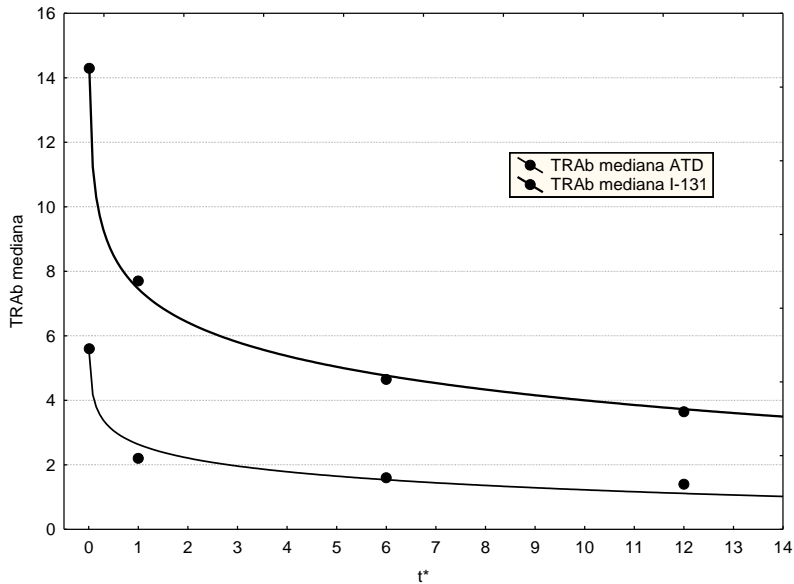
Dane: Model: $v3=a+b*\log_{10}(v5)$ (Średnie i medi...)

Zmn. zal.: TRAb Mediana I-131
 Poziom ufności: 95.0% (alpha=0.050)

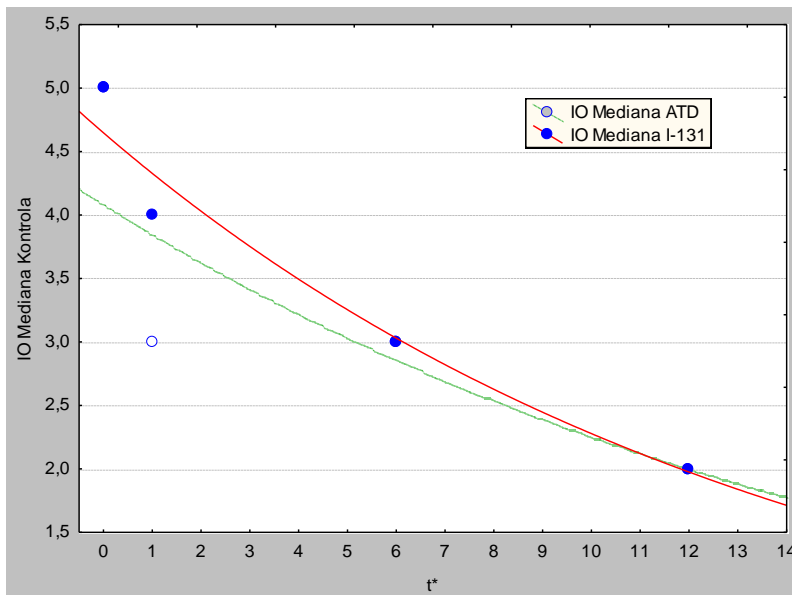
	Ocena	Błąd stand.	Wart. t df = 590	p	Doln. uf Granica	Górn. uf Granica
a	7,45987	0,006123	1218,374	0,0000	7,44785	7,47190
b	-3,45658	0,004988	-692,988	0,0000	-3,46637	-3,44678



Model jest wysoce istotny statystycznie. Współczynnik determinacji wynosi niemal 0,999. W podobny sposób oszacowano trend dla grupy ATD. Skorzystano z klawisza *Dopasowana funkcja 2W i wart. obserw.*, a następnie scalono obydwie rysunki.



