

Uniwersytet Jagielloński

Collegium Medicum

Wydział Lekarski

Joanna Szaleniec

**Wykorzystanie sztucznych sieci neuronowych  
w prognozowaniu efektów leczenia operacyjnego chorych  
z przewlekłym zapaleniem zatok przynosowych**

*Praca doktorska*

Promotor: Prof. dr hab. n. med. Paweł Strępek

Pracę wykonano w Katedrze i Klinice Otolaryngologii UJ CM

Kierownik jednostki: Prof. dr hab. n. med. Jacek Składzień

Kraków, 2014

*Pracę tę dedykuję moim córkom,  
dziewięcioletniej Beatce i czteromiesięcznej Justynce,  
za wyrozumiałość dla pracującej mamy*  
Joanna Szaleniec

# Podziękowania

Pragnę najserdeczniej podziękować mojemu Promotorowi, Panu Profesorowi Pawłowi Strękowski, za jego życzliwość dla młodego pokolenia lekarzy otolaryngologów i pomoc w prowadzonych przeze mnie badaniach naukowych. Gdyby nie nowoczesne spojrzenie Pana Profesora na problematykę leczenia przewlekłego zapalenia zatok przynosowych połączone z otwartością na nowe idee z pogranicza medycyny i informatyki, ta praca nie mogłaby powstać.

Serdeczne podziękowania składam także Panu Profesorowi Jackowi Składzieniowi za nieustanne wsparcie w mojej codziennej pracy klinicznej, a przede wszystkim za zaufanie dla możliwości zastosowania modeli informatycznych w medycynie, które stało się impulsem do rozpoczęcia badań będących przedmiotem mojej pracy doktorskiej.

Dziękuję również mojej Rodzinie za ogromny wysiłek organizacyjny, jakiego wymagało pogodzenie mojej pracy naukowej z opieką nad dwojgiem dzieci, z których jedno przyszło na świat już w czasie pisania tej pracy.

Szczególnie gorące podziękowania należą się Pani Pielęgniarce Ewie Kmiecik za wieloletnią bezinteresowną pomoc w gromadzeniu danych o wynikach leczenia chorych z przewlekłym zapaleniem zatok przynosowych oraz cierpliwe wsparcie w przeprowadzeniu setek kontroli endoskopowych u operowanych pacjentów.

Równie serdecznie dziękuję wszystkim osobom, które wspierały mnie w trakcie prowadzenia badań naukowych - Koleżankom i Kolegom z Kliniki Otolaryngologii w Krakowie, Pielęgniarkom Oddziałowym – Pani Marii Klicz i Pani Teresie Mikoń oraz pracownikom Administracji – Paniom Agnieszce Kruk, Agnieszce Włosik-Dziedzinie i Weronice Bujas.

# Spis treści

1. Wprowadzenie.....	7
1.1 Cele prognozowania w medycynie .....	7
1.2 Przykłady metod prognozowania stosowanych w medycynie.....	8
2. Cel pracy .....	12
3. Przewlekłe zapalenie zatok przynosowych – wprowadzenie i przegląd piśmiennictwa.....	13
3.1 Przewlekłe zapalenie zatok przynosowych (PZZP) – definicja i objawy .....	13
3.2 Ocena zaawansowania PZZP w tomografii komputerowej .....	16
3.3 Leczenie chirurgiczne chorych z PZZP.....	19
3.4 Wyniki chirurgii endoskopowej stosowanej u chorych z PZZP i metody ich oceny.....	21
3.5 Przewidywanie wyników leczenia PZZP – czynniki prognostyczne.....	23
3.6 Uwarunkowania nasilenia poszczególnych objawów PZZP oraz możliwości prognozowania ich zmian po leczeniu operacyjnym .....	28
3.7 Modele predykcyjne i sztuczne sieci neuronowe w PZZP.....	43
3.8 Podsumowanie .....	44
4. Sztuczne sieci neuronowe – wprowadzenie i przegląd piśmiennictwa.....	46
4.1 Wprowadzenie.....	46
4.2 Zastosowania sztucznych sieci neuronowych – przegląd piśmiennictwa.....	46
4.3 Modelowanie z wykorzystaniem sieci neuronowych.....	49
4.4 Budowa i zasada działania sieci neuronowych .....	50
4.5 Architektura sieci .....	51
4.6 Uczenie sieci neuronowej.....	53
4.7 Kodowanie danych na potrzeby sieci neuronowych .....	55
4.8 Typy problemów rozwiązywanych przez sieci neuronowe .....	55
4.9 Zbiór uczący i zmienne wejściowe .....	56
4.10 Zdolność sieci do generalizacji, walidacja .....	58
4.11 Testowanie .....	59
4.12 Problem wyboru zbioru uczącego, walidacyjnego i testowego .....	60
4.13 Poszukiwanie optymalnego modelu .....	61
4.14 Interpretacja działania modelu - ograniczenia.....	61
4.15 Tworzenie modeli neuronowych – zadania badacza (podsumowanie).....	62
5. Materiał i metoda .....	63
5.1 Grupa badana i jej charakterystyka kliniczna.....	63
5.2 Kwestionariusz samooceny objawów .....	68

5.3 Zakres i wiarygodność analizowanych danych .....	70
5.4 Analiza statystyczna .....	70
5.5 Sztuczne sieci neuronowe .....	71
6. Wyniki .....	84
6.1 Uwagi wstępne .....	84
6.2 Nasilenie dolegliwości po zabiegu operacyjnym .....	85
6.3 Zmiana nasilenia dolegliwości po zabiegu operacyjnym w porównaniu do stanu przed zabiegiem .....	112
6.4 Upośledzenie drożności nosa .....	125
6.5 Wydzielina z nosa .....	137
6.6 Spływanie wydzieliny po tylnej ścianie gardła .....	148
6.7 Zaburzenia węchu .....	160
6.8 Ból twarzy .....	169
6.9 Ból głowy .....	180
6.10 Przewlekłe zmęczenie .....	191
7. Dyskusja .....	201
7.1. Ocena zasadności podjętych badań .....	201
7.2. Zagadnienie przewidywalności wyników leczenia chorych z PZZP w kontekście przeglądu piśmiennictwa i w świetle własnych badań .....	201
7.3. Porównanie modeli opracowanych w ramach niniejszej pracy z modelami predykcyjnymi opisywanymi w literaturze .....	204
7.4. Identyfikacja nie opisywanych dotychczas czynników warunkujących nasilenie dolegliwości u chorych z PZZP .....	205
7.5. Możliwości zastosowania opracowanych modeli neuronowych w praktyce klinicznej oraz perspektywy ich dalszego rozwoju .....	211
7.6. Obiektywne ograniczenia możliwości modeli predykcyjnych w PZZP .....	216
7.7. Podsumowanie – znaczenie modeli prognozujących wyniki leczenia operacyjnego chorych z PZZP .....	218
8. Wnioski .....	219
9. Streszczenie .....	221
10. Summary .....	222
11. Piśmiennictwo .....	223

## Wykaz skrótów

AIA – astma aspirynowa ( <i>aspirin-intolerant asthma</i> )
APS – Automatyczny projektant sieci, narzędzie programu Statistica 7.0
CCS – <i>Chronic Sinusitis Survey</i> , kwestionariusz do oceny kontroli choroby u chorych z PZZP
EPOS – <i>European Position Paper on Rhinosinusitis and Nasal Polyps</i> , europejski konsensus dotyczący zapaleń zatok przynosowych
ESS – endoskopowa chirurgia zatok przynosowych, <i>endoscopic sinus surgery</i>
FESS – czynnościowa mikrochirurgia wewnątrznosowa, <i>functional endoscopic sinus surgery</i>
GERD – choroba refluksowa przełyku ( <i>gastroesophageal reflux disease</i> )
LPR – refluks krtaniowo-gardłowy ( <i>laryngopharyngeal reflux</i> )
NPR – refluks treści żołądkowej do nosogardła i jamy nosowej ( <i>nasopharyngeal reflux</i> )
PZZP – przewlekłe zapalenie zatok przynosowych
RSDI – <i>Rhinosinusitis Disability Index</i> , kwestionariusz do oceny wpływu objawów związanych z PZZP na ograniczenia w codziennym funkcjonowaniu pacjenta
RSI – <i>Rhinosinusitis Symptom Inventory</i> , kwestionariusz do oceny nasilenia dolegliwości związanych z PZZP
SNOT-20 - <i>20-item Sino-Nasal Outcome Test</i> , kwestionariusz do oceny jakości życia uwarunkowanej stanem zdrowia w PZZP
TK – tomografia komputerowa ( <i>computer tomography</i> )
VAS – skala wizualno-analogowa ( <i>visual-analogue scale</i> )

# 1. Wprowadzenie

## 1.1 Cele prognozowania w medycynie

Wyjaśniając cel prognozowania w medycynie, jeden z badaczy zajmujących się zastosowaniem sieci neuronowych w analizie przeżycia, Harry B. Burke, napisał: *„Jeśli wiedza oznacza władzę, to zdolność przewidywania i kontrolowania przyszłości zapewnia władzę ostateczną. Nauki medyczne umożliwiają poznanie przyszłości pacjenta i dają władzę, by odmienić jego los.”* [1].

