

Andrzej Żarowski^{1,2}, Henryk Skarżyński¹, Erwin F. Offeciers²

¹ Instytut Fizjologii i Patologii Słuchu
Warszawa

² University ORL Department, Medical Institute St. Augustinus
Antwerp

Propozycja ujednolicenia zasad prezentacji wyników funkcjonalnych w otocirurgii

The Rules of Presentation of Functional Results in Otosurgery – the Proposal of Standardization

Słowa kluczowe: otocirurgia, wyniki leczenia.

Key words: otosurgery, results of treatment.

Streszczenie

Autorzy pracy przedstawiają propozycję ujednolicenia zasad prezentacji wyników słuchowych chirurgicznego leczenia różnych schorzeń uszu. Jednolite podejście do opracowywanych danych powinno pozwolić na możliwości porównywania różnych technik operacyjnych, stosowanych materiałów do rekonstrukcji aparatu przewodzącego itd. Przedstawiona propozycja ujednolicenia zasad prezentacji wyników funkcjonalnych uwzględnia obecnie stosowane w praktyce zalecenia Amerykańskiej Akademii Otolaryngologii Chirurgii Głowy i Szyi oraz nasze własne przemyślenia.

Summary

Authors present unified rules for presentation of hearing results of surgical treatment of various ear pathologies. Standardized approach to these data should allow comparing different surgical techniques, materials used for reconstruction of conductive apparatus etc. Our proposal of standardization the rules of outcomes presentation takes in the account currently used in practice recommendations of American Academy of Otolaryngology Head and Neck Surgery and our own considerations.

Rezultaty badań klinicznych oraz wyniki leczenia publikowane w fachowej literaturze stanowią podstawowe źródło informacji dla lekarzy oraz umożliwiają im ocenę i porównanie poszczególnych metod leczniczych pomiędzy sobą. Jednym z podstawowych problemów jest jednakże fakt stosowania w publikacjach różnych sposobów prezentacji i opracowania statystycznego danych. Powoduje to trudności w interpretacji wyników i czyni przedstawiane metody lecznicze niemożliwymi do porównania. Jedynie w efekcie różnic w sposobie opracowywania statystycznego to samo badanie może dać istotnie różne wyniki (por. przykład 1).

I. CEL PRACY

Celem pracy jest standaryzacja metod prezentacji i opracowywania statystycznego technicznych wyników otocirurgicznych w aspekcie funkcjonalnym. Standaryzacja ta jest konieczna dla:

- zapewnienia powszechnej poprawności statystycznej,
- wymuszenia dyscypliny i staranności w zbieraniu i prezentowaniu danych,
- umożliwienia wykonywania metaanaliz i porównywania między sobą wyników stosowania różnych technik otocirurgicznych.

II. STAN OBECNY

W 1995 r. Komitet ds. Słuchu i Równowagi Amerykańskiej Akademii Otolaryngologii, Chirurgii Głowy i Szyi podał jako pierwszy na świecie spójny zestaw prostych zaleceń dotyczących sposobu prezentacji i metod opracowywania statystycznego danych audiologicznych w otocirurgii. Poziom I zaleceń dotyczy jednolitych metod prezentowania danych sumarycznych, poziom II – nie przetworzonych danych pomiarowych wszystkich pacjentów.

W zakresie poziomu I (dane sumaryczne) podane zostały następujące grupy zaleceń:

1. Każda publikacja otocirurgiczna musi zawierać minimum danych dotyczących stanu wyjściowego łańcucha kosteczek, w szczególności stanu strzemiączka (rekonstrukcja do główki strzemiączka, do podstawy czy do otwartego okienka owalnego) oraz młoteczka (rękojeść obecna lub jej brak). W chirurgii przewlekłych zapaleń konieczne jest podanie oceny upowietrznienia ucha środkowego i stanu błony śluzowej. Publikacje muszą zawierać również dane dotyczące wskazań do operacji, decyzji dotyczących rozłożenia rekonstrukcji funkcjonalnych

na etapy, liczby rewizji oraz liczby występujących komplikacji (np. ekstruzji protezek).

