

Analiza statystyczna wybranych parametrów średnich pEMG, uzyskanych podczas oceny diagnostycznej zwieracza zewnętrznego odbytu

Paulina Trybek¹, Michał Nowakowski², Łukasz Machura¹

¹ Zakład Fizyki Teoretycznej, Instytut Fizyki, Uniwersytet Śląski, Bankowa 10, 40-006 Katowice,

² Zakład Dydaktyki Medycznej, Wydział Lekarski, Uniwersytet Jagielloński, Collegium Medicum, Łazarza 16, 31-530 Kraków

* paulina.trybek@us.edu.pl

Streszczenie

Typowe metody diagnostyczne wykorzystywane przy ocenie dysfunkcji zwieracza zewnętrznego odbytu bazują na technikach manometrycznych bądź wykorzystujących ultrasonografię endoskopową i następujących po nich testach statystycznych pre-post, które umożliwiają automatyczną klasyfikację. W tej pracy opisujemy szeroko zakrojone testy zarówno parametryczne jak i nieparametryczne, które potwierdzają znaczenie nieinwazyjnej techniki, jaką jest elektromiografia powierzchniowa, jako istotnej techniki diagnostycznej. W szczególności zaprezentowane testy nieparametryczne wydają się być wysoce obiecującymi metodami, pozwalającymi na grupowanie pacjentów w zależności od rodzaju dysfunkcji.

Abstract

The typical diagnostic evaluation of the dysfunction of the external anal sphincter is based on manometry or endoscopic ultrasound techniques and following pre-post statistical tests which enables the semi-automated classification. Here we discuss the extensive statistical parametric and nonparametric tests which support the importance of the non-invasive surface electromyography as a significant diagnosis tool. Especially presented nonparametric tests seem to constitute the promising methodology which allow grouping of the patient according to the certain dysfunction type.

Interpretacja wyników szeroko pojętych badań medycznych (od wspomaganie decyzji diagnostycznych po analizy dotyczące badań klinicznych) bez zastosowania narzędzi statystycznych, bezdyskusyjnie nie jest możliwa we współczesnej medycynie. Problem dotyczy raczej wyboru optymalnych procedur obliczeniowych, spośród coraz szerszej gamy dostępnych pakietów i programów, konkurujących ze sobą na

poziomie specjalizacji odpowiednich metod, mających na celu jak najwierniej oddać przebieg badanych zależności. Jako przykład zostanie przedstawiona analiza statystyczna wyników badania elektromiograficznego, konkretnie elektromiografii powierzchniowej (pEMG) [1], oparta głównie o analizę porównawczą pomiędzy grupą zdrowych, stanowiących grupę kontrolną a grupą badaną, z dysfunkcją zwieracza zewnętrznego odbytu.

Elektromiografia

EMG jest techniką uzyskiwania sygnałów bioelektrycznych, powstałych na skutek zmian fizjologicznych w błonie komórkowej i składzie elektrolitowym płynu wewnątrz i zewnątrzkomórkowego, które warunkują istnienie w poprzek błony potencjału. Podstawą funkcjonowania każdej żywej komórki jest stała i odpowiednio szybka reakcja na zmiany zachodzące w jej otoczeniu. Najszybsza, wykształcona sygnalizacja opiera się na zmianie potencjałów błonowych. Sam potencjał błonowy jest charakterystyczny dla każdej komórki, jednak jego zmienność w czasie zarezerwowana jest dla neuronów oraz komórek mięśniowych. W dużym uproszczeniu to właśnie dzięki tej zmienności, możemy zmierzyć aktywność mięśniową, używając odpowiedniej aparatury pomiarowej. Pomijając szereg fizjologicznych i typowo biochemicznych procesów, charakteryzujących proces aktywacji komórki mięśniowej prowadzący w konsekwencji do skurczu mięśnia, w odniesieniu do rejestracji sygnału bioelektrycznego należy wspomnieć dokładniej o wspomnianym wyżej potencjale błonowym. Możemy mówić o trzech fazach: depolaryzacji, repolaryzacji i fazie spoczynkowej. Wszystkie uwarunkowane są zmianą przepuszczalności błony komórkowej dla konkretnych jonów, z których najważniejsze to kationy sodowe, kationy potasowe i aniony chlorkowe. W stanie spoczynku stężenie kationów potasowych wewnątrz komórki znacznie przewyższa zewnątrzkomórkowe stężenie tych jonów. Dokładnie odwrotna sytuacja rysuje się dla kationów sodowych i anionów chlorkowych. Ich zewnętrzne stężenie jest większe. Równowagę elektrolitową dla stanu spoczynku opisuje zależność Goldmana:

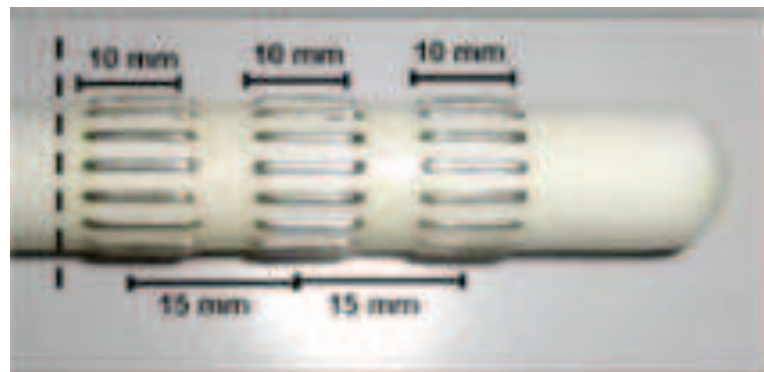
$$E = \left(\frac{RT}{F}\right) \ln \frac{P_k[K^+]_z + P_{Na}[Na^+]_z + P_{Cl}[Cl^-]_w}{P_k[K^+]_w + P_{Na}[Na^+]_w + P_{Cl}[Cl^-]_z} \quad (1)$$

Wzór przypomina dobrze znaną fizyce i elektrochemii zależność Nernsta, gdzie odpowiednio R, F to stałe gazowa i Faradaya, P oznacza przepuszczalność dla poszczególnych jonów, natomiast w nawiasach kwadratowych przedstawiono stężenia zaindeksowane „w” lub „z” w odniesieniu do stężenia wewnątrz i zewnątrzkomórkowego. Wartość takiego spoczynkowego potencjału dla komórki mięśniowej wynosi ok. -70mV. W momencie, kiedy do komórki dociera sygnał przez złącze synaptyczne, w wyniku szeregu procesów fizjologicznych, poprzez otwieranie się kanałów, specyficznych dla odpowiednich jonów, następuje przeorientowanie wyżej opisanego stanu spoczynku. Do komórki gwałtownie napływa prąd sodowy i następuje depolaryzacja

błony określana jako skok potencjału czynnościowego, z wartości ujemnej (-70 mV) do wartości w okolicach $+30$ mV. Jest on wywołany nagromadzeniem wewnątrz komórki ładunku dodatniego. Natomiast podczas repolaryzacji następuje powrót do stanu wyjściowego. Z punktu widzenia pomiaru aktywności mięśniowej najważniejszy jest etap szerzenia się fali potencjału czynnościowego wzdłuż błony komórek mięśniowych, czyli jej depolaryzacja. Bowiem do różnicy potencjałów wywołanym gradientem koncentracji jonów po obu stronach błony dochodzi różnica pomiędzy obszarami już zdepolaryzowanymi, a tymi w stanie spoczynku. Ta ostatnia tworzy swoistego rodzaju dipol, który możemy obserwować za pomocą zewnętrznej aparatury pomiarowej w postaci zmiany sygnału w czasie.

Rejestracja sygnału

Na uwagę zasługuje fakt, że powierzchniowa elektromiografia jest techniką nieinwazyjną. Sygnał zbierany jest przy użyciu elektrod powierzchniowych w przeciwieństwie do technik igłowych, w których elektroda jest bezpośrednio wkłuwana w mięsień. W technice powierzchniowej elektrody umieszczone są na skórze lub jak analizowanym przypadku, gdzie badany jest specyficzny mięsień, skonstruowany został specjalny zestaw elektrod doodbytniczych w postaci tzw. sondy anorektalnej [2]. Zwieracz zewnętrzny odbytu należy do mięśni okrężnych, tworzących pierścienie dookoła tkanki ciała, dlatego elektrody zostały umieszczone pierścieniowo wokół sondy – po 16 par elektrod na trzech poziomach odpowiadających anatomii kanału odbytu, według odpowiednio ustalonej geometrii (przedstawionej poniżej) [3].



Rys. 1. Sonda anorektalna.

Analizowane parametry

Podstawowymi parametrami poddanymi analizie, są średnie wartości opisujące charakter sygnału, konkretnie jego amplituda i częstotliwość. Ponieważ ideą projektu, w trakcie którego otrzymano dane, była ocena przydatności elektromiografii w diagnostyce nietrzymania stolca, równoległe dokonano analizy statystycznej wyników niektórych badań wchodzących w skład standardu postępowania w takim przypadku.

Wśród badań znajdowała się m.in. manometria anorektalna, polegająca na pomiarze ciśnienia w kanale odbytu, a także ocena odruchów odbytowych, pozwalająca

na sprawdzenie prawidłowości unerwienia. Tę ostatnią można było sklasyfikować poprzez dane kategoryjne.

Średnia amplituda sygnału pEMG wyliczona została, jako \bar{x} RMS (Root Mean Square), czyli średnia kwadratowa. Jak już wspomniano przy opisie konstrukcji sondy, elektrody w liczbie 16-par zostały umieszczone na trzech poziomach sondy, odpowiadając kolejno pierścieniom dystalnemu, środkowemu oraz proksymalnemu, aby móc zarejestrować sygnał bioelektryczny z różnych głębokości kanału odbytu. W celach diagnostycznych dokonywano pomiaru w stanie skurczu i relaksacji mięśnia. Było to możliwe z uwagi na fakt, że zwieracz zewnętrzny odbytu zalicza się do kategorii mięśni poprzecznie prążkowanych, unerwianych przez układ somatyczny, którego działanie w dużym stopniu podlega kontroli. Układ ten kieruje bowiem czynnościami zależnymi od woli organizmu. W związku z tym dla każdej badanej osoby mamy następujący zestaw parametrów opisujących wartość amplitudy: średnią wartość (średnią kwadratową) z 16 kanałów, odpowiadających 16 parom elektrod na konkretnym pierścieniu, czyli średnią kwadratową z wartości od K1 do K16, liczoną osobno dla każdego pierścienia oraz stanu skurczu i spoczynku. Łącznie, dla 1 osoby mamy zestaw 6 wartości średnich: \bar{x} RMS 1M; \bar{x} RMS 2M; \bar{x} RMS 3M; \bar{x} RMS 1R; \bar{x} RMS 2R i \bar{x} RMS 3R. Numery 1, 2, 3 odpowiadają położeniu pierścieni (1 oznacza pierścień dystalny, najbliższy wierzchołka sondy, 2-pierścień środkowy, 3-pierścień proksymalny). Oznaczenia M i R odnoszą się odpowiednio do skurczu maksymalnego mięśnia (M) i jego stanu spoczynkowego (R). Analogiczny zestaw parametrów występuje dla średniej częstotliwości sygnału MF (Mean Frequency). Jeśli chodzi o wyniki już wcześniej wspomnianych pozostałych badań, to dla manometrii anorektalnej mamy dwie wartości ciśnienia: MCP- maksymalne ciśnienie podstawowe, jako ciśnienie spoczynkowe mierzone w dowolnym miejscu kanału odbytu oraz MCS- maksymalne ciśnienie skurczowe, maksymalna wartość ciśnienia, zmierzona w tym samym miejscu, w którym uprzednio zarejestrowano MCP.