Dla obydwu grup oszacowano logarytmiczne funkcje trendu. Obydwie dobrze oddają zmiany w medianach TRAb. Podobne analizy przeprowadzone dla indeksu oftalmopatii dały wynik zobrazowany na kolejnym rysunku.





IO przyjmuje wartości całkowite i stąd mediany są też wartościami całkowitymi. W porównywanych grupach różniły się tylko w pierwszym miesiącu: 4 dla 131-I oraz 3 dla ATD. Naturalne jest więc, że ważone funkcje trendu zbliżają się do siebie.

Prognozowanie non-responders

Do grupy non-responders zaliczono wszystkich pacjentów w grupie leczonej ADT i grupie leczonej 131-I, którzy wymagali dalszego stosowania glikokortykoidów doustnie po zakończonym leczeniu pulsami SoluMedrolu.

W grupie 131-I 13/46 pacjentów (11 kobiet i 2 mężczyzn) wymagało dalszego leczenia doustnymi glikokortykoidami. Przeprowadzono analizę dyskryminacyjną (z krokową eliminacją zmiennych nieistotnych) uwzględniającą: płeć, wiek, czas rozpoznania nadczynności tarczycy, czas rozpoznania GO, TSH-0, FT4-0, TRAb-0, CAS-0 oraz IO-0, w wyniku której pozostały dwie zmienne mogące mieć znaczenie różnicujące szansę bycia non-responder: TSH-0 i IO-0.

Nie pokazujemy tutaj w szczegółach przebiegu analizy dyskryminacyjnej. Pozwoliła ona wstępnie zidentyfikować czynniki umożliwiające prognozowanie wariantu zmiennej dwustanowej. W dalszej kolejności zbudowano model regresji logistycznej. Znajduje się on w już prezentowanym module *Estymacja nieliniowa* jako opcja *Szybka regresja logistyczna*. W otrzymanym modelu obydwie zmienne, mierzone w miesiącu zero, wykazują jednak istotność statystyczną na poziomie 0,10. Zastosowanie kryterium eliminacji zmiennej z wykorzystaniem wartości p wskazuje, że z modelu należy usunąć poziom TSH, a pozostawić indeks oftalmopatii. Niestety model tylko z jedną zmienną w dalszym ciągu nie wykazuje istotności na poziomie 0,05.

Dane: Model: Regr. logistyczna (logit) N zer: 32 jedynek: 13 (AW 110126)*

Model: Regr. logistyczna (logit) N zer: 32 jedynek: 13 (AW 110126)
Zmn. zal.: sterydy po pulsach Strata: Największe prawd. bł.średnkw.skala.
Całkowita strata: 23,665922187 Chi2(2)=6,7720 p=,03386
Warunek uwzględniania: v1=101

	Stała B0	TSH 0'	IO 0		
N=45					
p	0,0150	0,0930	0,0729		
iloraz szans z.jedn.	0,03	1,04	1,54		
-95%CL	0,00	0,99	0,96		
+95%CL	0,50	1,09	2,46		

W grupie ADT 69/167 pacjentów było leczonych glikokortykoidami doustnymi po zakończonym leczeniu pulsacyjnym. W modelu wstępnym analizy dyskryminacyjnej uwzględniono tylko 109 pacjentów mających komplet analizowanych danych: płeć, wiek, czas rozpoznania GO, TSH-0, FT4-0, TRAb-0, Cas-0, IO-0. Na podstawie modelu wstępnego ustalono, że tylko IO-0 może mieć znaczenie różnicujące szansę bycia non-responder. W modelu ostatecznym regresji logistycznej wyliczono, że każdy punkt IO-0 więcej przed leczeniem GO zwiększa ryzyko względne bycia non-responder średnio o 30%.