Prognozowanie w medycynie ma zasadnicze znaczenie dla planowania terapii, stanowi podstawę świadomej zgody pacjenta na leczenie, a także pozwala na bardziej adekwatne planowanie organizacji opieki zdrowotnej [2].

Ocena rokowania jest nieodzownym elementem pracy klinicznej. **Wybór metody leczenia** zawsze opiera się na przewidywaniu jego efektu oraz na porównaniu szans uzyskania poprawy przy zastosowaniu wybranego postępowania w stosunku do metod alternatywnych. Właściwa predykcja umożliwia podjęcie terapii dającej potencjalnie najlepsze wyniki i odstąpienie od działań obarczonych dużym ryzykiem niepowodzenia. Należy zauważyć, że w praktyce każda decyzja terapeutyczna w mniej lub bardziej jawny sposób odnosi się do analizy rokowania. Nawet podejmując typową, codzienną decyzję, czy zaordynować antybiotyk u gorączkującego dziecka, lekarz intuicyjnie dokonuje oszacowania szansy uzyskania poprawy po włączeniu leku i porównuje wynik z szansą wyleczenia bez jego zastosowania. Analizuje przy tym czynniki indywidualnie charakteryzujące danego pacjenta (objawy kliniczne, wyniki badań dodatkowych) i porównuje je z podobnymi przypadkami, które zna z doświadczenia bądź literatury (rys. 1).

Przekazanie informacji o przewidywanym efekcie terapii jest również warunkiem uzyskania **świadomej zgody chorego** na proponowane leczenie. Dysponując informacją o rokowaniu pacjent jest w stanie lepiej zrozumieć własną sytuację, rozważyć potencjalne korzyści z wybranej metody postępowania w kontekście ewentualnego ryzyka i na tej podstawie może podejmować świadome, racjonalne decyzje dotyczące swojego leczenia [3].

Ponadto pacjenci często odczuwają **potrzebę poznania rokowania**, nawet jeśli ta informacja nie ma wpłynąć w żaden sposób na decyzje odnośnie dalszego postępowania. Przykładem może być wykonywanie w I trymestrze ciąży badania ultrasonograficznego określającego ryzyko wad chromosomalnych u płodu (tzw. „USG genetyczne”). W Polsce badanie to u większości kobiet nie jest refundowane, a jego cena jest stosunkowo wysoka. Mimo to wiele osób decyduje się je wykonać, chociaż deklarują, że nawet niekorzystny wynik nie będzie skutkowało decyzją o przerwaniu ciąży. Głównym celem korzystania z tej metody prognostycznej wydaje się zatem w wielu przypadkach wyłącznie potrzeba redukcji lęku w sytuacji związanej nieodłącznie z niepewnością i oczekiwaniem.

Zidentyfikowanie pacjentów z dużym prawdopodobieństwem zagrożonych niepowodzenia leczenia ma także znaczenie po jego zakończeniu, tj. przy planowaniu wizyt kontrolnych. Świadomość wysokiego ryzyka niepomyślnego przebiegu choroby skłania do ściślejszego monitorowania określonych pacjentów w celu możliwie szybkiego **rozpoznania ewentualnego niepowodzenia terapii** (np. wznowy po operacji). Dzięki temu konieczna interwencja (np. reoperacja) jest możliwa do podjęcia bez zbędnej zwłoki.

Powyższe uwagi odnosiły się do prognozowania w skali jednostkowej (pojedynczego pacjenta). W skali szpitala, miasta czy kraju prognozowanie ma jednak także istotne znaczenie, ponieważ przewidywanie wyników leczenia w odniesieniu do całej populacji leczonych chorych umożliwia racjonalne **zaplanowanie organizacji opieki zdrowotnej** i właściwą dystrybucję środków - stosownie do prognozowanych potrzeb.

## 1.2 Przykłady metod prognozowania stosowanych w medycynie

Jak wspomniano powyżej, podstawą każdej predykcji jest określenie zmiennych wpływających na wynik, który ma być przewidywany. Zmienne te są często nazywane **czynnikami prognostycznymi** lub **czynnikami ryzyka**. Z tego względu pierwszym krokiem w większości prac mających na celu przewidywanie efektu leczenia są zwykle analizy statystyczne zmierzające do zidentyfikowania takich czynników. Na ogół wyboru potencjalnych czynników prognostycznych dokonuje się na podstawie obserwacji klinicznych bądź teorii tłumaczących patogenezę badanego schorzenia. Ewaluacji znaczenia wybranych czynników najczęściej dokonuje się metodą porównania wyników leczenia w grupach chorych różniących się pod względem występowania (lub braku) rozważanych czynników. Wnioskowanie na temat znaczenia określonego czynnika najczęściej prowadzi się, poszukując statystycznie istotnych różnic dotyczących przebiegu choroby między tymi pacjentami, u których dany czynnik ryzyka występuje, a tymi, u których nie jest on obecny.

W przypadku większości schorzeń tego rodzaju analizy wykazują, że ostateczny wynik leczenia zależy od oddziaływania nie pojedynczych czynników ryzyka, ale wielu czynników równocześnie. Dlatego różnymi metodami dąży się do tego, żeby tworzyć modele predykcyjne uwzględniające łączne oddziaływanie kilku zmiennych (**modele wieloczynnikowe**).

Najprostszymi tego rodzaju modelami są różnego rodzaju skale kliniczne, które pozwalają lekarzowi w łatwy sposób wyznaczać prawdopodobieństwo wystąpienia u danego pacjenta określonego efektu zależnie od obecności czynników ryzyka. Jednym z najczęściej stosowanych w praktyce modeli tego rodzaju jest **system TNM**, stosowany do oceny zaawansowania nowotworów złośliwych. Na podstawie trzech czynników klinicznych: stopnia zaawansowania miejscowego guza (T), zaawansowania przerzutów w węzłach chłonnych (N) i obecności przerzutów odległych (M) można między innymi przewidywać szansę konkretnego pacjenta na przeżycie (rys. 1). Jest to możliwe, ponieważ dla różnych guzów zebrano obszernie dane historyczne dotyczące losów wielu chorych z nowotworami



w różnym stopniu zaawansowania, co stanowi podstawę wnioskowania przez analogię, a system jest stale modyfikowany w miarę gromadzenia nowych danych klinicznych. Dzięki temu po określeniu stopnia zaawansowania w skali TNM można na przykład sprawdzić, jaki odsetek chorych wykazujących podobieństwo do pacjenta poddawanego aktualnie analizie przeżywa 5 lat od rozpoznania choroby [4, 5].

Podobnego rodzaju modelem jest niezwykle popularny **system SCORE** (*Systematic Coronary Risk Evaluation*) [6], służący do oceny ryzyka sercowo-naczyniowego w zależności od płci, wieku, palenia tytoniu, ciśnienia tętniczego i stężenia cholesterolu całkowitego. Dla ułatwienia posługiwania się systemem w codziennej praktyce stworzono tzw. karty ryzyka, na których w barwnych tabelach łatwo można odszukać wartość prawdopodobieństwa zgonu z powodu incydentu sercowo-naczyniowego odpowiednią dla konkretnego pacjenta (rys. 1). Zgodnie z wytycznymi, posługując się kartą należy obrazowo pokazać pacjentowi, jak zmieni się jego indywidualne ryzyko jeśli np. przestanie palić papierosy (w celu lepszego zobrazowania odpowiednie pola diagramów są kolorowe – od zielonych oznaczających małe ryzyko, poprzez żółte i pomarańczowe aż do ciemnoczerwonych). Należy zaznaczyć, że w celu stworzenia modelu (tj. wyznaczenia procentowego ryzyka zgonu dla każdego z „pól” diagramu SCORE) przeanalizowano 3 miliony osobo-lat obserwacji i 7934 incydentów sercowo-naczyniowych zakończonych zgonem.

Opisane powyżej metody są bardzo proste do zastosowania i dzięki temu szeroko stosowane. Jednak - jak się okazuje - dzięki powszechnej dostępności komputerów w gabinetach lekarskich, nawet bardziej wymagające modele bywają powszechnie używane w praktyce. Wspomniana wyżej **metoda oceny ryzyka wad chromosomalnych u płodu** na podstawie badania ultrasonograficznego bazuje na wielowymiarowym modelu (podobnie jak system SCORE), jednak do wyliczenia wyniku niezbędne jest posiadanie specjalnego oprogramowania komputerowego. Pozwala ono wyliczyć indywidualne ryzyko zależnie od kilku zmiennych wejściowych (wieku matki, szerokości przezierności karkowej, oceny kości nosowej, kąta twarzowego, przepływu na zastawce trójdzielnej i w przewodzie żylnym oraz czynności serca płodu) [7] –rys. 1.