W przypadku oceny odruchów mamy podział na odruch kaszlowy OK i odruch parcia OP. Dla OK wynik jest dodatni jeżeli odruch jest obecny (a dodatkowo wartość ciśnienia w kanale odbytu jest wyższa niż w bańce odbytnicy). Dla odruchu parcia występuje trzystopniowe kodowanie: 0-odruch nieobecny, 1-obecny odruch patologiczny, 2-obecny odruch prawidłowy. Dla sumarycznego przedstawienia wyników z OK i OP skala rozszerza się na zakres od 0-3. Omówione parametry zostały porównane zarówno na poziomie grup choroby-zdrowi, a także w obrębie samej grupy badanej, dokonując podziału ze względu na etiologię schorzenia. Do wykonywania potrzebnych obliczeń oraz prezentacji wyników wykorzystywany był program STATISTICA 10 PL.

Charakterystyka grupy badanej i kontrolnej

Grupę kontrolną stanowili zdrowi ochotnicy, natomiast grupę badaną osoby ze stwierdzeniem problemu nietrzymania stolca o różnej etiologii tego schorzenia [4, 5]. W przypadku rekrutacji do obydwu grup obowiązywały restrykcyjne zasady włącza-

nia, a zarazem wykluczenia z grupy, które dotyczyły głównie wskazań medycznych, w tym także przebytych schorzeń oraz odpowiedniego przedziału wiekowego. Charakterystyka poszczególnych grup, uwzględniająca podział na płeć i średni wiek w grupach, została przedstawiona w Tabeli 1.

Tabela 1. Podział na płeć i średnią wieku dla grupy badanej i kontrolnej.

		N ważnych	Średnia [lata]	Minimum	Maksimum	Odchylenie Standardowe
Grupa Badana	Kobiety	39	60.2	25	81	12.6
	Mężczyźni	10	53.8	21	81	17.6
Grupa kontrolna	Kobiety	39	42.7	23	75	15.3
	Mężczyźni	10	55.9	45	75	9.2

Grupy są równoliczne pod względem płci, jednak w obydwu przypadkach (grupy badanej i kontrolnej) to kobiety stanowią 80 procent. Prawie czterokrotnie wyższa liczba kobiet jest odzwierciedleniem proporcji pomiędzy zapadalnością na badaną chorobę u kobiet i mężczyzn w populacji ogólnej.

Rezultaty statystyk wstępnych

Aby analiza podążała właściwym torem, krokiem podstawowym było przeanalizowanie rozkładów badanych zmiennych. Ścisłej rzecz ujmując, wymagane było sprawdzenie, czy rozkład rozpatrywanej zmiennej możemy opisać przy pomocy rozkładu normalnego (gaussowskiego), który opisuje następująca zależność:

$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad (2)$$

gdzie: Y – wartość krzywej rozkładu,

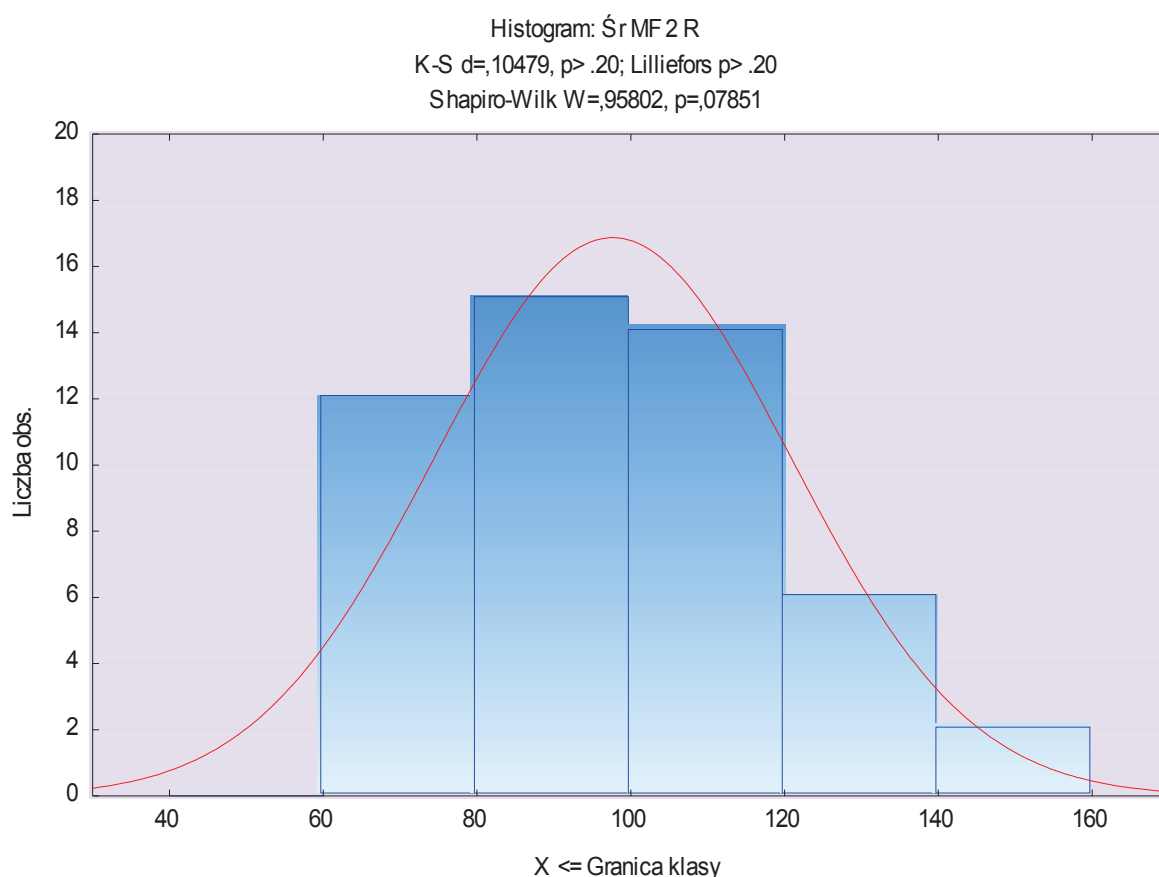
μ, σ – odpowiednio średnia i odchylenie standardowe.

Normalność rozkładu jest warunkiem koniecznym w przypadku analiz parametrycznych (np. grupy testów), a także analizie wariancji, która pozwala na podział zmienności zaobserwowanej wśród otrzymanych wyników na oddzielne części oraz rozpatruje wpływ czynników na analizowaną zmienną, z osobna lub łącznie, w zależności od modelu analizy wariancji. Można powiedzieć, że rozkład normalny jest wymagany w analizach o największej mocy wykrywania rzeczywistych efektów, dlatego test normalności powinien stanowić etap poprzedzający właściwą analizę. W celu uzyskania empirycznego rozkładu częstości danych można utworzyć wykres danych punktowych, histogram lub wykres łodyga-liście i sprawdzić czy kształt wykresu odpowiada rozkładowi normalnemu. Alternatywę może stanowić wykres normalności, mianowicie wykres rozkładu częstości danych w funkcji rozkładu normalnego. Jeżeli taki wykres będzie linią prostą, to badany rozkład w przybliżeniu jest rozkładem normalnym [6, 7]. Ponieważ opisane wyżej metody są wysoce subiektywne i nie ce-

chuje ich duża dokładność, ocena normalności zostanie przeprowadzona z użyciem testów Lilieforsa, Shapiro-Wilka, a także Kołmogorowa Smirnowa. Wyniki dla analizowanych parametrów przedstawia poniższa tabela, w której wyróżniono, dla jakich parametrów wyliczona wartość p jest większa od przyjętego poziomu istotności $\alpha = 0.05$, który określa prawdopodobieństwo odrzucenia hipotezy zerowej, gdy wyliczona wartość p jest od niego mniejsza [6]. Dla $p > 0.05$ nie ma wystarczających podstaw do odrzucenia hipotezy zerowej, która zakłada normalność rozkładu.

Tabela 2. Wyniki testów Lilieforsa i Shapiro Wilka.

Parametr	Grupa badana	Grupa kontrolna	Razem
Śr RMS	-	-	-
Śr MF	$p > 0.05$ dla 2R	-	-
MCP	-	-	$p > 0.05$
MCS	-	$p > 0.05$	$p > 0.05$



Rys. 2. Histogram dla śr MF 2R.

Z wyników przedstawionych w tabeli widać, że dla parametrów sygnału pEMG tylko dla jednej zmiennej, przedstawionej na powyższym histogramie (średniej częstotliwości na pierścieniu środkowym w stanie spoczynku) nie mamy podstaw do odrzucenia hipotezy zerowej o normalności rozkładu. Dla parametrów manometrii anorektalnej, biorąc pod uwagę obydwie grupy łącznie, zarówno ciśnienie spoczynkowe jak i skurczowe charakteryzuje się normalnością rozkładu. Dla grup rozpatrywanych

osobno nie możemy odrzucić hipotezy o normalności rozkładu dla maksymalnego skurczowego ciśnienia w grupie kontrolnej. Na powyższym histogramie widać również wartość testu Kołmogorowa-Smirnowa (K-S). Wyniki dla tego testu w grupach rozpatrywanych z osobna dają znacznie większą liczbę parametrów, dla których nie mamy podstaw odrzucenia hipotezy o normalności.

grupa kontrolna: śr RMS 1M, śr RMS 2M, śr RMS 3M, śr MF

grupa badana: śr RMS 1M, śr RMS 2M, śr RMS 3R, śr MF1M, śr MF 2M, śr MF 3M, śr MF 2R.