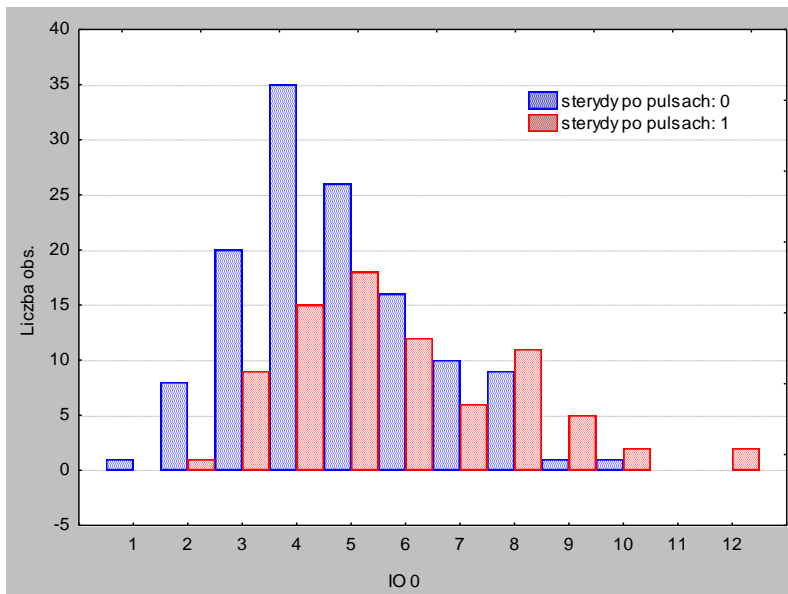
Dane: Model: Regr. logistyczna (logit) N zer: 95 jedynek: 68 (AW 110126)*

Model: Regr. logistyczna (logit) N zer: 95 jedynek: 68 (AW 110126)
Zmn. zal.: sterydy po pulsach Strata: Największe prawd. bł. średnk.w. skal
Całkowita strata: 105,30363941 Chi2(1)=10,866 p=,00098
Warunek uwzględniania: v1=102

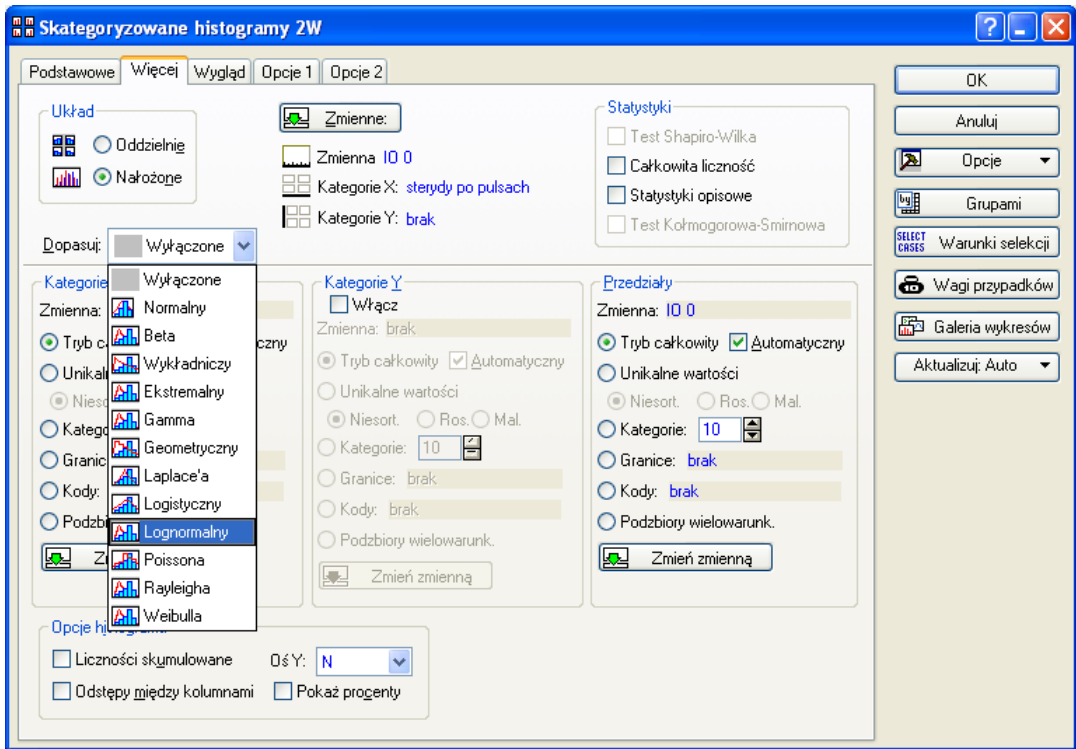
	Stała B0	IO 0				
N=163						
p	0,0004	0,0019				
Iloraz szans z jedn.	0,18	1,30				
-95%CL	0,07	1,10				
+95%CL	0,45	1,54				