2. Do prezentacji danych funkcjonalnych zalecono posługiwanie się średnią wartością progów słyszenia na częstotliwościach 0.5, 1, 2 i 3 kHz w zaokrągleniu do najbliższej liczby całkowitej. Częstotliwość 3 kHz wybrano ze względu na jej znaczenie dla rozumienia mowy (górna część pasma mowy). Pomiary audiometryczne powinny być wykonywane dla oktaowych częstotliwości 0.5-8 kHz wraz z 3 kHz dla przewodnictwa powietrznego oraz 0.5-4 kHz wraz z 3 kHz dla przewodnictwa kostnego.

3. Komitet ds. Słuchu i Równowagi zalecił wyliczanie wartości rezerwy ślimakowej na podstawie danych zebranych w tym samym czasie, tj. przedoperacyjne progi przewodnictwa powietrznego względem przedoperacyjnych progów przewodnictwa kostnego oraz pooperacyjne progi przewodnictwa powietrznego względem pooperacyjnych progów przewodnictwa kostnego. Pooperacyjna zmiana wartości przewodnictwa kostnego (poprawa będąca efektem zjawiska Carharta lub pogorszenie wskutek urazu śródoperacyjnego) powinna być podawana jako różnica średniej wartości przedoperacyjnych progów przewodnictwa kostnego na częstotliwościach 1, 2 i 4 kHz minus średnie wartości progów pooperacyjnych. Stopień pooperacyjnego zmniejszenia rezerwy ślimakowej wyliczać należy jako różnicę pomiędzy przed- i pooperacyjną rezerwą ślimakową.

4. Dopuszczalne jest podawanie w publikacjach wyników audiometrii słownej w wolnym polu; zaleca się w tym przypadku podawanie indeksu dyskryminacji dla ujednoliconego poziomu natężenia równego 60 dB.

5. Publikacje powinny zawierać wyliczone wartości średnie, odchylenia standardowe oraz zakresy zmienności dla: (1) pooperacyjnej rezerwy ślimakowej (tu dla wygody autorów dopuszcza się podawanie wyników w przedziałach 0-10 dB, 11-20 dB, 21-30 dB oraz >30 dB), (2) stopnia pooperacyjnego zmniejszenia rezerwy ślimakowej oraz (3) pooperacyjnej zmiany wartości przewodnictwa kostnego. Komitet zachęca również do podawania (4) sumarycznego opisu przed- i pooperacyjnych wartości progów słyszenia powietrznego. Dla (1) i (2) rezultaty podawane być mogą po minimum roku obserwacji, dla (3) po minimum 6 tygodniach. Jedynie w analizach porównujących czysto techniczne aspekty rekonstrukcji (np. dwa różne typy protezek) można podawać dla (1) i (2) wyniki wczesne.

W zakresie poziomu II podane zostały następujące zalecenia:

1. W celu umożliwienia precyzyjnej obróbki statystycznej danych w metaanalizach Komitet zalecił podawanie przez autorów pełnych, nie przetworzonych danych wszystkich pacjentów. Ze względów praktycznych dane te nie muszą być konieczne publikowane, powinny być jednak dostępne u wydawcy.

2. Dane nie przetworzone powinny obejmować przed- i pooperacyjne wartości progów słyszenia na częstotliwościach 0.5-8 kHz wraz z 3 kHz dla przewodnictwa powietrznego oraz 0.5-4 kHz wraz z 3 kHz dla przewodnictwa kostnego.

III. NASZE STANOWISKO

W wyniku wielu dyskusji, m.in. na forum Europejskiej Akademii Otologii i Neuro-Otologii, autorzy, akceptując ogólnie zalecenia Akademii Amerykańskiej, przedstawiają własne uwagi dotyczące metod prezentacji funkcjonalnych wyników technicznych w otolaryngologii:

1. Uwzględnienie częstotliwości 3 kHz jest niepraktyczne ze względu na zakorzenione przyzwyczajenia do pomiarów audiometrycznych jedynie w zakresie częstotliwości oktaowych. W przypadku braku wartości pomiarowej dla 3 kHz proponowane przez Komitet ds. Słuchu i Równowagi Akademii Amerykańskiej uwzględnienie średniej z 2 i 4 kHz powoduje wprowadzenie do statystyki zmiennej zależnej. Dlatego też autorzy proponują uwzględnienie średniej z częstotliwości 0.5, 1, 2 i 4 kHz.