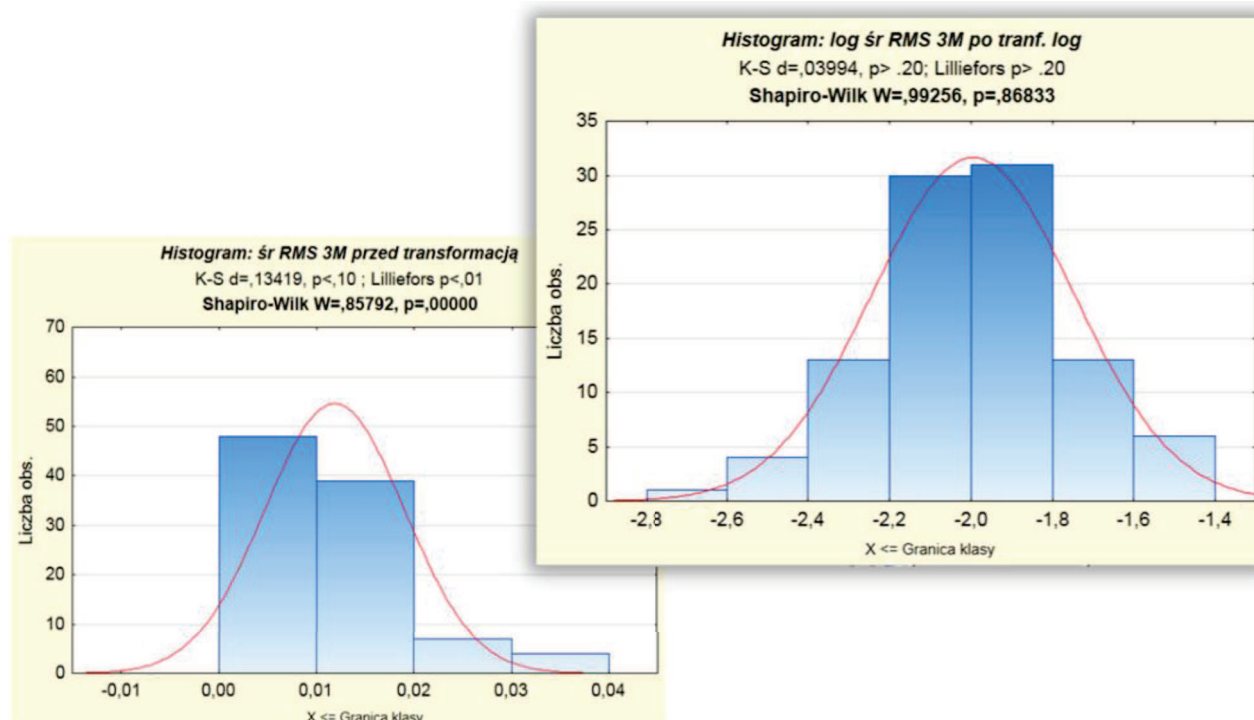
Zwiększona tolerancja ostatniego testu, nie zmienia faktu, że nie wszystkie parametry cechuje normalność rozkładu, a test Shapiro-Wilka i Lilieforsa można powiedzieć zdecydowanie wykluczył rozkłady gaussowskie dla parametrów amplitudy i częstotliwości. Zatem w związku z niespełnieniem założeń istnieje konieczność w dalszych analizach stosowania statystyk wyłącznie nieparametrycznych.

Transformacja danych

Statystyki parametryczne oraz analiza wariancji są źródłem niezwykle cennych informacji o zależnościach pomiędzy badanymi zmiennymi i nie można postawić znaku równości pomiędzy wartością statystyk para i nieparametrycznych, te ostatnie bowiem są przede wszystkim testami istotności, które nie pozwalają na uzyskanie estymatorów interesujących nas efektów. Dlatego przed ostatecznym wykluczeniem możliwości zastosowania statystyk parametrycznych warto przeanalizować dane np. pod kątem możliwości ich transformacji. W analizowanym przypadku, po uprzednim wyliczeniu współczynników asymetrii, dla których wyliczona wartość była dodatnia i dla prawie wszystkich parametrów amplitudy oraz częstotliwości wskazywała na silną asymetrię (wyliczony współczynnik asymetrii > 1), najodpowiedniejszą transformacją była transformacja logarytmiczna. Dodatni, współczynnik, wskazuje na prawoskośność rozkładu. Jeżeli rozkład y jest prawo skośny, to $z = \log(y)$ posiada w przybliżeniu rozkład normalny (6). Rezultat takiej transformacji został przedstawiony na poniższych histogramach (Rys. 3).

Z przedstawionych histogramów widać, że transformacja logarytmiczna osiągnęła zamierzony skutek dla rozpatrywanego przypadku. Rozkład po transformacji logarytmicznej jest rozkładem normalnym. Tabela 3 przedstawia rezultaty dla wszystkich analizowanych parametrów.

Transformacja logarytmiczna przyniosła zaskakująco obiecujące wyniki dla parametrów amplitudy i częstotliwości w przypadku grupy badanej, w której wszystkie rozkłady za wyjątkiem śrMF 1R są rozkładami normalnymi. Ponieważ jednak nie dla wszystkich parametrów zlogarytmowanie danych pozwoliło na uzyskanie normalności rozkładów, ostatecznie zachodzi konieczność zastosowania statystyk nieparametrycznych w dalszych etapach analizy.



Rys. 3. Rezultat transformacji logarytmicznej (po lewej stronie rozkład po transformacji).

Tabela 3.

	Grupa badana	Grupa kontrolna	Razem
śr RMS	$p > \alpha$ dla wszystkich	$p > \alpha$ dla 1M, 3M	$p > \alpha$ dla 1M,2M,3M,1R
śr MF	$p > \alpha$ dla wszystkich z wyjątkiem 1R	$p < \alpha$	$p > \alpha$ dla 3R
MCP	$p < \alpha$	$p > \alpha$	$p > \alpha$ przed transformacją
MCS	$p > \alpha$	$p > \alpha$	$p > \alpha$ przed transformacją

Testowanie hipotez na podstawie statystyk nieparametrycznych

Istnieje wiele grup testów nieparametrycznych, dostosowanych tak jak ich parametryczne odpowiedniki do rodzaju zależności, którą chcemy zbadać, charakteru zmiennej, liczebności grup etc. Ogólnie przyjęty podział obejmuje grupy testów zgodności, weryfikujących hipotezę o postaci funkcyjnej rozkładu zmiennej w populacji, testy losowości badające charakter losowy próby oraz testy wnioskujące o niezależności dwóch zmiennych losowych. Najczęściej stosowanymi testami na poziomie analizy porównawczej (np. przypadek-kontrola) są nieparametryczne odpowiedniki testów t oraz analizy wariancji [6].

Poniżej zostaną przedstawione wyniki wybranych analiz nieparametrycznych dla odpowiednich zmiennych w wybranych grupach. W zależności od stosowanego testu dla określonych parametrów badano różnice na poziomie stanu zdrowia, poziomu 1, 2, 3 odpowiadającemu danemu pierścieniowi na sondzie (anatomicznie odpowied-

niemu poziomowi w kanale odbytu) oraz etiologii schorzenia. W przypadku ostatniego kryterium, w grupie chorych, liczącej 49 osób znajdowały się 24 osoby o nieokreślonej przyczynie nietrzymania stolca (idiopatyczne); 19 osób o pochodzeniu neurogennym, oraz 6 osób z uszkodzeniem aparatu zwieraczowego. Brak zatem grup równolicznych dla wyżej wymienionych etiologii. Rodzaj zastosowanych statystyk wraz z przyporządkowaniem badanej zależności został przedstawiony w Tabeli 4.

Tabela 4.

Test/ statystyka	Sumy Rang Wilcoxon	Kruskala- Walisa	Anova Fried- mana	U Manna Whitneya
Kryterium porównawcze	Stan zdro- wia/etiologie	Etiologia	Stan mięśnia	Stan zdrowia z jednoczesnym względnieniem etiologii
Porównywane Grupy	Zdrowi-Chorzy Idiopat.- neurogenne	Neurogenne uszkodzenie idiopatyczne	skurcz – stan spoczynkowy	Zdrowi-Chorzy (idiopatyczne) Zdrowi-Chorzy (neurogenne) Zdrowi-Chorzy (uszkodzenie)

Pierwszy z wymienionych w tabeli testów jest nieparametrycznym odpowiednikiem testu t studenta dla zmiennych niepowiązanych. Oparty jest na rangowaniu obserwacji, a jego prosta postać statystyki testowej zależna jest od liczebności grup i dla analizowanego przypadku wynosi:

$$Z = \frac{T - \mu_T}{\sigma_T} \quad (3)$$

gdzie: T oznacza sumę rang, natomiast

μ_T, σ_T odpowiednio wyliczone liczebności porównywanych grup [6, 8].

Z użyciem tej statystyki zweryfikowana została hipoteza o równości mediany porównywanych parametrów w odpowiednich grupach. Parametry amplitudy, częstotliwości oraz ciśnienia porównano pomiędzy grupą kontrolną i badaną, a także grupą kontrolną i wybranych etiologiach. Wyniki zawierają poniższe tabele, w których zawarto wartość wyliczonej statystyki Z , oraz p dla przyjętego poziomu istotności $\alpha = 0.05$.

Tabela 5.

Zmienna	Grupa		N ważnych	Z	P
MCP	Chorzy	Zdrowi	49/49	-6.6	0
	Idiopatyczne	neurogene	24/19	-2.091	0.037
RMS	Chorzy	Zdrowi	49/49	-5.6	0
	Idiopatyczne	neurogene	24/19	-3.852	0

Dla parametrów ciśnienia istnieją wystarczające podstawy dla odrzucenia hipotezy zerowej i przyjęcia hipotezy o istotnych statystycznie różnicach w medianach dla wszystkich porównywanych grup. W związku ze zbyt małą liczebnością grup, niemożliwe było porównanie grup: chorzy (idiopatyczne)-chorzy (uszkodzenie) oraz chorzy (neurogenne)-chorzy (uszkodzenie).