Zgodnie z przedstawionymi założeniami, do opracowania wszystkich opisanych powyżej modeli niezbędne było określenie **zmiennych wejściowych** (czynników ryzyka), na podstawie których przewiduje się będące przedmiotem zainteresowania **zmienne wyjściowe**, a podstawę predykcji stanowią **dane historyczne**, tj. informacja o przebiegu choroby u pacjentów obserwowanych w przeszłości.

Na podobnych zasadach budowana jest większość modeli predykcyjnych znajdujących zastosowanie w medycynie. Należą do nich przede wszystkim rozmaite **modele statystyczne**. [8]. Dotychczas opracowano bardzo wiele metod obliczeniowych służących do rozwiązywania problemów związanych z prognozowaniem, poczynając od bardzo prostych po niezwykle wyrafinowane. Jako przykłady technik często stosowanych w medycynie należy wymienić przede wszystkim modele regresyjne (liniowe – rys. 1 – lub nieliniowe), mające postać równań, których argumentami (zmiennymi niezależnymi) są odpowiednie czynniki ryzyka, a wynik oznacza wartość prognozowaną –

np. prawdopodobieństwo wystąpienia przewidywanego zdarzenia. Szczególny przypadek takiego modelu to tzw. regresja logistyczna. Tego typu model służy do przyporządkowania przypadków do dwóch klas, np. w celu określenia, u których pacjentów prognozowane zdarzenie nastąpi, a u których nie. W kontekście technik regresyjnych warto również wspomnieć o modelu proporcjonalnego hazardu Coxa, powszechnie stosowanym w analizie przeżycia (tj. prognozowaniu czasu, jaki upłynie do wystąpienia określonego zdarzenia – np. zgonu z powodu choroby). Model ten opisuje, w jaki sposób czynniki rokownicze wpływają na czas przeżycia, a jego wyróżniającą cechą jest możliwość wykorzystania charakterystycznych dla analizy przeżycia niepełnych obserwacji, tzw. obserwacji uciętych [9]. Wśród innych metod często stosowane bywają również na przykład tzw. drzewa decyzyjne, tj. modele mające postać stopniowo rozgałęziającego się diagramu służącego do stopniowego podziału zbioru danych na coraz mniejsze grupy (np. w zależności od obecności lub braku określonych czynników ryzyka) – rys. 1. Dzięki temu w przypadku prognostycznego modelu klinicznego określonego pacjenta można ostatecznie zakwalifikować do pewnej wąskiej „kategorii” (analogicznie do przyporządkowania określonego „pola” na karcie ryzyka w systemie SCORE).

Punkt wyjścia do zastosowania opisanych powyżej technik, jak również do budowania innych, bardziej zaawansowanych modeli statystycznych stanowią zawsze pewne **wstępne założenia dotyczące charakteru modelowanej zależności**. Jednak w przypadku zagadnień predykcyjnych, zwłaszcza w medycynie, charakter zależności pomiędzy zmiennymi wejściowymi i wyjściowymi często nie jest znany. W takiej sytuacji przewagę nad opisanymi powyżej technikami regresyjnymi zyskują metody tzw. **uczenia maszynowego** (*machine learning*), stanowiące gałąź sztucznej inteligencji, a wśród nich zwłaszcza sztuczne sieci neuronowe – rys. 1.

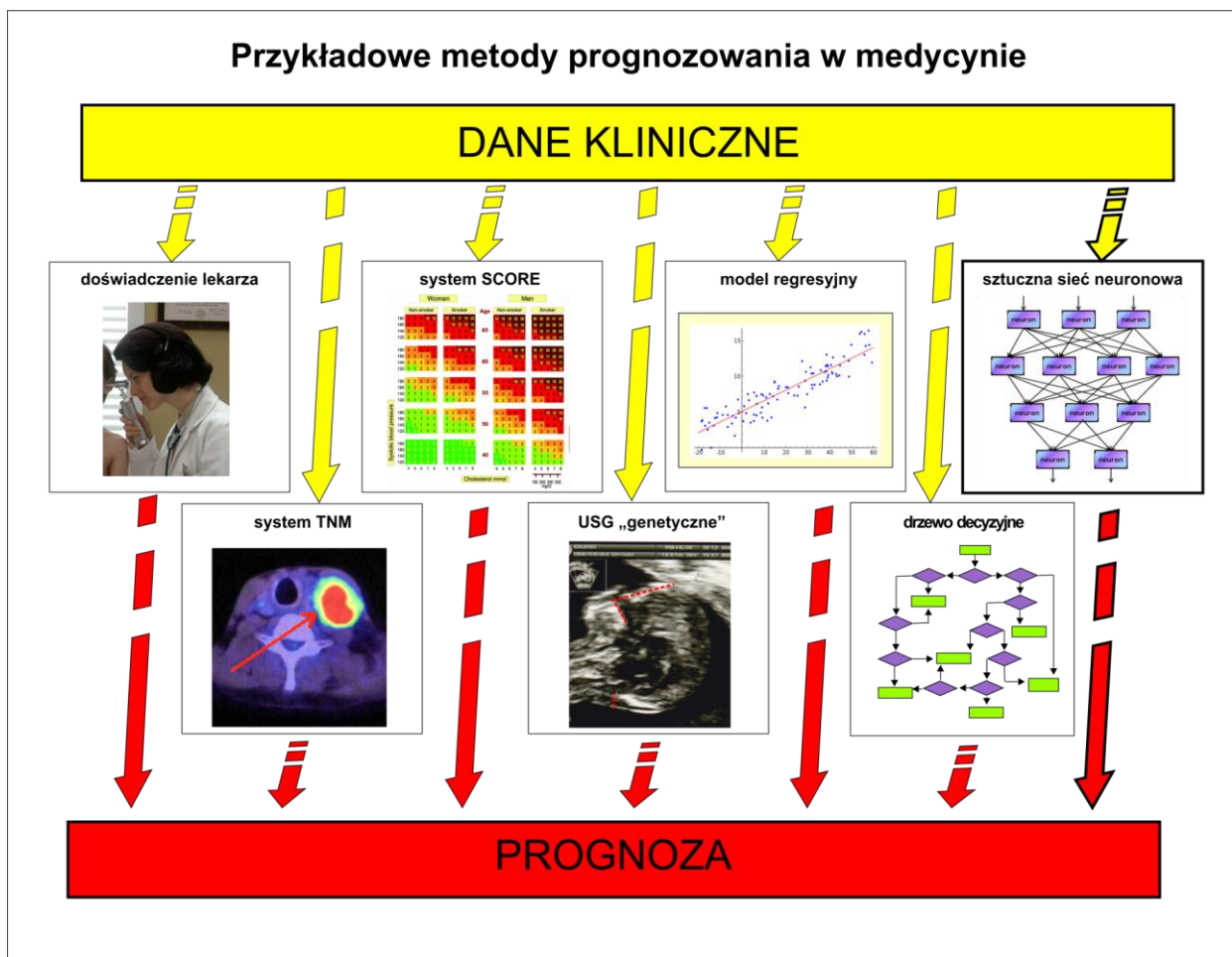
**Sztuczne sieci neuronowe** są systemami informatycznymi, których działanie naśladuje (w dużym uproszczeniu) funkcjonowanie układu nerwowego. Wyróżniające je cechy to **zdolność uczenia się na podstawie przykładów** i **możliwość automatycznego uogólniania zdobytej wiedzy**. Modele prognostyczne wykorzystujące sztuczne sieci neuronowe mają kilka istotnych dla medycyny zalet:

- dobrze nadają się do tworzenia **modeli wielowymiarowych** (gdzie wynik zależy od łącznego oddziaływania wielu różnych czynników, czyli zmiennych wejściowych),
- w procesie modelowania mogą uwzględniać wzajemne **interakcje** pomiędzy poszczególnymi zmiennymi wejściowymi,
- z założenia są przystosowane do modelowania zależności **nieliniowych**,
- nie są ograniczone do wykorzystania zmiennych ilościowych wyrażanych w skali ilorazowej (jak np. wiek), ale są w stanie przetwarzać zmienne wyrażone w skali przedziałowej lub porządkowej (np. poziom wykształcenia), a nawet zmienne jakościowe typu opisowego (np. płeć, zawód),
- nie wymagają żadnych **założeń formalnych** dotyczących metody wyliczania zmiennych wyjściowych na podstawie zmiennych wejściowych,

- prawidłowo działający model neuronowy może zostać łatwo przekształcony w prosty w obsłudze i tani program komputerowy, co umożliwia jego **praktyczne wykorzystywanie na dowolnym komputerze.**

**Rysunek 1. Wybrane przykłady metod prognozowania w medycynie.**

(Pochodzenie zapożyczonych ilustracji i licencje na ich udostępnienie: lekarz - [http://commons.wikimedia.org/wiki/File:Female\\_doctor\\_examines\\_a\\_child.JPG](http://commons.wikimedia.org/wiki/File:Female_doctor_examines_a_child.JPG) – domena publiczna; powiększony węzeł chłonny szyi uwidoczniiony w badaniu PET-CT - [http://commons.wikimedia.org/wiki/File:PET-CT\\_scanning\\_of\\_lymph\\_node\\_metastases\\_in\\_cancer.jpg](http://commons.wikimedia.org/wiki/File:PET-CT_scanning_of_lymph_node_metastases_in_cancer.jpg) – licencja Creative Commons Attribution-Share Alike 3.0 Unported, autor Akira Kouchiyama; system SCORE - European guidelines on cardiovascular disease prevention in clinical practice. [6]; regresja liniowa - [http://en.wikipedia.org/wiki/File:Linear\\_regression.svg](http://en.wikipedia.org/wiki/File:Linear_regression.svg) – domena publiczna).