2. Zalecane przez Komitet ds. Słuchu i Równowagi używanie w analizie danych audiometrycznych statystyk parametrycznych, uwzględniających wartości średnie i odchylenia standardowe, wydaje się niekorzystne. Statystyki parametryczne opierają się na założeniach dotyczących rozkładu zmiennej w populacji i z reguły wymagają, aby rozkład ten był rozkładem normalnym. Dane audiometryczne nie zawsze mają rozkład normalny (szczególnie w niewielkich próbach) i trudno jest je traktować jako normalne bez wykonania odpowiednich testów statystycznych (por. przykład 2). Jednocześnie statystyki nieparametryczne, nie czyniąc żadnych założeń co do rozkładu zmiennej w populacji, w niczym nie ustępują statystykom parametrycznym w zakresie testowania hipotez i umożliwiają wykonanie większości analiz (por. przykład 3). Dlatego też autorzy proponują stosowanie statystyk nieparametrycznych do analizy wyników funkcjonalnych w otolaryngologii.

3. Wartości średnie i odchylenia standardowe są w niewielkich próbach również bardzo wrażliwe na zmienność w zakresie wartości ekstremalnych. Wystąpienie choćby u jednego pacjenta pooperacyjnej głuchoty może zaważyć na ocenie technicznego rezultatu całej metody poprzez zaniżenie średniego wyniku (por. przykład 4). W tym aspekcie istotny jest również brak zgodności co do sposobu kodowania wartości słuchu spoza zakresu danego audiometru – kodowanie jako 100 dB ma inny wpływ na wynik średni niż np. 120 dB. Jest to kolejny argu-

ment przeciwko posługiwaniu się średnimi arytmetycznymi (i innymi słabymi parametrami statystycznymi) w analizie danych audiologicznych.

4. Autorzy, proponując stosowanie w analizie wyników otolaryngicznych statystyk nieparametrycznych (*rang*), charakteryzując populację za pomocą pięciu wartości: mediany, górnego i dolnego kwartyla oraz górnego i dolnego zakresu. Graficzną prezentacją powyższych parametrów może być wykres typu „prostokąt z ramionami” (*box and whisker plot*), pozwalający na precyzyjne przedstawienie populacji o niekoniecznie normalnym rozkładzie i doskonale uwidaczniający asymetrię rozkładu (por. przykład 5). Na wykresie takiego typu prostokąt przedstawia rozproszenie środkowych 50% danych, odcinek wewnątrz prostokąta jest wartością mediany, a ramiona obejmują wartości ekstremalne (lub też np. 2.5 oraz 97.5 percentyl). Jeżeli w populacji istnieją wartości znacznie odbiegające od mediany (*outliers*), to zostają zaznaczone jako osobne punkty poza ramionami.

5. Kilka wykresów typu „prostokąt z ramionami” może zostać przedstawionych na jednym rysunku, tak że możliwe jest uwzględnienie wszystkich parametrów zalecanych przez Komitet ds. Słuchu i Równowagi Akademii Amerykańskiej (parametry te zaznaczono pogrubioną czcionką). Możliwe jest jednak uwzględnianie na wykresie również i innych dodatkowych wielkości (por. przykład 6):

- przedoperacyjny poziom przewodnictwa powietrznego (średnia z częstotliwości 0.5, 1, 2 i 4 kHz),
- przedoperacyjny poziom przewodnictwa kostnego (średnia z częstotliwości 0.5, 1, 2 i 4 kHz),
- wielkość przedoperacyjnej rezerwy ślimakowej,
- pooperacyjny poziom przewodnictwa powietrznego (średnia z częstotliwości 0.5, 1, 2 i 4 kHz),
- pooperacyjny poziom przewodnictwa kostnego (średnia z częstotliwości 0.5, 1, 2 i 4 kHz),
- wielkość pooperacyjnej rezerwy ślimakowej,
- zmiana poziomu przewodnictwa kostnego (średnia z częstotliwości 1, 2 i 4 kHz przedoperacyjna minus pooperacyjna),
- stopień zmniejszenia rezerwy ślimakowej (rezerwa przedoperacyjna minus pooperacyjna).