W tabeli pogrubione zostały wartości $p > \alpha$, dla których nie mamy wystarczających podstaw na odrzucenie hipotezy zerowej, wnioskującej o równości mediany w analizowanych grupach. Zatem porównując całą grupę badaną z grupą kontrolną, dla $\overline{\text{rRMS}}$ na trzech poziomach kanału odbytu, wykazano istotnie statystycznie różnice w stanie skurczu. Brak wykrytych różnic dla stanu spoczynkowego. W zestawieniu idiopatyczne-neurogenne różnice dla pierścienia dystalnego i proksymalnego w fazie skurczowej oraz dystalnego w fazie spoczynkowej. Dla średniej częstotliwości różnice istotne statystycznie dla pierścienia środkowego i proksymalnego w stanie skurczu, tylko dla porównania na poziomie chorzy-zdrowi.

Tabela 6.

Parmetr	Grupa		N		Z			P		
$\overline{\text{rRMS}}$ M	cho- rzy	Zdro- wi	49	49	3.016	2.327	4.316	0.002 56	0.019 97	0
	Idiop .	Neu- rog.	24	19	-2.446	-1.271	-2.372	0.014 46	0.203 5	0.017 7
$\overline{\text{rRMS}}$ R	cho- rzy	Zdro- wi	49	49	-0.876	-1.744	-0.089	0.380	0.081	0.929
	Idiop .	Neu- rog.	24	19	-2.128	-1.125	-1.418	0.033	0.260	0.156
$\overline{\text{rMF}}$ M	cho- rzy	Zdro- wi	49	49	-1.039	-2.390	-3.321	0.301	0.016 8	0
	Idiop .	Neu- rog.	24	19	0.514	1.516	0.391	0.608	0.129	0.696
$\overline{\text{rMF}}$ R	cho- rzy	Zdro- wi	49	49	1.453	-0.124	-1.808	0.146	0.901	0.071
	Idiop .	Neu- rog.	24	19	0.196	0.033	1.492	0.845	0.067	0.136
Nr pierścienia (poziom kanał)					1	2	3	1	2	3

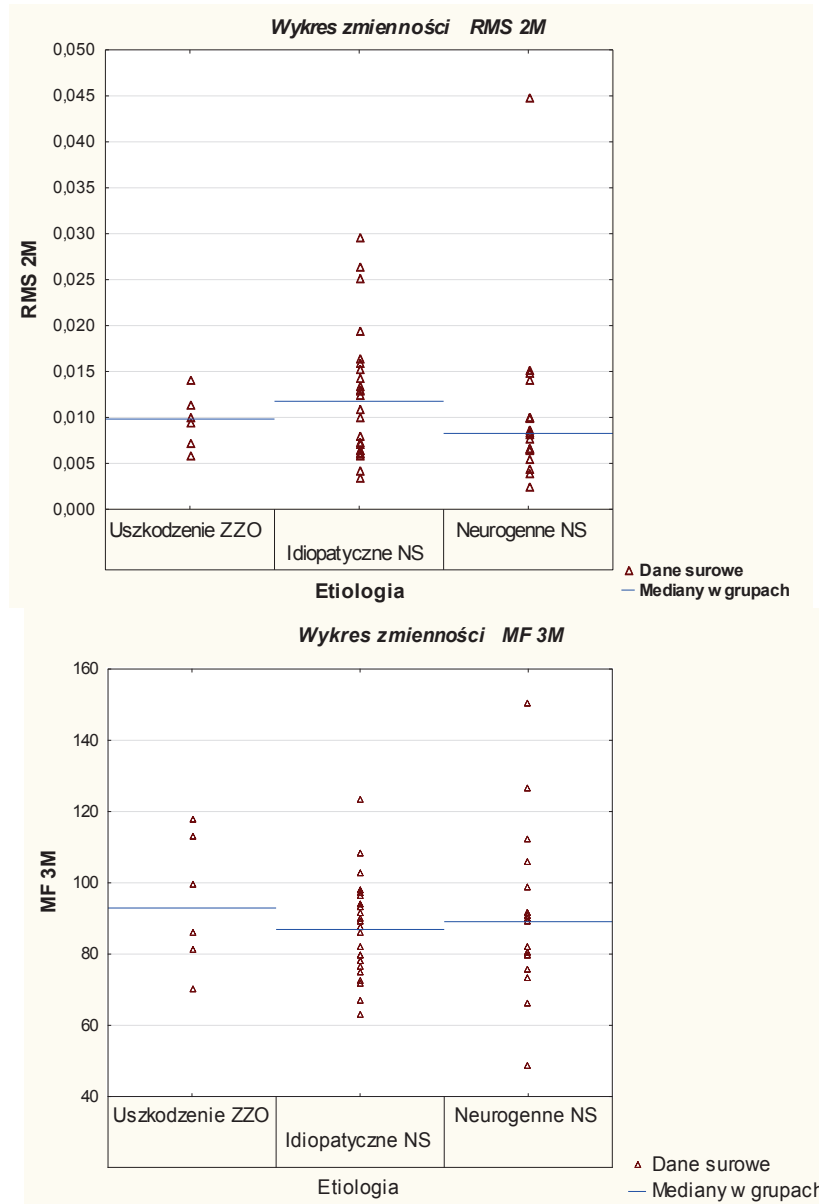
Test Kruskala-Walisa (dla wielu prób niepowiązanych)

Test Kruskala-Walisa stanowi rozszerzenie testu sumy rang Wilcoxon [6]. Statystyka testowa podlega rozkładowi χ^2 i jest postaci:

$$H = \frac{12}{n(n+1)} \sum \frac{R_i^2}{n_i} - 3(n+1) \quad (4)$$

gdzie: n – liczebność grup,
 R_i^2 – suma rang.

Testowi prawdziwości została poddana hipoteza o równości rozkładów wśród porównywanych grup etiologicznych. Poniżej zostały przedstawione wykresy punktowe przykładowych parametrów częstotliwości i amplitudy z zaznaczoną medianą w poszczególnych grupach etiologicznych.



Rys. 4. Wykresy z zaznaczoną medianą.

Wyniki testu zostały przedstawione w tabeli, gdzie zamieszczono wartość statystyki testowej i wyliczoną wartość p .

Tylko w dwóch przypadkach wyliczona wartość p jest mniejsza od przyjętego poziomu istotności $\alpha = 0.05$. Dla śr RMS na pierścieniu proksymalnym w stanie skurczu maksymalnego oraz dla średniej częstotliwości w stanie spoczynkowym (pierścień dystalny) mamy wystarczające podstawy do odrzucenia hipotezy zerowej o równości rozkładów w porównywanych trzech grupach etiologicznych.

Tabela 7.

Parametr	Poziom kanału	h	p	parametr	Poziom Kanału	h	p
Śr RMSM	1	5.937	0.514	Śr MF M	1	3.273	0.195
	2	1.923	0.382		2	4.775	0.092
	3	6.905	0.032		3	0.991	0.609
Śr RMS R	1	4.446	0.108	Śr MF R	1	3.150	0.207
	2	1.258	0.533		2	6.580	0.037
	3	2.085	0.353		3	5.032	0.081

Wyniki dla parametrów ciśnienia przedstawione są poniżej w Tabeli 8.

Tabela 8.

Ciśnienie	h	p
MCP	5.478	0.065
MCS	16.964	0.0002

Tylko dla maksymalnego ciśnienia skurczowego możemy mówić o istnieniu statystycznie istotnych różnic w porównywanych grupach etiologicznych.

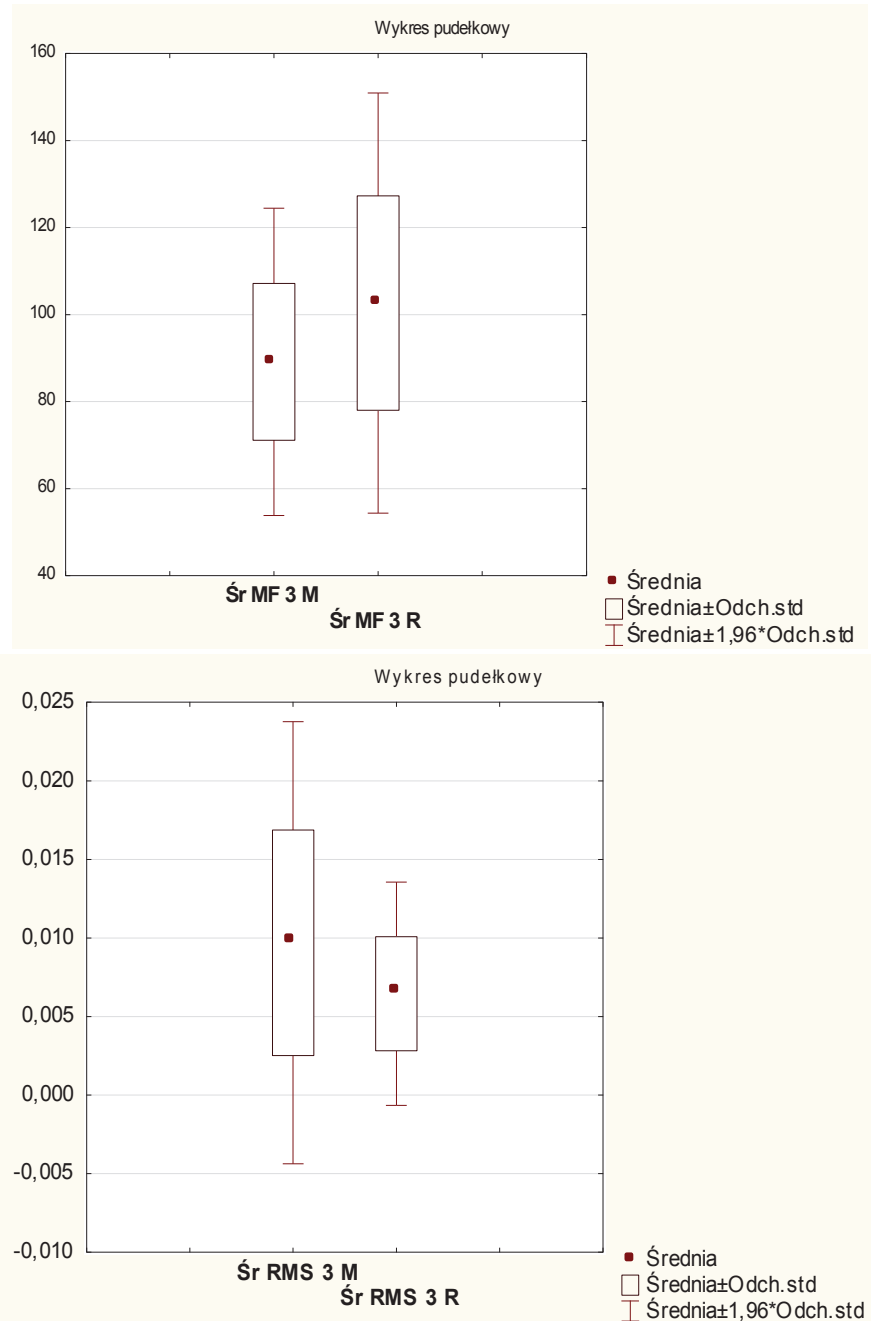
Statystyka Anova Friedmana

Statystyka Anova Friedmana dotyczy zmiennych zależnych. Stosujemy ją wtedy, gdy pomiaru analizowanej zmiennej dokonujemy kilkakrotnie w różnych warunkach ($k \geq 2$) [6]. W analizowanym przypadku zmienność warunków będzie dotyczyła stanu skurczu i spoczynku mięśnia. Zatem testowana będzie hipoteza o równości mediany badanego parametru w kolejnych pomiarach. Porównywane parametry to pary średnich parametrów amplitudy i częstotliwości dla odpowiadających sobie poziomów kanału w stanach skurczu i spoczynku np. śr RMS 1M-śr RMS 1R.