## 2. Cel pracy

Celem niniejszej pracy była ocena możliwości zastosowania sztucznych sieci neuronowych w prognozowaniu wyników leczenia operacyjnego chorych z przewlekłym zapaleniem zatok przynosowych (PZZP). Przegląd dostępnego piśmiennictwa wykazał, że jakkolwiek sztuczne sieci neuronowe okazały się skutecznymi narzędziami predykcyjnymi w różnych dziedzinach nauki, w tym w medycynie, nie były dotychczas wykorzystywane do prognozowania efektów terapii schorzeń o podłożu zapalnym w otolaryngologii. Na podstawie przeprowadzonego przeglądu literatury stwierdzono także, iż brakuje również innych wiarygodnych metod umożliwiających przewidywanie z wyprzedzeniem wyniku leczenia operacyjnego indywidualnie u każdego pacjenta z PZZP, zależnie od występujących u niego czynników ryzyka niepowodzenia operacji.

Aktualnie za najistotniejszy cel leczenia pacjentów z PZZP uważa się redukcję lub eliminację dolegliwości subiektywnych, których uciążliwość w samoocenie pacjentów często nie koreluje z wynikami badań obiektywnych [10]. Trudne do zaakceptowania nasilenie objawów choroby jest również dominującą przyczyną poszukiwania pomocy medycznej przez chorych oraz najczęstszą motywacją do poddania się leczeniu chirurgicznemu w przypadku nieskuteczności postępowania zachowawczego. Pacjenci decydujący się na operację zatok przynosowych na ogół potrafią wskazać jeden lub kilka objawów, które są dla nich szczególnie uciążliwe i których wyeliminowanie w przebiegu leczenia traktują jako priorytet [3]. Informacja na temat szansy na osiągnięcie założonych celów stanowi podstawę świadomej zgody pacjenta na zabieg, a także jest kluczem do prawidłowej kwalifikacji chorych do leczenia operacyjnego. Z tego względu w niniejszej pracy skupiono się na **analizie czynników wpływających na nasilenie objawów subiektywnych** po zabiegu operacyjnym oraz opracowano **modele służące do przewidywania skutków leczenia** w odniesieniu do dolegliwości związanych z PZZP.

**Szczegółowe cele pracy obejmowały prognozowanie:**

- sumarycznego nasilenia dolegliwości związanych z PZZP po leczeniu operacyjnym,
- zmiany nasilenia dolegliwości po leczeniu operacyjnym w porównaniu do stanu sprzed zabiegu,
- pooperacyjnego nasilenia następujących objawów:
  - upośledzenie drożności nosa,
  - wydzielina z nosa,
  - spływanie wydzieliny po tylnej ścianie gardła,
  - zaburzenia węchu,
  - ból i uczucie rozpierania w obrębie twarzy,
  - ból głowy,
  - przewlekłe zmęczenie.

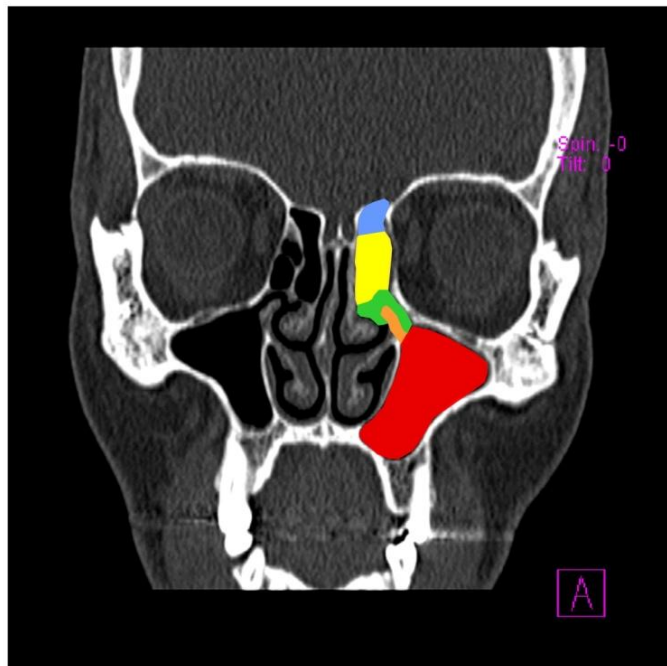
## 3. Przewlekłe zapalenie zatok przynosowych – wprowadzenie i przegląd piśmiennictwa

### 3.1 Przewlekłe zapalenie zatok przynosowych (PZZP) – definicja i objawy

Zgodnie z europejskim konsensusem dotyczącym zapaleń zatok przynosowych i polipów nosa (*European Position Paper on Rhinosinusitis and Nasal Polyps*, EPOS 2012) [10] zapalenie zatok przynosowych definiuje się jako **proces zapalny błony śluzowej nosa i zatok** (rys. 2 i 3), który charakteryzują:

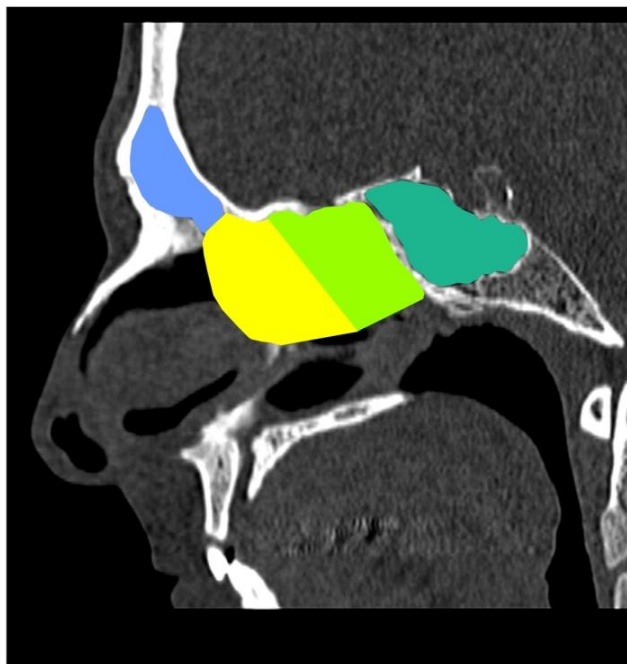
- co najmniej 2 spośród wymienionych **objawów**, z których jednym powinno być upośledzenie drożności nosa lub wydzielina z nosa:
  - upośledzenie drożności nosa,
  - wydzielina z nosa (w jamach nosa lub spływająca po tylnej ścianie gardła),
  - ból lub rozpieranie w obrębie twarzy,
  - upośledzenie lub utrata węchu (u dorosłych) bądź kaszel (u dzieci).
- zmiany widoczne w **badaniu endoskopowym** (rys. 4):
  - polipy nosa,
  - i/lub wydzielina śluzowo-ropna w przewodzie nosowym środkowym,
  - i/lub obrzęk błony śluzowej w przewodzie nosowym środkowym;
- i/lub zmiany stwierdzone w **tomografii komputerowej** (rys. 5):
  - pogrubienie błony śluzowej w obrębie kompleksu ujściowo-przewodowego i/lub zatok przynosowych.

**Rysunek 2. Zatoki przynosowe – lokalizacja najważniejszych struktur anatomicznych (tomografia komputerowa, przekrój w płaszczyźnie czołowej).**



- zachyłek czołowy
- sitowie przednie
- kompleks ujściowo-przewodowy
- wyrostek haczykowy
- zatoka szczękowa

**Rysunek 3. Zatoki przynosowe – lokalizacja najważniejszych struktur anatomicznych (tomografia komputerowa, przekrój w płaszczyźnie strzałkowej).**



- zatoka czołowa
- sitowie przednie
- sitowie tylne
- zatoka klinowa

**Rysunek 4. Obraz z badania endoskopowego jamy nosowej – polipy nosa i wydzielina śluzowo-ropna.**

Zdjęcie pochodzi ze strony internetowej <http://www.78steps.com>, gdzie zostało udostępnione do kopiowania, dystrybucji i użytkowania na mocy licencji Creative Commons Attribution 3.0 - Uznanie Autorstwa.