IV. DODATKOWE UWAGI

1. Autorzy uważają, że prezentowanie wyników w poszczególnych przedziałach nie powinno być podstawową metodą prezentacji ze względu na słabość tego typu statystyki (por. przykład 7).

2. W analizach statystycznych należy unikać przypadków zależnych (np. dwoje uszu lub ponowna operacja u tego samego pacjenta).

3. Wyniki audiometryczne powinny być przedstawiane na wykresach w taki sposób, aby na osi Y przedział uwzględnianych wartości progów słyszenia był w całej publikacji jednakowy. Stosowanie na kolejnych wykresach różnych skal na osi Y prowadzi bowiem do optycznego zafałszowania wyników.

PRZYKŁADY

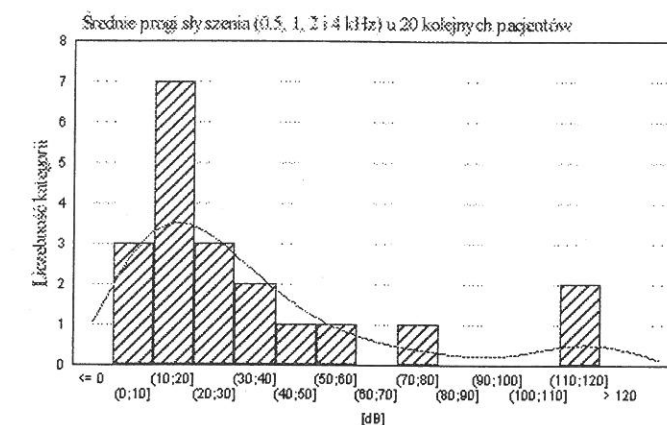
Przykład 1

Analiza 2521 przypadków otosklerozy operowanych w Uniwersyteckiej Klinice Otolaryngologii Instytutu Św. Augustyna w Antwerpii w latach 1961-1991 wykazała zamknięcie rezerwy ślimakowej na częstotliwościach 0.5, 1 i 2 kHz (*Fletcher index*) w granicach 0-10 dB w 81% przypadków. Jeżeli uwzględniona zostałaby częstotliwość 4 kHz, to zamknięcie rezerwy ślimakowej w granicach 0-10 dB wystąpiłoby w 90% przypadków.

Przykład 2

Poniższy histogram przedstawia wyniki audiometrii tonalnej (średnia z częstotliwości 0.5, 1, 2 i 4 kHz) kolejnych 20 pacjentów testowanych w ciągu jednego dnia w Uniwersyteckiej Klinice Otolaryngologii Instytutu Św. Augustyna w Antwerpii. Linia interpolacyjna wskazuje na rozkład, który jest asymetryczny i dwumodalny, a więc daleki od normalnego. Zostało to również potwierdzone w teście W. Shapiro-Wilk, gdzie wyliczone prawdopodobieństwo, że rozkład uzyskany w próbie może pochodzić z populacji o rozkładzie normalnym, było mniejsze od 0.0001.

(seria wyników: 3, 10, 118, 20, 36, 33, 46, 12, 15, 13, 22, 29, 120, 78, 26, 14, 56, 6, 18, 11 [dB])



Przykład 3

Poniższa tabela podaje nieparametryczne odpowiedniki klasycznych testów parametrycznych stosowanych w typowych przypadkach weryfikacji hipotez statystycznych:

TYP ANALIZY STATYSTYCZNEJ	TEST PARAMETRYCZNY	TEST NIEPARAMETRYCZNY
Jedna grupa obserwacji (porównanie średniej z próby z ustaloną wartością)	test t dla jednej próby	- test znaków - test rang Wilcozona
Dwie grupy obserwacji parowanych (porównanie średnich z obu prób)	test t dla dwóch prób sparowanych	test rang Wilcozona dla prób sparowanych
Dwie grupy niezależnych obserwacji (porównanie średnich z obu prób)	- test t dla dwóch prób niezależnych (przy założeniu jednakowych wariancji w próbach) - test Welcha (przy założeniu niejednakowych wariancji w próbach)	test Mann-Whitneya
Kilka grup niezależnych obserwacji	ANOVA	ANOVA Kruskal-Wallis