H_0 : Brak istotnej różnicy dla porównywanych parametrów w stanie skurczu maksymalnego (M) i stanie spoczynkowym (R).

H_1 : Istnieje istotna statystycznie różnica dla odpowiadających sobie parametrów w stanie skurczu i spoczynku.

Wynik przeprowadzonej analizy dla wszystkich porównywanych parametrów wskazuje na istotnie statystycznie różnice na poziomie skurcz-spozynek. Dla każdego z porównywanych przypadków (śr RMS 1-śr RMS 1R; śr RMS 2M-śr RMS 2R; śr RMS 3M-śr RMS 3R; śr MF 1-śr MF 1R; śr MF 2M-śr MF 2R; śr MF 3M-śr MF 3R) wyliczona wartość $p > \alpha = 0.05$. Poniżej przy użyciu wykresów pudełkowych zobrazowano te różnice dla wybranych przypadków.



Rys. 5. Wykresy pudełkowe lewej śrMF 3M - śr MF 3R, po prawej śrRMS 3M-śrRMS 3R.

Test U Manna-Whitneya

Test U Manna-Whitneya dotyczy porównania zmiennych zależnych i wnioskuje o równości median w porównywanych grupach [6, 9]. Jego statystyka uzależniona jest od liczebności grup i dla analizowanego przypadku przedstawia się jako:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \quad (1)$$

gdzie: n_1, n_2 – liczebność grup,

R_1 – suma rang przyznawanych wartościom pierwszej próby.

Weryfikacji została poddana hipoteza o równości mediany w porównywanych grupach: grupy kontrolnej z odpowiednią etiologią, a także dla porównania grupy badanej z kontrolną. Rezultaty wyliczeń przedstawia poniższa tabela, w której dla wyróżnionych parametrów, w postaci numeru odpowiadającego odpowiedniemu pierścieniowi sondy, odrzucono hipotezę zerową o równości mediany dla porównywanych przypadków.

Tabela 9.

Porównywane grupy		N ważnych		śr RMS M	śr RMS R	śr MF M	śr MF R
zdrowi	chorzy	49	49	1, 2, 3	-	2, 3	-
zdrowi	idiopatyczne	49	24	3	1, 2	3	-
zdrowi	Neurogenne	49	19	1, 2, 3	-	2, 3	-
Zdrowi	uszkodzenie	49	6	3	-	2, 3	2, 3

Z wyników przedstawionych w tabeli widać, że analogiczna sytuacja występuje dla parametrów śrRMS i śrMF w stanie skurczu maksymalnego dla porównania grup chorzy-zdrowi oraz zdrowi-chorzy (neurogenne). Dla tych grup dokładnie na tych samych poziomach występują statystycznie istotne różnice.

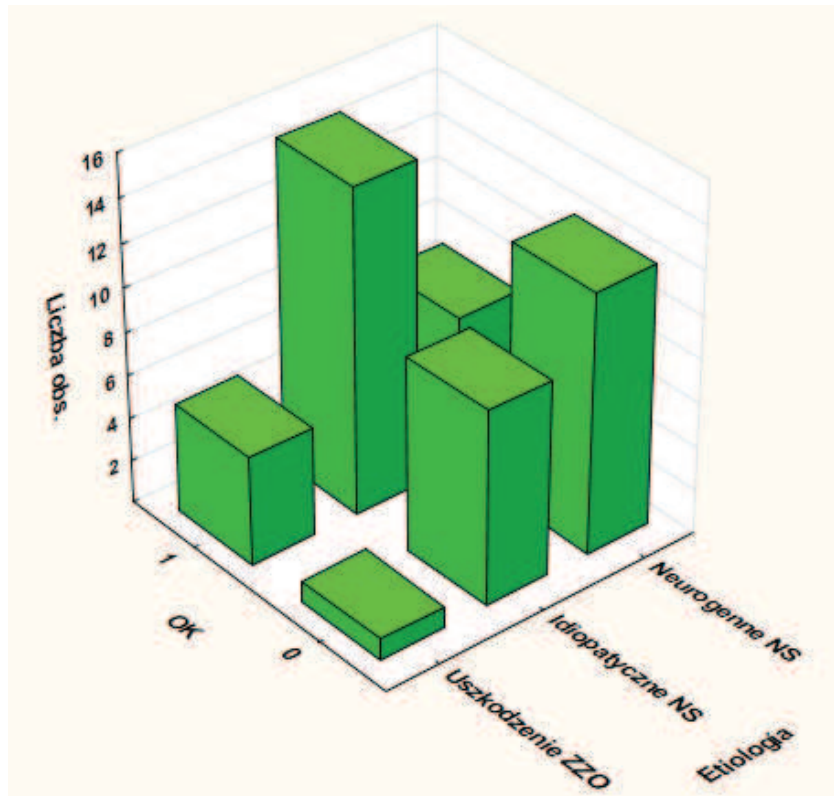
Analiza odruchów odbytowych

Analiza wcześniej opisanych odruchów zostanie przeprowadzona z użyciem testu χ^2 dla tablic wielodzzielczych [7]. Poniżej została przedstawiona Tabela 10 licznosci dla odruchu kaszlowego, w której zamieszczono ilości prawidłowych odruchów-1 i braku odruchu-0 w poszczególnych etiologiach.

Tabela 10.

Ok.	Uszkodzenie ZZO	Idiopatyczne NS	Neurogenne NS	Razem
0	1	9	12	22
1	5	15	7	27
Ogół	6	24	19	49

Zależności obecności odruchów, w odniesieniu to etiologii zostały poniżej zobrazowane, jako rozkład dwuwymiarowy.



Rys. 6. Histogram dwóch zmiennych OK względem Etiologia.

Dla odruchu parcia (OP), gdzie kodowanie obejmuje trzy wartości: 0- odruch obecny, 1-obecny odruch patologiczny, 2-obecny odruch prawidłowy, tabela licznosci oraz wykres dwuwymiarowy przedstawiają się następująco:

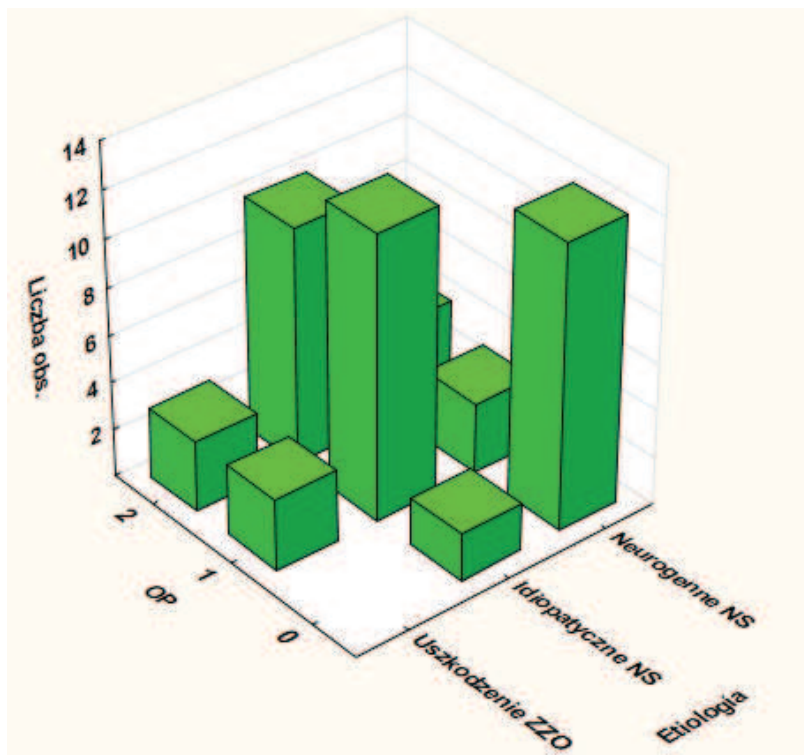
Tabela 11.

OP	Uszkodzenie ZZO	Idiopatyczne NS	Neurogenne NS	Razem
0	0	2	12	14
1	3	12	3	18
2	3	10	4	17
Ogół	6	24	19	49

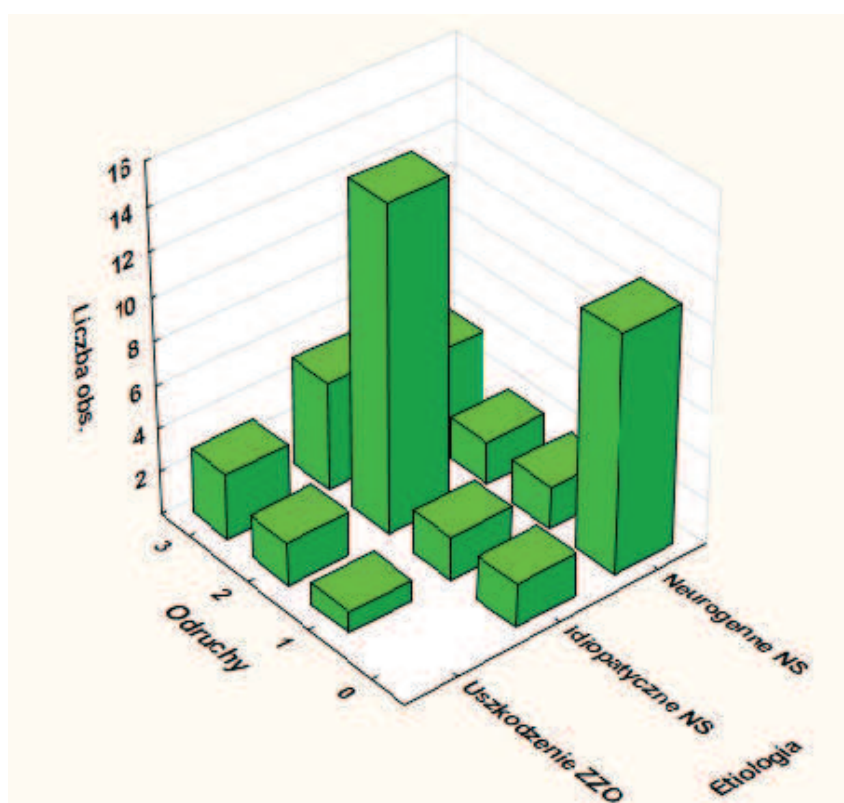
Dla sumarycznego efektu, biorąc pod uwagę odruchy OP i OK razem:

Tabela 12.