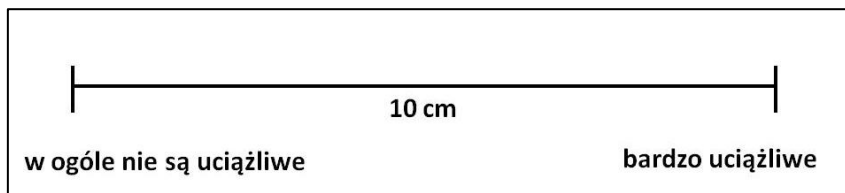


**Rysunek 5. Tomografia komputerowa zatok przynosowych – widoczne pogrubienie błony śluzowej w obrębie kompleksu ujściowo-przewodowego, w zatokach szczękowych i sitowiu.**



W celu oceny stopnia nasilenia schorzenia wykorzystuje się 10-punktową **skalę wizualno-analogową** (*visual-analogue scale, VAS*), na której pacjent wskazuje, jak uciążliwe są dla niego dolegliwości związane z zapaleniem zatok przynosowych (rys. 6).

**Rysunek 6. Skala wizualno-analogowa (VAS) – pacjent zaznacza, jak bardzo uciążliwe są objawy zapalenia zatok przynosowych.**



Wyróżnia się trzy **stopnie nasilenia** zapaleń zatok przynosowych [11]:

- łagodne – VAS = 0-3 pkt.
- umiarkowane – VAS >3-7 pkt.
- ciężkie VAS > 7-10 pkt.

Zapalenie zatok przynosowych określa się jako **przewlekłe** jeśli objawy utrzymują się powyżej 12 tygodni. [10]

Poza wymienionymi powyżej objawami definiującymi przewlekłe zapalenie zatok przynosowych wg kryteriów EPOS, u chorych często występują dolegliwości takie jak: bóle głowy, przykry zapach z ust, zmęczenie, osłabienie, ból zębów, kaszel, ból, ucisk lub uczucie pełności w uchu [12].

Częstość występowania PZZP jest bardzo wysoka. Szacuje się, że w populacji europejskiej podane powyżej kryteria rozpoznania choroby spełnia 10.9% populacji (6.9-27.1% zależnie od lokalizacji geograficznej) [13].

### **3.2 Ocena zaawansowania PZZP w tomografii komputerowej**

Jak zaznaczono powyżej, tomografia komputerowa stanowi złoty standard w rozpoznawaniu PZZP, a także jest niezbędnym badaniem obrazowym w przypadku planowania leczenia operacyjnego. Spośród różnorodnych metod klasyfikacji zaawansowania zmian u chorych z PZZP na podstawie wyników tomografii komputerowej, jedną z najczęściej stosowanych jest skala zaproponowana przez Lund i Mackay'a w 1993 roku [14]. Skala ta była kilkakrotnie modyfikowana przez różne grupy ekspertów i poddawana walidacji [15, 16]. Obecna wersja wydaje się dobrze odpowiadać potrzebom klinicznym ze względu na prostotę oceny i dobrą powtarzalność wyników [17]. Zasady **klasyfikacji Lund-Mackay** przedstawiono w tabeli 1.

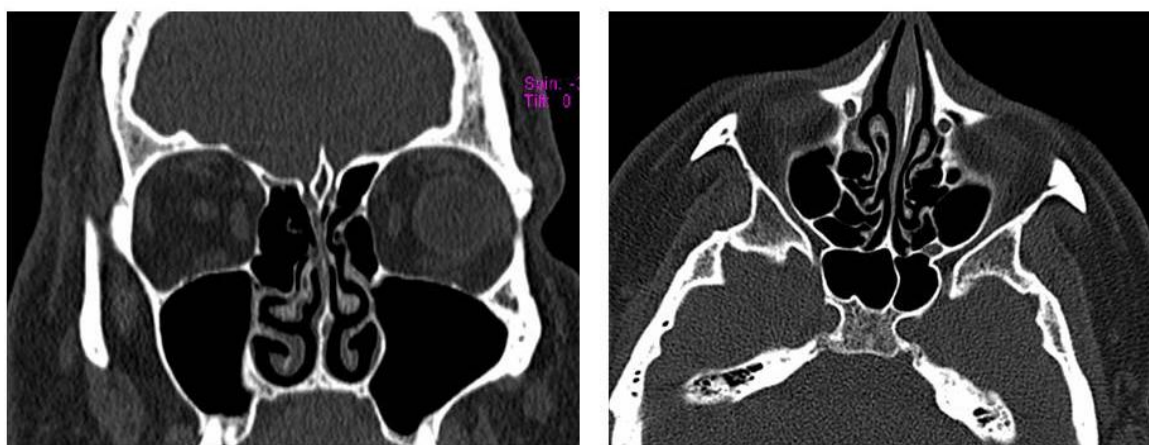


Każdej z grup zatok (zatoka szczękowa, sitowie przednie, sitowie tylne, zatoka klinowa, zatoka czołowa) przydziela się odpowiednio: 0 punktów w przypadku braku zmian zapalnych, 1 punkt w przypadku częściowego zaciemnienia lub 2 punkty w przypadku całkowitego zaciemnienia. Dodatkowo ocenia się drożność kompleksu ujściowo-przewodowego: 0 punktów oznacza kompleks drożny, 2 punkty – niedrożny. Sumarycznie każdej ze stron (prawej i lewej) można przypisać od 0 do 12 punktów (łącznie 24 punkty) - rys. 7, 8, 9.

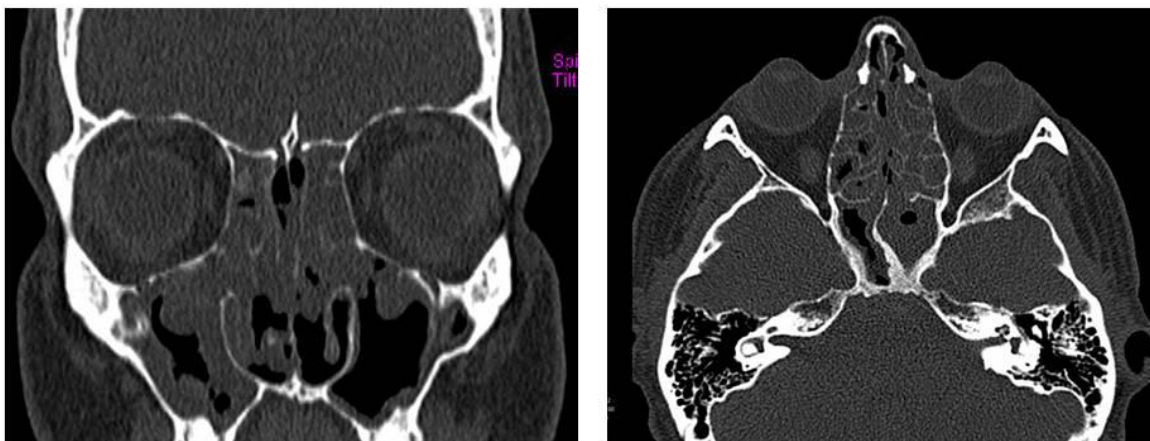
**Tabela 1. Stopień zaawansowania zmian zapalnych w przewlekłym zapaleniu zatok przynosowych na podstawie badania tomografii komputerowej (wg Lund i Mackay'a).**

Zatoka	Strona lewa	Strona prawa
Szczękowa	0-2	0-2
Sitowie przednie	0-2	0-2
Sitowie tylne	0-2	0-2
Klinowa	0-2	0-2
Czołowa	0-2	0-2
Kompleks ujściowo-przewodowy	0 lub 2	0 lub 2
Całkowita liczba punktów	0-12	0-12
	0-24	

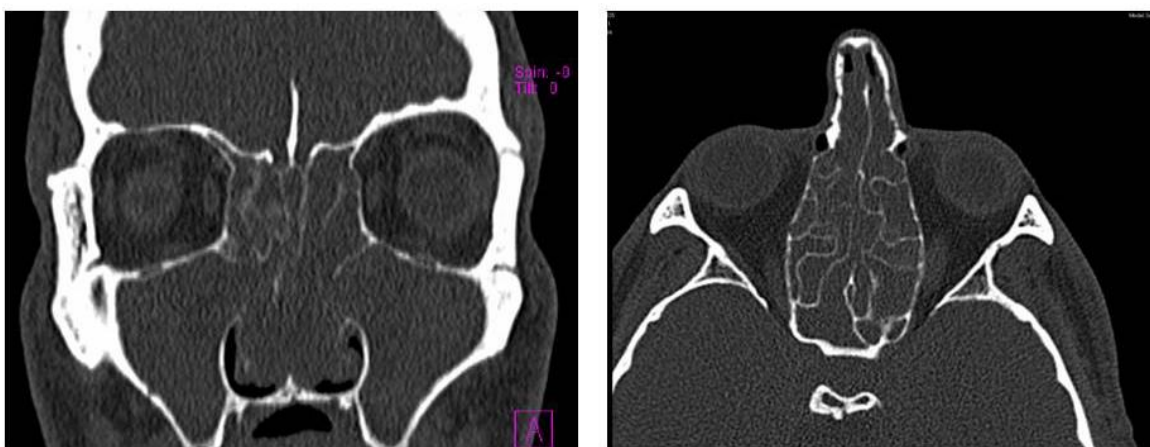
**Rysunek 7. Tomografia komputerowa zatok przynosowych – niski stopień zaawansowania zmian zapalnych (2 punkty w skali Lund-Mackay).**