Przykład 4

Weźmy dla przykładu dwie serie danych audiometrycznych:

seria I: 10, 10, 10, 10, 20 [dB]
seria II: 10, 10, 10, 10, 120 [dB]

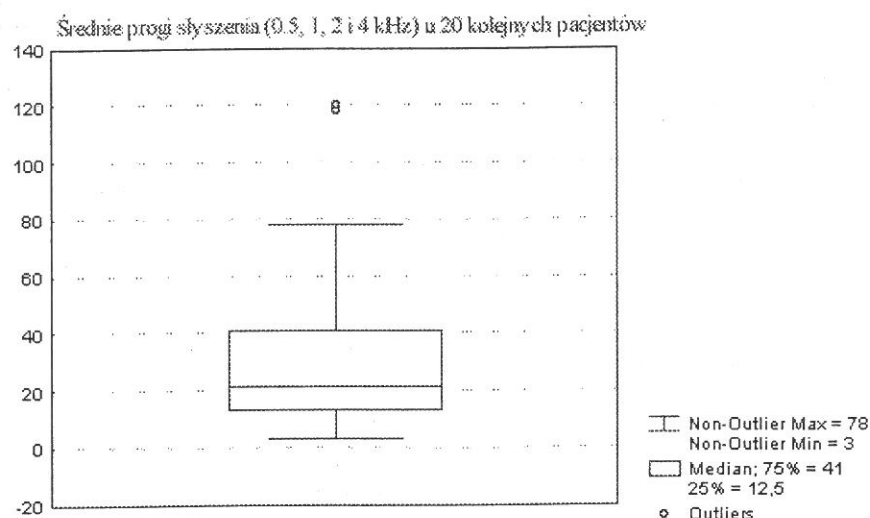
średnia = 12 dB
średnia = 32 dB

mediana = 10 dB
mediana = 10 dB

Na przykładzie tym widać, jak bardzo klasyczne statystyki sumacyjne (średnia) wrażliwe są na zmienność w zakresie wartości skrajnych. Pojedynczy przypadek pooperacyjnej głuchoty bądź też zwykły błąd przy wprowadzaniu danych mogą w zasadniczy sposób wpłynąć na wynik analizy. Słabość ta nie występuje przy korzystaniu z mediany.

Przykład 5

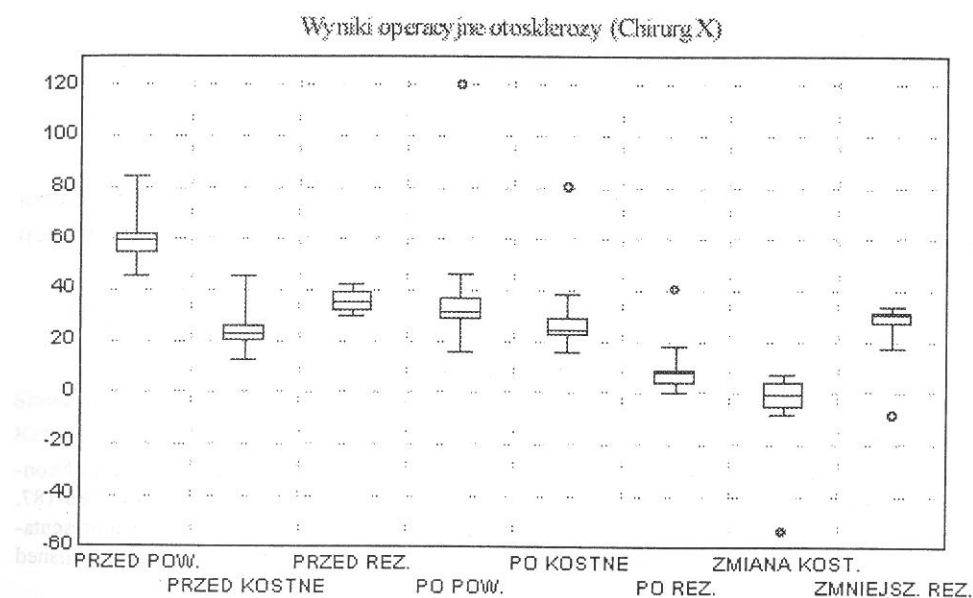
Weźmy dane audiometryczne z przykładu 2. Opis tej populacji za pomocą klasycznych parametrów sumacyjnych: wartości średniej = 34.3 dB oraz odchylenia standardowego = 34.2 dB indukuje założenie normalności rozkładu i powoduje dalekie od prawdziwości założenie o rozkładzie cechy w populacji generalnej. Znacznie bliższa prawdy jest charakterystyka rozkładu za pomocą mediany = 21 dB oraz wartości dolnego kwartyla = 12.5 dB i górnego kwartyla = 41 dB. Graficzna reprezentacja za pomocą wykresu typu „prostokąt z ramionami” (*box and whisker plot*) przedstawiona jest poniżej.