OP	Uszkodzenie ZZO	Idiopatyczne NS	Neurogenne NS	Razem
0	0	2	11	13
1	1	2	2	5
2	2	15	2	19
3	3	5	4	12
Ogół	6	24	19	49



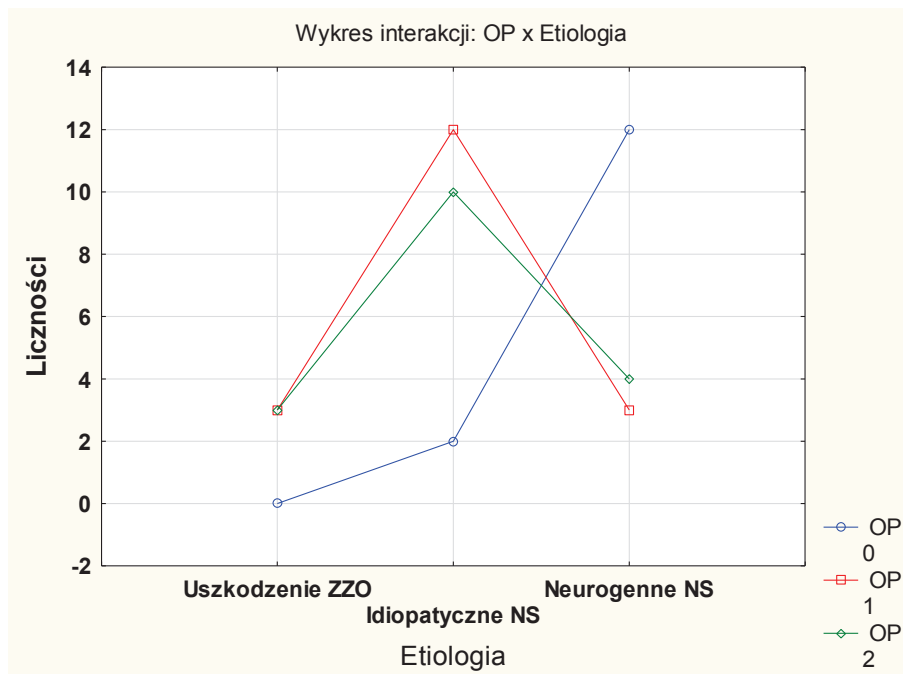
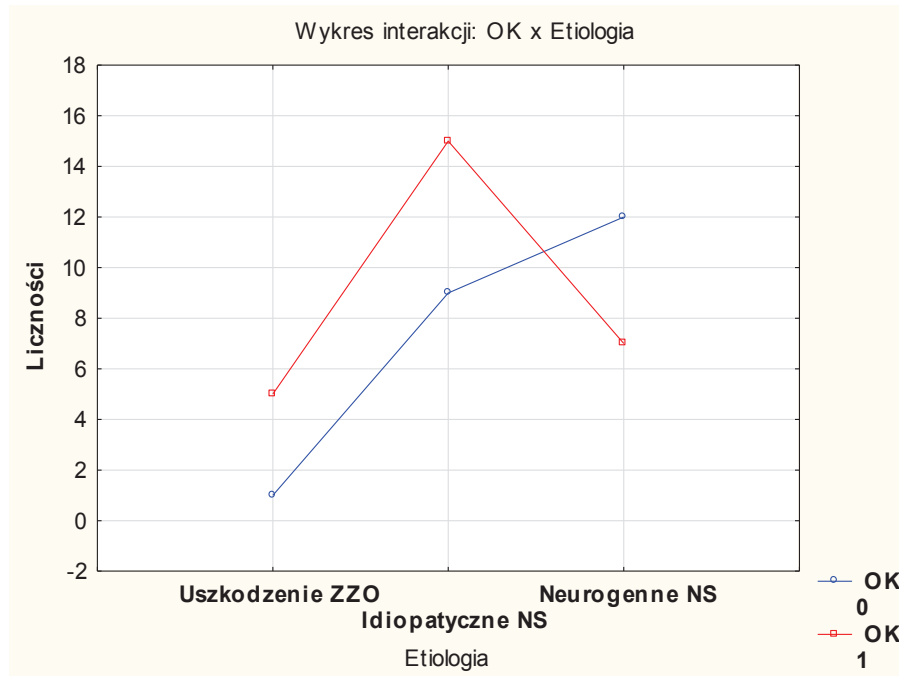
Rys. 7. Histogram dwóch zmiennych OP względem Etiologia



Rys. 8. Histogram dwóch zmiennych. Odruchy względem Etiologia.

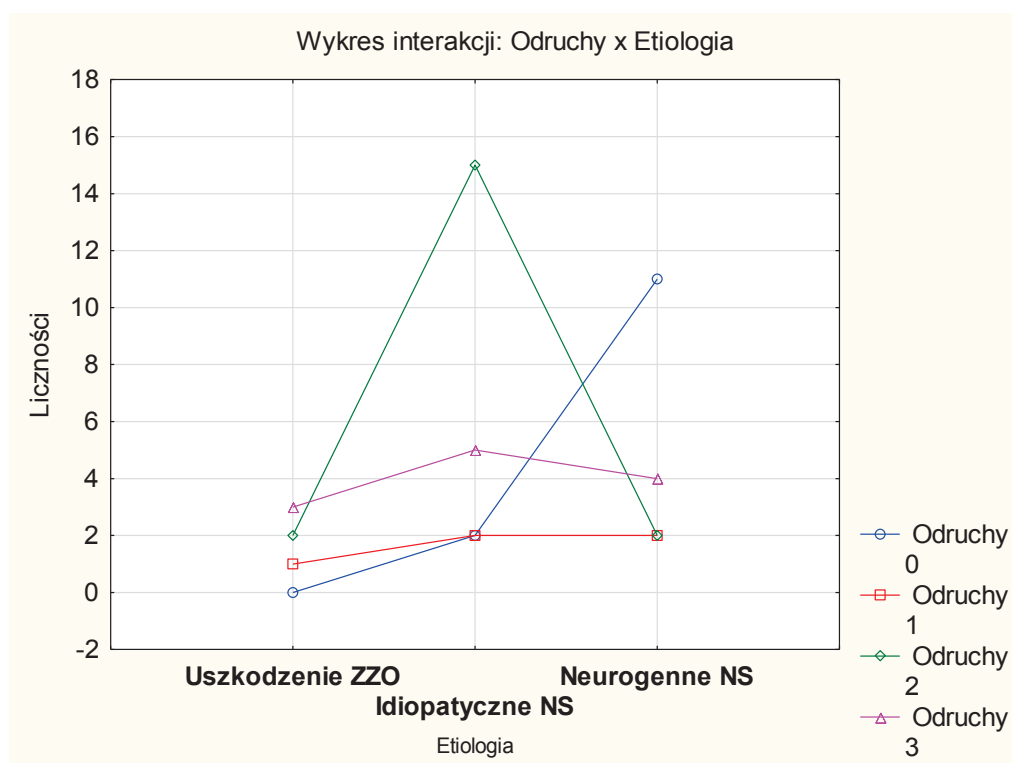
Przedstawione zostały także wykresy interakcji [7]. O efekcie interakcji możemy mówić wtedy, gdy łączny wpływ czynników powoduje skutek, którego nie można by

było przewidzieć na podstawie sumowania efektów czynników z osobna. Badaną interakcją jest zależność pomiędzy występowaniem odruchów (w tym także prawidłowych lub nie), a etiologią schorzenia. Wykresy zostały przedstawione osobno dla OP i OK oraz dla przypadku gdzie badamy efekt sumaryczny odruchów.



Rys. 9. Wykresy interakcji dla OK i OP.

Poniżej został przedstawiony efekt sumaryczny.



Rys. 10. Odruchy – efekt sumaryczny.

Przy braku interakcji krzywe, opisujące wpływ jednego czynnika przy danym poziomie drugiego czynnika, są równoległe, z dopuszczalnymi niewielkimi odchyleniami. Wystąpienie interakcji prezentuje natomiast przecięcie się łamanych. Dla odruchu OK brak interakcji pomiędzy etiologią a wystąpieniem prawidłowego odruchu na poziomie uszkodzenia zwieracza oraz dla grupy o niezdiagnozowanym pochodzeniu NS, natomiast krzywe przecinają się, zbliżając do grupy o poziomie etiologii neurogennej NS. W przypadku OP brak interakcji dla idiopatyczne NS, dla efektu sumarycznego interakcja na każdym poziomie etiologii, jednak dla odpowiednich odruchów (brak punktu przecięcia wszystkich krzywych). Poniżej przedstawiono tabelę wyników odpowiednich statystyk testowych dla odruchów OP i OK branych pod uwagę łącznie.

Tabela 13.

Statystyka	χ^2	df	P
χ^2 Pearsona	21.21901	6	0.00168
χ^2 NW	22.78706	6	0.00087
Wsp. Kontyngencji	0.5497120		
V Cramera	0.4653177		
R. Spearmana	-0.438270		0.00163

Statystyka χ^2 Pearsona i NW (największej wiarygodności) wykorzystuje liczebności oczekiwane obliczone dla tabeli dwudzielczej. Test χ^2 NW oparty jest o teorię największej wiarygodności. Pozostałe, zamieszczone w tabeli, współczynniki (kontyn-

gencji, V Cramera, rang Spearmana) są miarami zależności i ich wartość mówi o sile korelacji. Wartość p, wyliczona dla statystyk χ^2 Pearsona oraz NW, pozwala na przyjęcie hipotezy o istnieniu zależności odruchów w poszczególnych grupach etiologicznych.