**Rysunek 8. Tomografia komputerowa zatok przynosowych – średni stopień zaawansowania zmian zapalnych (13 punktów w skali Lund-Mackay).**



**Rysunek 9. Tomografia komputerowa zatok przynosowych – wysoki stopień zaawansowania zmian zapalnych (24 punkty w skali Lund-Mackay).**



### 3.3 Leczenie chirurgiczne chorych z PZZP

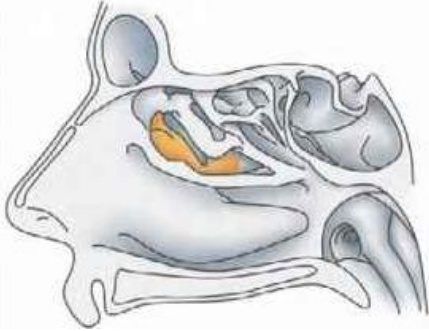
Zasadniczym celem leczenia chorych z PZZP jest **zlikwidowanie lub zmniejszenie uciążliwości objawów choroby** i – co za tym idzie - poprawa jakości życia pacjentów. U większości chorych w pierwszej kolejności rekomenduje się stosowanie leczenia zachowawczego (eliminacja czynników predysponujących do rozwoju PZZP, farmakoterapia). Do leczenia chirurgicznego kwalifikowani są pacjenci, u których prawidłowo prowadzone leczenie zachowawcze nie powoduje trwałej poprawy.

Współcześnie podstawową techniką operacyjną w leczeniu przewlekłych stanów zapalnych zatok przynosowych jest **mikrochirurgia wewnątrznosowa**, zazwyczaj przeprowadzana pod kontrolą endoskopu (chirurgia endoskopowa zatok przynosowych, *endoscopic sinus surgery – ESS*). Jej podstawowym celem jest zwykle przywrócenie wentylacji i drenażu, tj. prawidłowej funkcji zatok (stąd określenie: czynnościowa chirurgia endoskopowa zatok przynosowych, *functional endoscopic sinus surgery – FESS*). Ograniczona interwencja chirurgiczna polega na wytworzeniu trwałego połączenia między jamami nosa a zatokami w miejscu ich naturalnych ujść oraz na usunięciu zmian zapalnych, przy czym potencjalnie zdrowa błona śluzowa zatok jest usuwana możliwie oszczędnie [18]. Taką koncepcję leczenia opracowano wychodząc z założenia, iż istotną rolę w patogenezie zapaleń zatok przynosowych odgrywa zaburzenie drożności ich ujść, zwłaszcza tzw. kompleksu ujściowo-przewodowego - por. rys. 2 [19, 20]. Aktualnie uważa się, że dla rozwoju PZZP kluczowe znaczenie ma zaburzenie interakcji między organizmem a środowiskiem. Tłumacząc skuteczność chirurgii endoskopowej u chorych z PZZP w kontekście tej hipotezy zakłada się, że istotnym efektem operacji jest usunięcie tkanki zapalnej i ładunku antygenów nasilających reakcję zapalną, a także ułatwienie penetracji leków donosowych do jamy nosowej i zatok przynosowych [10]. Zakres interwencji chirurgicznej zależy od rozległości procesu chorobowego i może obejmować: resekcję wyrostka haczykowatego, poszerzenie naturalnego ujścia zatoki szczękowej, otwarcie komórek sitowych przednich, komórek sitowych tylnych, zatoki klinowej i zatoki czołowej (rys. 10) [19].

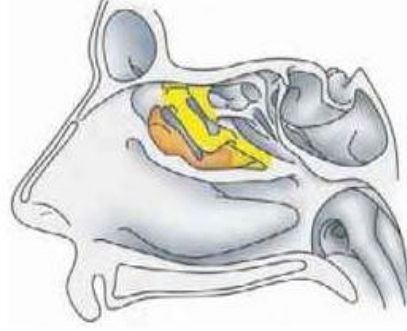
**Rysunek 10. Przykłady procedur chirurgicznych stosowanych w leczeniu przewlekłego zapalenia zatok przynosowych (kolory odpowiadają oznaczeniom zastosowanym na rys. 2 i 3).**

Rysunki pochodzą ze strony internetowej <http://www.78steps.com>, gdzie zostały udostępnione do kopiowania, dystrybucji i użytkowania na mocy licencji Creative Commons Attribution 3.0 (Uznanie Autorstwa).

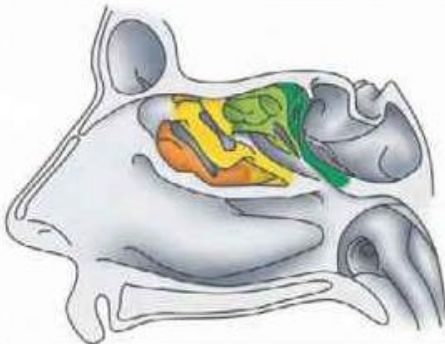
**A – infundibulotomia, inaczej uncinektomia (częściowe usunięcie wyrostka haczykowego – otwarcie zatoki szczękowej)**



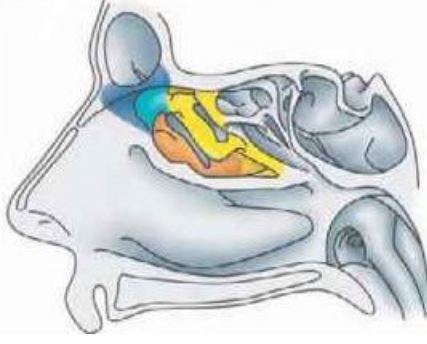
**B – ethmoidektomia przednia (infundibulotomia i otwarcie komórek sitowych przednich)**



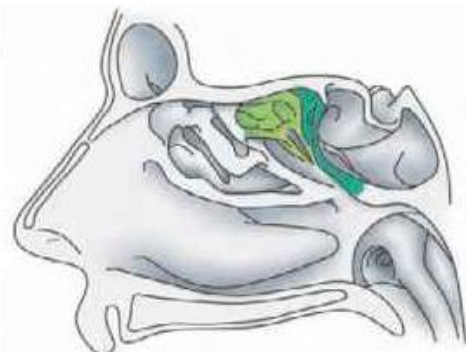
**C – sphenothmoidektomia (infundibulotomia, otwarcie komórek sitowych przednich i tylnych oraz zatoki klinowej)**



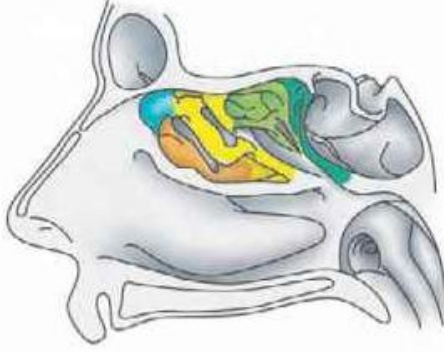
**D – frontoethmoidektomia (infundibulotomia, ethmoidektomia przednia i otwarcie zatoki czołowej)**



**E – sphenoidotomia (otwarcie zatoki klinowej oraz dodatkowo komórek sitowia tylnego)**



**F – frontosphenothmoidektomia (otwarcie zatoki szczękowej, sitowia przedniego i tylnego, zatoki czołowej i zatoki klinowej)**



### 3.4 Wyniki chirurgii endoskopowej stosowanej u chorych z PZZP i metody ich oceny

Za najistotniejszy cel leczenia chorych z PZZP uznaje się osiągnięcie **kontroli choroby**, definiowanej jako stan, w którym pacjent jest wolny od objawów klinicznych lub objawy te nie są uciążliwe. Jeżeli to możliwe, dąży się również do uzyskania prawidłowego lub bliskiego prawidłowemu obrazu śluzówki w badaniu endoskopowym oraz do ograniczenia dalszego leczenia wyłącznie do preparatów donosowych. Ze względów etycznych tomografia komputerowa w okresie pooperacyjnym jest wykonywana jedynie wtedy, gdy utrzymują się dolegliwości, nie może zatem być traktowana jako standardowa metoda oceny wyników leczenia. [10]

Ponieważ najważniejszym celem terapii jest osiągnięcie subiektywnej poprawy i uwolnienie chorego od dolegliwości, opracowano różnorodne narzędzia mające na celu weryfikację skuteczności chirurgii endoskopowej w odniesieniu do nasilenia objawów klinicznych i ich wpływu na jakość życia chorych. Konsensus EPOS 2012 rekomenduje zastosowanie **skali wizualno-analogowej VAS** (rys. 6) w celu oceny sumarycznego nasilenia dolegliwości. W badaniach mających na celu niezależną ocenę uciążliwości poszczególnych objawów oprócz skali VAS stosuje się np. **skale 4-stopniową**, gdzie 0 oznacza „brak objawu”, 1 – „objaw mało uciążliwy”, 2 – „średnio uciążliwy”, 3 – „bardzo uciążliwy” [10].