Ostatecznie wartość indeksu oftalmopatii w momencie zero pozostała zmienną możliwą do wykorzystania do prognozowania braku odpowiedzi na leczenie. Ponieważ ryzyka względne odnoszące się do jednostkowej różnicy IO-0 były podobne w obu grupach, zdecydowano się więc w tej części analizy potraktować te grupy łącznie.

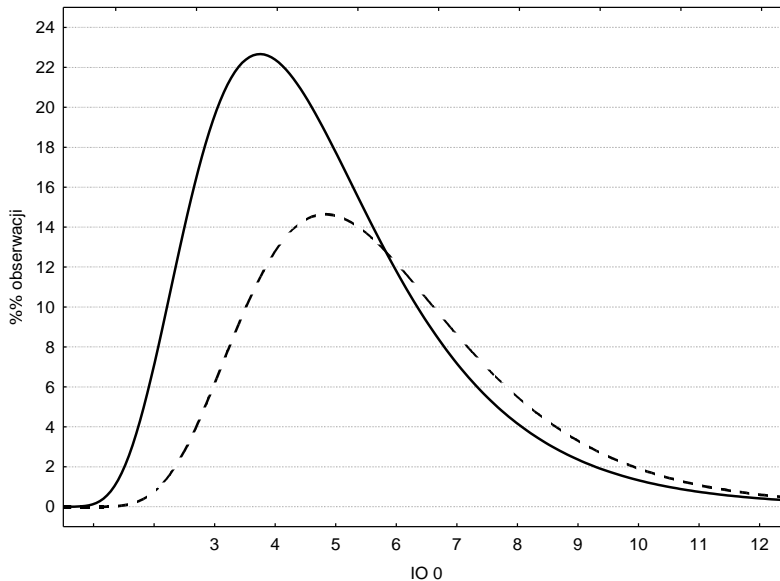
Sporządzamy histogram skategoryzowany wartości IO-0 w zależności od kodu zmiennej *sterydy po pulsach*, która identyfikuje non-responders. Na pierwszym rysunku pokazujemy histogramy bez funkcji aproksymujących.



Na karcie *Więcej* można wybrać rozkład aproksymujący histogram. My wybieramy rozkład log normalny.



Teraz z otrzymanego rysunku trzeba usunąć histogramy pozostawiając same funkcje gęstości rozkładów.





Na takim wykresie łatwiej zidentyfikować optymalny punkt odcięcia. Tu widać, że większe ryzyko bycia non-responder (niż zareagowania na leczenie) mają pacjenci, których indeks oftalmopatii w momencie wyjściowym wynosi co najmniej 5. Po przekroczeniu tej wartości krzywe mają i tak dość zbliżony przebieg. Proponowana reguła decyzyjna bierze pod uwagę fakt, że non-responders stanowią około 40% ogółu pacjentów. Czulość wynosi 0,47, specyficzność 0,71, dodatnia wartość predykcyjna 0,51, a ujemna wartość predykcyjna 0,68. Nie są to zbyt zachwycające charakterystyki.