Korelacje

Ponieważ dla badanych zmiennych założenia o normalności rozkładu nie zostały spełnione, istniała potrzeba stosowania nieparametrycznych odpowiedników współczynników korelacji. Nieparametrycznym odpowiednikiem współczynnika korelacji Pearsona jest współczynnik korelacji rangowej Spearmana (6) (7), który opisany jest zależnością:

$$r_s = \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)} \quad (6)$$

gdzie: d_i to różnica rang dla cechy X i Y

Interpretacja współczynnika dotyczy, tak jak dla Pearsona, jego wartości i znaku. Dla większych wartości dodatnich lub ujemnych zależność pomiędzy zmiennymi jest silniejsza, znak dotyczy zależności rosnącej (+) lub malejącej(-). Gdy współczynnik jest bliski zeru należy wykluczyć zależność monotoniczną, spodziewając się np. sinusoidalnej. Analiza współczynników korelacji została przeprowadzona pomiędzy odpowiadającymi sobie pierścieniami dla stanu skurczu i spoczynku, dla grupy chorych i zdrowych łącznie, a także oddzielnie, oraz z podziałem na poszczególne etiologie schorzenia. Współczynnik został wyliczony również pomiędzy parametrami amplitudy i częstotliwości, badając czy istnieje korelacja pomiędzy tymi dwoma zmiennymi. Dla grupy badanej i kontrolnej łącznie, biorąc pod uwagę stan napięcia mięśnia, czyli korelacje pomiędzy skurczem a spoczynkiem, w przypadku odpowiadających sobie parametrów, wyniki prezentują się następująco:

Tabela 14.

śrRMS	r	błąd r	statystyka t	P
1M-1R	0.6952	0.07336	9.4769	<0.000001
2M-2R	0.4121	0.09229	4.4312	0.000025
3M-3R	0.4446	0.09142	4.8628	0.000005
Śr MF				
1M-1R	0.6071	0.08110	7.4855	<0.000001
2M-2R	0.6741	0.07539	8.9411	<0.000001
3M-3R	0.6650	0.07623	8.7233	<0.000001

Wyniki dla grupy badanej:

Tabela 15. Wyniki dla grupy badanej.

śrRMS	r	błąd r	statystyka t	P
1M-1R	0.9689	0.0361	23.8331	<0.000001
2M-2R	0.9524	0.4444	21.4299	<0.000001
3M-3R	0.8874	0.0672	13.2001	<0.000001
Śr MF				
1M-1R	0.8058	0.08638	9.3291	<0.000001
2M-2R	0.9046	0.06218	14.5482	<0.000001
3M-3R	0.7811	0.09108	8.5767	<0.000001

Tabela 16. Wyniki dla grupy kontrolnej.

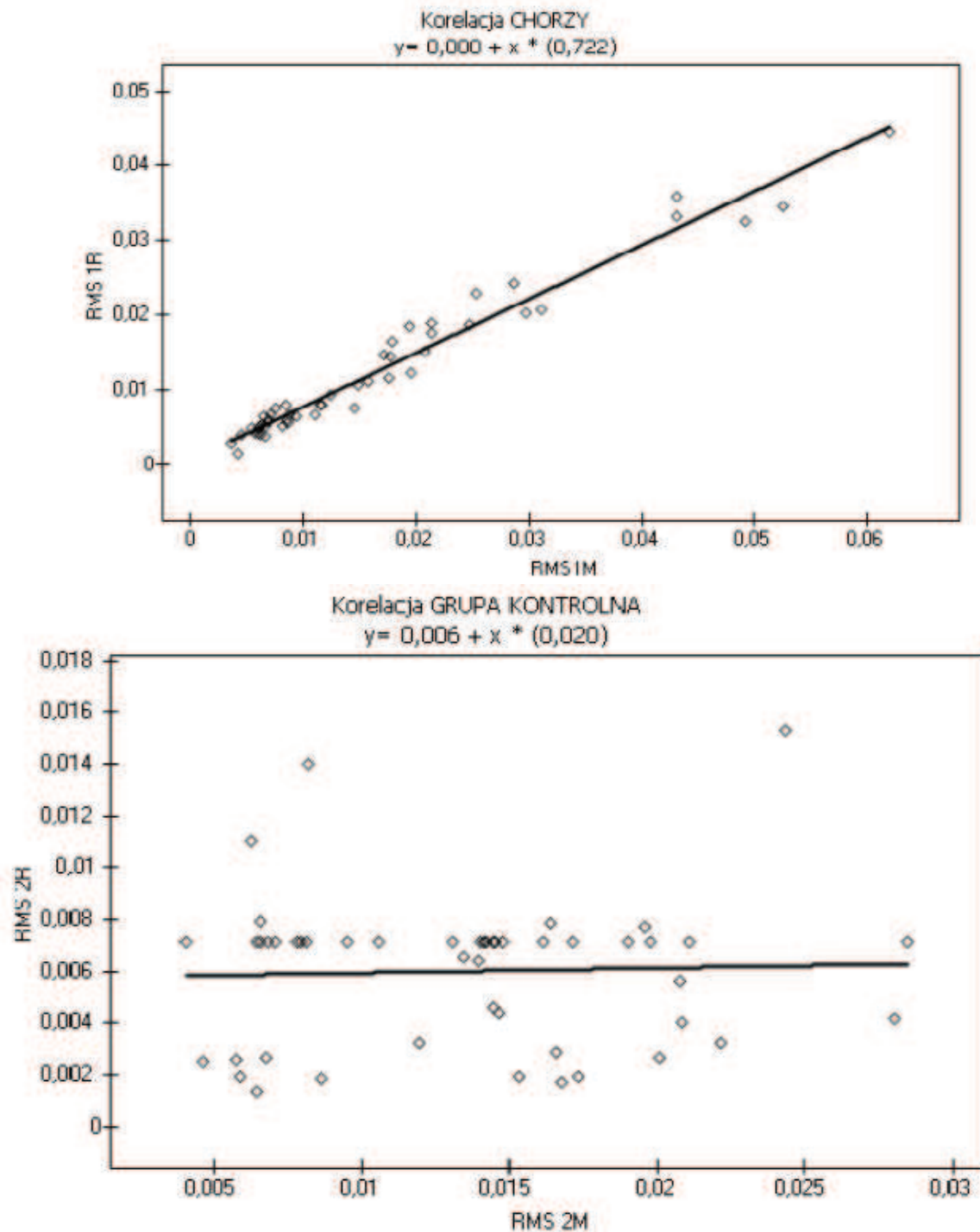
śrRMS	r	błąd r	statystyka t	P
1M-1R	0.53890	0.1229	4.3854	0.000065
2M-2R	0.00603	0.1459	0.0414	0.9672
3M-3R	0.03838	0.1458	0.2633	0.7935
Śr MF				
1M-1R	0.4620	0.1293	3.5712	0.000883
2M-2R	0.4122	0.1329	3.1014	0.003253
3M-3R	0.4450	0.1303	3.4544	0.001179

Przedstawione wyniki wskazują na to, że siła korelacji dla odpowiadających parametrów pomiędzy stanem skurczu i spoczynku w grupie badanej jest znacznie większa niż w grupie kontrolnej (dla wszystkich analizowanych przypadków), co obrazują również wykresy poniżej; po lewej zależność śr RMS 1R-1M dla grupy badanej, natomiast po prawej śr RMS 2R-2M dla grupy kontrolnej.

Taka sama analiza została przeprowadzona w grupie chorych z podziałem na etiologię, aby zbadać, w której grupie istnieje najsilniejsza korelacja.

Tabela 17.

Śr RMS	Idiopatyczne	Neurogenne	Uszkodzenie
1M-1R	0.98087	0.9386	0.9429
2M-2R	0.9478	0.9439	1.0000
3M-3R	0.9052	0.9614	-0.02857
Śr MF			
1M-1R	0.6617	0.8579	0.8258
2M-2R	0.8809	0.9105	0.8286
3M-3R	0.6983	0.8439	0.8286



Rys. 11. Korelacja: grupa badana(po lewej) i kontrolna.

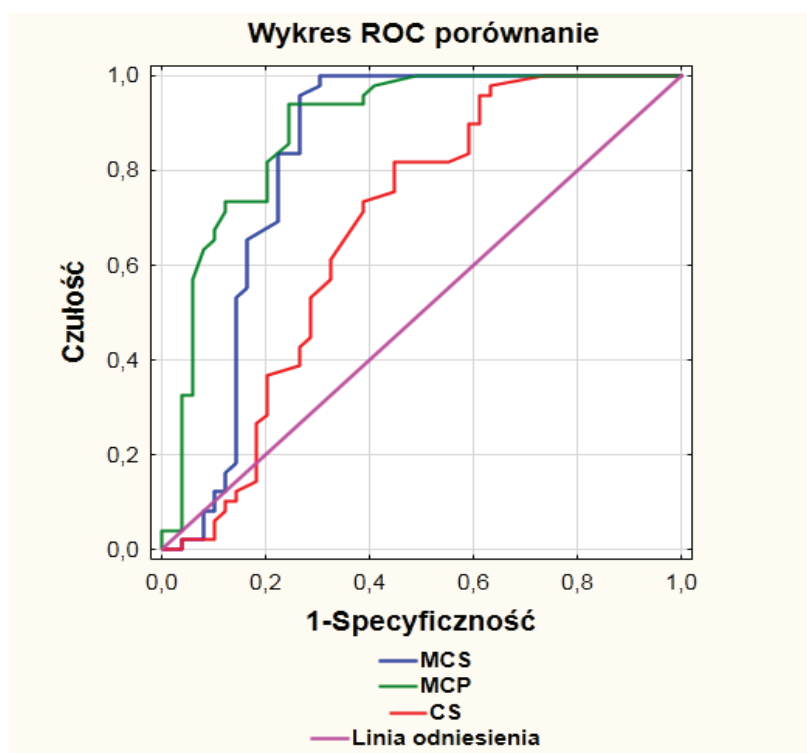
Współczynniki korelacji w porównywanych grupach są zbliżone, nieco silniejsze dla wartości odpowiadających sobie amplitud w porównaniu ze średnią częstotliwością. Odchylenie stanowi (ujemna i bardzo słaba) zależność, występująca dla porównania $\bar{r}RMS$ 3M-3R w grupie z uszkodzeniem ZZO.

Korelacje badano również na poziomie odpowiadających sobie kanałów pomiędzy amplitudą i częstotliwością, ale dla wszystkich porównanych przypadków. Tylko w dwóch wykryto bardzo słabe zależności.