Ponadto w celu oceny uciążliwości objawów PZZP i ich wpływu na jakość życia chorych opracowano rozmaite **kwestionariusze** [21], z których najpopularniejsze to:

- *20-item Sino-Nasal Outcome Test (SNOT-20)* - jedno z najwartościowszych narzędzi do oceny jakości życia uwarunkowanej stanem zdrowia (*health-related quality of life*) w PZZP, służy ocenie nasilenia 20 dolegliwości w skali 6-stopniowej (od 0 do 5): 10 pytań dotyczy objawów charakterystycznych dla PZZP, pozostałe 10 dolegliwości takich jak męczliwość, zaburzenia produktywności czy rozdrażnienie [22],
- *Chronic Sinusitis Survey (CSS)* – zawiera pytania o czas trwania objawów (pytania dotyczą objawów takich jak: ból głowy, wydzielina z nosa, spływanie wydzieliny po tylnej ścianie gardła, niedrożność nosa) oraz czasu stosowania farmakoterapii (antybiotyki, leki donosowe, leki doustne) w ciągu 2 miesięcy poprzedzających badanie [23],
- *Rhinosinusitis Disability Index (RSDI)* – pozwala na ocenę wpływu objawów związanych z PZZP na ograniczenia w codziennym funkcjonowaniu pacjenta [24].

Oceniając wyniki leczenia podczas **badania endoskopowego** zwraca się uwagę na obecność polipów, obrzęku, wydzieliny przezroczystej lub ropnej, blizn i zrostów oraz zasychania wydzieliny [17].

Liczne obserwacje potwierdzają, że wyniki badań obiektywnych (tomografii komputerowej i badania endoskopowego) rzadko korelują z ocenianym subiektywnie przez chorych nasileniem dolegliwości typowych dla PZZP. U wielu pacjentów po leczeniu operacyjnym stwierdza się znaczną poprawę w zakresie objawów klinicznych mimo utrzymywania się

nieprawidłowego obrazu w badaniu endoskopowym. Również kwestionariusze służące do oceny wyniku leczenia różnią się między sobą i często uzyskuje się odmienne wyniki zależnie od rodzaju zastosowanego narzędzia pomiarowego lub sposobu sformułowania pytań [25]. W związku z powyższym wyniki badań naukowych wykorzystujących niejednakowe metody oceny efektów leczenia bywają trudne do porównania.

Mimo kontrowersji wokół wyboru optymalnego sposobu oceny wyników, analiza piśmiennictwa dowodzi **wysokiej skuteczności chirurgii endoskopowej w leczeniu chorych z PZZP**. Większość publikacji podaje, że odsetek pacjentów uzyskujących zmniejszenie nasilenia dolegliwości po operacji jest bliski 90% [26, 27], a prawidłowy obraz w badaniu endoskopowym notuje się u około 70% operowanych chorych [28]. Istotna poprawa w odniesieniu do objawów subiektywnych utrzymuje się na ogół przez co najmniej kilka lat po operacji [29]. Reoperacja w ciągu 5 lat od pierwszego zabiegu jest niezbędna w około 13 do 20% przypadków [29]. Odsetek poważnych powikłań waha się w granicach 1-1,5%, jakkolwiek w przypadku reoperacji może sięgać nawet 7% [10]. Skuteczność chirurgii endoskopowej dobrze ilustruje wynik jednego z badań, w którym 88% pacjentów w okresie od 12 do 85 miesięcy po operacji endoskopowej zatok przynosowych stwierdziło, że poleciliby tego rodzaju zabieg przyjacielowi cierpiącemu z powodu podobnych dolegliwości [30].

Analizując zestawione powyżej wyniki leczenia operacyjnego pacjentów z PZZP należy mieć na uwadze fakt, że w praktyce terapia tego schorzenia nigdy nie sprowadza się wyłącznie do samego zabiegu, ale zawsze obejmuje również **leczenie farmakologiczne**. Zgodnie z przyjętymi aktualnie zasadami operacja powinna być poprzedzona odpowiednio długotrwałym i intensywnym leczeniem zachowawczym, tak aby do zabiegu kwalifikować wyłącznie tych chorych, u których nie da się uzyskać zadowalających efektów na drodze farmakologicznej. Kontynuacja leczenia zachowawczego, w tym odpowiednia terapia chorób towarzyszących, ma kluczowe znaczenie również w okresie pooperacyjnym [19]. U większości pacjentów po zabiegu konieczne jest długotrwałe stosowanie leków (steroidów donosowo i ogólnie, antybiotyków, leków przeciwobrzękowych, przeciwalergicznym i przeciwrefluksowym), a rezygnacja z ich stosowania może nawet okazać się niemożliwe. Tę prawidłowość uwzględniono przy formułowaniu przytoczonej wcześniej definicji kontroli choroby, gdzie za zadowalającą uważa się już sytuację, w której leczenie zachowawcze ograniczone jest wyłącznie do preparatów donosowych [10]. Ponadto w opiece pooperacyjnej istotną rolę pełni odpowiednia toaleta jamy nosowej i zatok przynosowych, zwłaszcza usuwanie skrzepów i strupów pod kontrolą endoskopu oraz płukanie nosa [31]. W związku z tym przytaczane w większości publikacji wyniki leczenia operacyjnego należy praktycznie w każdym przypadku traktować jako sumę efektów samej operacji oraz opieki pooperacyjnej. Co więcej dotychczas nie ustalono jednolitego schematu postępowania, który byłby optymalny dla wszystkich pacjentów po ESS, więc szczegółowy plan leczenia zależy od doświadczenia chirurga i indywidualnych potrzeb każdego pacjenta.

## **3.5 Przewidywanie wyników leczenia PZZP – czynniki prognostyczne**

**Wpływ różnych czynników klinicznych na nasilenie objawów i wyniki leczenia operacyjnego chorych z PZZP** był dotychczas przedmiotem licznych badań. Ich wyniki dowodzą, że chirurgię endoskopową można uznać za skuteczną metodę postępowania w PZZP, jednak nie jest jasne, jakie zmienne najistotniej wpływają na ostateczny efekt terapii.

Dokonano przeglądu literatury celem wyselekcjonowania czynników mających potencjalne znaczenie w prognozowaniu wyników leczenia PZZP. Poniżej przedstawiono najważniejsze spostrzeżenia poczynione przez różnych autorów dla kolejnych czynników prognostycznych.

### **3.5.1 Płeć**

Jedynie nieliczne badania sugerują zależność wyniku leczenia operacyjnego chorych z PZZP od płci. Różnice między kobietami a mężczyznami dotyczą głównie samooceny objawów: mimo braku różnic w badaniach obiektywnych, w niektórych badaniach kobiety podawały większe nasilenie objawów przed i po operacji, niższą jakość życia oraz mniejszą poprawę po leczeniu niż mężczyźni [25, 32, 33].

### **3.5.2 Wiek**

W badaniach dotyczących zależności efektów leczenia operacyjnego od wieku chorych nie stwierdzono gorszych wyników u osób starszych [34]. Zależność między wiekiem a wynikami zabiegu wydaje się nawet odwrotna do oczekiwanej: w licznych pracach wykazano, że efekty u pacjentów po 65 roku życia mogą okazać się lepsze niż u osób młodszych w odniesieniu do samooceny objawów [35], bądź wyników badania endoskopowego [36, 37]. U pacjentów starszych stwierdzano jednak większą liczbę powikłań pooperacyjnych niż u osób młodych [35, 38].

### **3.5.3 Nasilenie zmian w tomografii komputerowej**

Analiza piśmiennictwa dotyczącego zależności między zaawansowaniem zmian w tomografii komputerowej a nasileniem objawów choroby przed i po zabiegu operacyjnym nie pozwala na wyciągnięcie jednoznacznych wniosków. Część autorów podaje, że rozległość zmian w TK wcale nie koreluje z nasileniem objawów PZZP lub że korelacja jest bardzo słaba, a zaawansowanie zmian w przedoperacyjnej tomografii komputerowej nie warunkuje nasilenia objawów po leczeniu operacyjnym [39-41]. Inni autorzy dowodzą jednak występowania istotnych zależności między wynikiem TK a nasileniem dolegliwości związanych z PZZP i wynikami leczenia [28, 32, 42-44]. Rozbieżności te mogą wynikać między innymi z faktu, że w cytowanych pracach stosowane są różne skale oceny wyniku tomografii komputerowej oraz różne metody ewaluacji efektu terapii. Wydaje się, że zaawansowanie choroby ocenione na podstawie przedoperacyjnej tomografii

komputerowej wpływa przede wszystkim na wyniki leczenia w odniesieniu do badania endoskopowego, potrzeby dodatkowego leczenia farmakologicznego (sterydoterapia ogólna i antybiotykoterapia) po zabiegu oraz kwalifikacji do reoperacji, słabiej natomiast koreluje z ocenianym subiektywnie nasileniem objawów PZZP po zabiegu operacyjnym [10, 28, 32, 37].