Wykres ten pozwala na precyzyjne przedstawienie rozkładu niesymetrycznego zawierającego dodatkowo wartości znacznie odbiegające od wartości środkowej. Na wykresie tego typu prostokąt przedstawia rozproszenie środkowych 50% danych, odcinek wewnątrz prostokąta jest wartością mediany, a ramiona obejmują wartości ekstremalne (lub też np. 2.5 oraz 97.5 percentyl). Jeżeli w populacji istnieją wartości znacznie odbiegające od mediany (*outliers*), to zostają zaznaczone jako osobne punkty poza ramionami.

Przykład 6

Poniższy graf przedstawia hipotetyczne wyniki operacyjne otosklerozy z uwzględnieniem wszystkich zmiennych proponowanych przez autorów. Wykresy typu „prostokąt z ramionami” dla uwzględnionych zmiennych (średnie progi przewodnictwa dla odpowiednich częstotliwości) przedstawiono obok siebie.



Analiza wykresów pozwala na proste odczytanie stanu przedoperacyjnego, jak i wyników operacyjnych. Przed operacją mediana przewodnictwa powietrznego wynosiła 59 dB, przewodnictwa kostnego 22 dB, a rezerwy ślimakowej 35 dB. Widoczna była również asymetria rozkładu wyników audiometrycznych. W wyniku operacji przewodnictwo powietrzne uległo poprawie do 31 dB, a przewodnictwo kostne praktycznie się nie zmieniło (poprawa o 0.5 dB). Rezerwa ślimakowa zmniejszyła się średnio o 30 dB. Pojedynczy przypadek pooperacyjnej głuchoty jest wyraźnie widoczny, jednakże nie wpływa on na ocenę techniczną rekonstrukcji.

Przykład 7

Weźmy dla przykładu dwie serie danych:

seria I: 1, 1, 2, 5, 6, 8, 8, 9 średnia = 5,0
seria II: 1, 1, 2, 6, 6, 8, 8, 9 średnia = 5,1

i przedstawmy je w poszczególnych kategoriach:

	Kategorie 1-5	Kategorie 6-10
seria I:	50%	50%
seria II:	38%	62%

Na przykładzie tym widać słabość wynikającą z kategoryzacji danych. Nawet minimalna różnica pomiędzy seriami może dać istotne różnice przy prezentacji w postaci kategorii.

Bibliografia

- Altman D. G. [ed.] (1991): Chapman and Hall. „Practical Statistics for Medical Research”.
- Committee on Hearing and Equilibrium guidelines for the evaluation of results of treatment of conductive hearing loss (1995). „Otolaryngology – Head and Neck Surgery” 113 (3), 186-187.
- Govaerts P., Somers Th., Offeciers F. E.: The use of box and whisker plots for the graphical presentation of the audiometrical results of treatment of conductive hearing loss. Unpublished information.
- Somers Th., Govaerts P., Marquet Th., Offeciers F. E. (1994): Statistical analysis of otosclerosis surgery performed by Jean Marquet. „Annals of Otology, Rhinology & Laryngology” 103 (12), 945-951.