Ocena jakości klasyfikatora z użyciem krzywych ROC

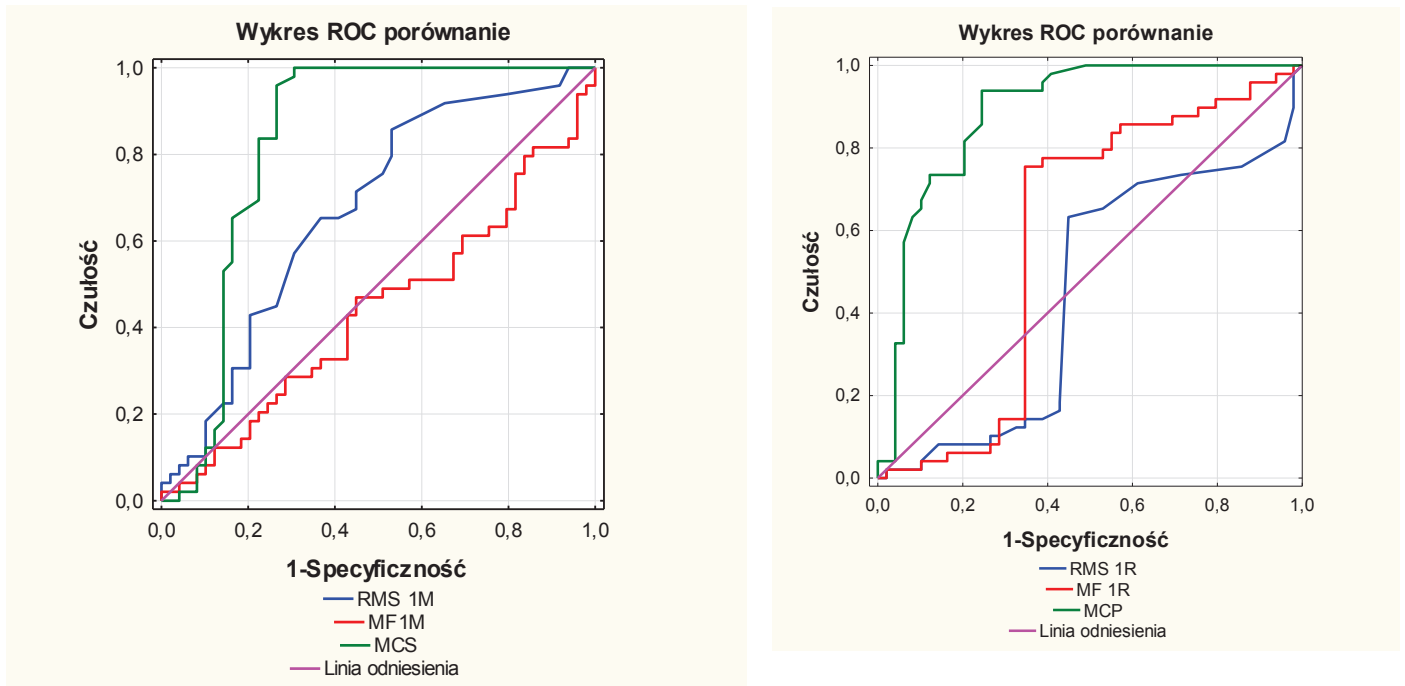
Przeprowadzone dotychczas analizy pozwoliły na zbadanie istnienia efektu (korelacja) lub ocenę jego istotności w przypadku testów nieparametrycznych.

Jednak szczególnie cenne dla analiz medycznych jest badanie i konstruowanie modeli, które klasyfikują przypadki ze względu na zmierzony parametr do grupy: chory lub zdrowy. Krzywa ROC (Receiver Operating Characteristic) służy do oceny poprawności klasyfikatora, zapewniając łączny opis jego czułości i specyficzności. Czułość jest definiowana, jako stosunek ilości przypadków trafnie przypisanych do klasy, do wszystkich przypadków w niej obecnych. Natomiast specyficzność, jako ilość przypadków błędnie przypisanych do klasy do liczby przypadków do niej nienależących. Konstrukcja krzywej opiera się na wyznaczeniu dla każdego punktu czułości i specyficzności i odznaczeniu wartości na wykresie (9). Poniżej został przedstawiony zbiorczy wykres krzywych ROC dla parametrów ciśnienia MCP, MCS i CS (będącego różnicą pomiędzy MCP i MCS).



Rys. 12. Krzywa ROC dla wartości ciśnienia.

Pole powierzchni pod krzywą AUC (Area Under Curve), mieści się w przedziale $[0,1]$. Im jego wartość jest bliższa jedynce tym wybrana zmienna może być stosowana jako klasyfikator z lepszym skutkiem. Optymalnym punktem odcięcia jest punkt, znajdujący się najbliżej punktu o współrzędnych $(0,1)$. Czułość takiego punktu wynosi 1, czyli nie sklasyfikowano błędnie żadnego przypadku (8). Zatem dla przedstawionego po lewej przypadku, najlepszym klasyfikatorem będzie MCP. Poniżej zostały przedstawione krzywe porównujące ze sobą wszystkie parametry RMS, MF, MCP, MSC odpowiednio dla stanów skurczu i spoczynku.



Rys. 1. Krzywe ROC porównujące wszystkie parametry.

Tabela przedstawia wyliczone pola powierzchni pod krzywymi, odpowiadające poszczególnym parametrom:

Tabela 18.

Parametr	AUC	SE
MCS	0.828	0.848
RMS 1M	0.672	0.055
MF 1M	0.439	0.058
MCP	0.887	0.035
RMS 1R	0.450	0.06
MF 1R	0.585	0.063

Z przedstawionych krzywych ROC i wyliczonych pól powierzchni pod tymi krzywymi wynika, że zarówno dla stanu skurczu jak i spoczynku najbardziej trafny model klasyfikacji opiera się na parametrach ciśnienia. To wartości ciśnienia najlepiej przyporządkowują osobę do grupy chory-zdrowy. Dużo trudniej zdecydować w kwestii, który z parametrów pEMG, średnia amplituda czy częstotliwość jest lepszym klasyfikatorem, gdyż dla \overline{srRMS} w stanie skurczu pole powierzchni pod krzywą jest większe niż dla \overline{srMF} , jednak dla stanu spoczynku sytuacja jest odwrotna. Dlatego dodatkowo przeprowadzono statystykę testową, która wskaże czy istnieje statystycznie istotna różnica w modelach generowanych przy pomocy analizowanych zmiennych. Wyniki testu przedstawiono w tabeli 19.

Tabela 19.

Parametry	Z	P
RMS 1M-MF 1M	-2.9601	0.0031
RMS 1R-MF 1R	1.3974	0.1623
RMS 2M-MF 2M	-3.6456	0.0003
RMS 2R-MF 2R	1.1829	0.1548
RMS 3M-MF 3M	-6.2518	0.0000
RMS 3R-MF 3R	-1.2313	0.0792

Z wyników przedstawionych w tabeli wynika, że istotne statystycznie różnice dla parametrów amplitudy i częstotliwości występują tylko w stanie skurczu.

Podsumowanie i wnioski

Przeprowadzone analizy z użyciem testów nieparametrycznych, korelacji, krzywych ROC pozwoliły na ocenę parametrów badania elektromiograficznego i relacji między nimi w grupie z problemem nietrzymania stolca, grupie kontrolnej, a także w poszczególnych etiologiach tego schorzenia. W przypadku tych ostatnich ocena zależności była najtrudniejsza z uwagi na znaczną nierówno liczebność w grupach. Dla odpowiednich testów statystycznych wykazano istotne statystycznie różnice między określonymi parametrami stosując różne kryteria porównawcze. Analiza korelacji wykazała jednoznaczną różnicę w sile korelacji w grupie kontrolnej i badanej. Równolegle dokonano także porównania innych badań wykonywanych podczas diagnostyki NS (manometrii i oceny odruchów odbytowych). Ostatnia analiza z użyciem krzywych ROC pozwoliła na ocenę, który z porównywanych parametrów najlepiej klasyfikuje osobę do grupy chory lub zdrowy, dokonując porównania wyników pEMG i manometrii anorektalnej łącznie. Na poziomie porównania parametrów samej elektromiografii różnice istotne statystycznie w wielu przypadkach istnieją. Nie można jednak wnioskować o ogólnej zależności dla konkretnych zmiennych i poziomach kanału odbytu, ze względu na zbyt małą próbę i przede wszystkim duże różnice we wspomnianych już wyżej liczebnościach poszczególnych grup etiologicznych.

Literatura

- [1] R. Merletti, A. Rainoldi, D. Farina, *Surface electromyography for noninvasive characterization of muscle*, Exerc Sport Sci Rev., 2001, 29, 20.
- [2] J. Sałówka, M. Nowakowski, P. Wałęga, E. N. Kamavuako, D. Farina, A. Korczyński, R. M. Herman, *Wpływ zakresu resekcji odbytnicy na zapis elektromiogramu powierzchniowego zwieracza zewnętrznego odbytu u chorych z rakiem odbytnicy*, Proktologia, 2008, 9, 237.
- [3] C. Cescon, L. Mesin, M. Nowakowski, R. Merletti, *Geometry assessment of anal sphincter muscle based on monopolar multichannel surface EMG signals*, J. Electromyogr. Kinesiol, 2011, 21, 394.
- [4] R. M. Herman, M. Nowakowski, *Chirurgiczne leczenie nietrzymania stolca*, Polish Gastroenterology, 2009, 16, p. 171.

- [5] R. Herman, M. Nowakowski, *Choroby okolicy anorektalnej*. [ed.] J. Kuźdzał J. Szmidt. *Podstawy chirurgii*, Kraków, Wydawnictwo Medycyna Praktyczna, 2010, Vol. 2, p. 829.
- [6] A. Petrie, C. Sabine, *Statystyka Medyczna w Zarysie*, Warszawa, Wydawnictwo Lekarskie PZWL, 2006, 22-27; 49-56.
- [7] A. Stanisław, *Przystępny kurs statystyki z zastosowaniem Statistica Pl na przykładach z medycyny*, Kraków, StatSoft, 2007, 123-127, 202-210, 263-279, Vol.1, Statystyki Podstawowe.
- [8] A. Łomnicki, *Wprowadzenie do statystyki dla przyrodników*, Warszawa, Wydawnictwo naukowe PWN, 2005, 127-136.
- [9] G. Harańczyk, *Krzywe ROC, czyli ocena jakości klasyfikatora i poszukiwanie optymalnego punktu odcięcia*, 2010, 79.