#### **3.5.4 Polipy nosa**

Polipy nosa stwierdzane są u około 20% pacjentów z PZZP, częściej u chorych z towarzyszącą astmą lub nietolerancją niesteroidowych leków przeciwzapalnych. Obecność polipów nosa jest uważana za negatywny czynnik prognostyczny. Stwierdzono, że u pacjentów z polipami nosa na ogół stwierdza się większe zaawansowanie zmian chorobowych w tomografii komputerowej oraz w przedoperacyjnym i pooperacyjnym badaniu endoskopowym, częściej też zachodzi u nich potrzeba reoperacji [28, 45, 46]. Prawidłowości dotyczące wyników badań obiektywnych nie zawsze znajdują jednak odzwierciedlenie w nasileniu objawów subiektywnych. Należy zaznaczyć, że przed operacją chorzy z polipami nosa paradoksalnie często zgłaszają mniejsze nasilenie objawów klinicznych (zwłaszcza bólu głowy) niż chorzy bez polipów nosa, oraz porównywalną lub nawet większą satysfakcję z przebytej operacji [25, 29, 30, 46]. Jedynie nieliczne badania sugerują większe nasilenie objawów PZZP u chorych z polipami nosa po zabiegu operacyjnym [47].

#### **3.5.5 Relacja między nasileniem objawów przed i po zabiegu operacyjnym**

Chirurgia endoskopowa jest skuteczną metodą leczenia powodującą w większości przypadków istotne zmniejszenie uciążliwości objawów PZZP. Zauważono jednak, że mimo to nasilenie poszczególnych dolegliwości po zabiegu operacyjnym jest na ogół tym większe, im bardziej uporczywe były one przed leczeniem [43]. Oznacza to, że u pacjentów zgłaszających na przykład bardzo dokuczliwe bóle głowy lub spływanie wydzieliny po tylnej ścianie gardła, po operacji można się spodziewać poprawy, ale dolegliwości te nadal mogą dominować w obrazie klinicznym choroby.

#### **3.5.6 Wcześniejsze leczenie operacyjne**

Okolo 10-20% pacjentów operowanych z powodu PZZP wymaga reoperacji [26, 29]. Co więcej, pacjenci reoperowani z większym prawdopodobieństwem będą wymagali kolejnych reoperacji [32, 48], a ryzyko powikłań jest u nich wyższe niż u pacjentów operowanych po raz pierwszy [49]. Zagadnienie skuteczności reoperacji w zakresie zmniejszania nasilenia objawów klinicznych budzi kontrowersje. Niektórzy badacze stwierdzili, że po zabiegu operacyjnym dolegliwości związane z PZZP u osób reoperowanych są bardziej uciążliwe niż u operowanych po raz pierwszy [43, 44, 50], inne prace sugerują



jednak, że - w przypadku odpowiedniego leczenia farmakologicznego w okresie pooperacyjnym – nie stwierdza się istotnych różnic między tymi grupami chorych [49, 51, 52].

### **3.5.7 Alergia**

Do niedawna uważano, że alergia odgrywa istotną rolę w patomechanizmie PZZP, jednak w świetle najnowszych badań związek przyczynowo-skutkowy między tymi dwoma schorzeniami nie wydaje się już tak oczywisty [10]. Analizując piśmiennictwo dotyczące badań wpływu alergii na efekty leczenia operacyjnego chorych z PZZP trudno wyciągnąć jednoznaczne wnioski. W części prac stwierdzono gorsze wyniki uzyskane po leczeniu operacyjnym z następową farmakoterapią u osób z alergią [47]. Na uwagę zasługuje fakt, że statystycznie istotny negatywny wpływ alergii na wynik leczenia zanotowano między innymi w badaniu dotyczącym chorych z masywnymi polipami nosa [48]. W tej grupie zauważono częstsze występowanie wznowy polipów u chorych z alergią niż u pacjentów bez alergii. Z drugiej strony w wielu innych doniesieniach nie potwierdzono tych obserwacji [25, 30, 44, 53]. Szczególnie istotne wydaje się spostrzeżenie, że zastosowanie prawidłowego leczenia farmakologicznego chorych z alergią istotnie poprawia wyniki leczenia operacyjnego [54, 55]. Wydaje się, że jeśli chorzy z alergią są właściwie leczeni, nie stwierdza się różnic w zakresie efektów zabiegu endoskopowego u osób z alergią i bez alergii [28].

### **3.5.8 Astma oskrzelowa**

Astma oskrzelowa często współistnieje z PZZP, a liczne obserwacje dowodzą, że schorzenia górnych i dolnych dróg oddechowych są wzajemnie powiązane (stąd określenie „choroba połączonych dróg oddechowych” – *united airway disease*) [10, 19, 56-58]. Stwierdzono także, że właściwe leczenie chirurgiczne i farmakologiczne chorych z PZZP często pozytywnie wpływa również na przebieg astmy [59-62]. Mimo to, podobnie jak w przypadku omawianego powyżej problemu alergii, wyniki badań dotyczących wpływu obecności astmy oskrzelowej na efekt leczenia operacyjnego PZZP nie są jednoznaczne. W badaniu przeprowadzonym przez Smitha i wsp. [25] u pacjentów z astmą zanotowano większe zaawansowanie zmian w tomografii komputerowej wykonanej przed zabiegiem oraz częstsze występowanie nieprawidłowego obrazu śluzówki w przedoperacyjnym badaniu endoskopowym, ale nie wiązało się to z większym nasileniem objawów subiektywnych po leczeniu. Analogiczną tendencję zaobserwowali Kountakis i Bradley u pacjentów poddawanych reoperacji zatok [63], a także Kim i wsp. [28] w odniesieniu do pooperacyjnej oceny endoskopowej u dzieci (jakkolwiek wyniki tego badania nie osiągnęły progu istotności statystycznej). Starannej oceny wpływu astmy na wyniki leczenia PZZP dokonali Dejima i wsp. [64]. Porównali oni dwie grupy chorych nie różniące się klinicznie pod względem zaawansowania choroby i współistniejących chorób towarzyszących – z wyjątkiem astmy. W tym badaniu u pacjentów z PZZP i astmą zaobserwowano gorsze wyniki leczenia operacyjnego oceniane zarówno subiektywnie, jak i endoskopowo. Ponadto w kilku pracach udokumentowano większą

częstość nawrotu choroby po leczeniu operacyjnym u pacjentów z astmą [48, 65]. Istnieją jednak również liczne doniesienia nie potwierdzające istotnego wpływu współistnienia astmy na wynik operacji endoskopowej zatok [23, 30, 32].

### **3.5.9 Nadwrażliwość na niesteroidowe leki przeciwzapalne**

Nadwrażliwość na niesteroidowe leki przeciwzapalne (określana również jako astma aspirynowa) często manifestuje się jako tzw. „triada Samtera”, tj. astma oskrzelowa, nadwrażliwość na aspirynę i przewlekłe zapalenie zatok przynosowych z polipami nosa [66]. U większości pacjentów z nadwrażliwością na niesteroidowe leki przeciwzapalne stwierdza się bardzo duże zaawansowanie zmian w zatokach przynosowych oraz masywne polipy nosa. Praktycznie wszystkie badania dotyczące tego zagadnienia potwierdzają, że pacjentów z astmą aspirynową niezmiennie charakteryzuje większe nasilenie objawów po operacji i gorsze wyniki w ocenie endoskopowej. Chorzy ci również zdecydowanie częściej wymagają licznych reoperacji. [25, 67-72]

### **3.5.10 Refluks żołądkowo-przełykowy (GERD) i refluks nadprzełykowy**

W ostatnich latach coraz większą rolę w patogenezie PZZP przypisuje się chorobie refluksowej przełyku (*gastroesophageal reflux disease – GERD*), a zwłaszcza refluksowi nadprzełykowemu: refluksowi krtaniowo-gardłowemu (*laryngopharyngeal reflux – LPR*) oraz refluksowi treści żołądkowej do nosogardła i jamy nosowej (*nasopharyngeal reflux - NPR*). Sugerowane mechanizmy nasilania procesu zapalnego w zatokach przynosowych przez chorobę refluksową obejmują: bezpośrednie narażenie śluzówki nosa i nosogardła na działanie kwasu żołądkowego, reakcję odruchową z udziałem nerwu błędnego oraz infekcję *Helicobacter pylori*. U chorych z refluksiem krtaniowo-gardłowym nie zawsze współwystępują objawy typowe dla refluksu żołądkowo-przełykowego (*gastroesophageal reflux disease, GERD*), jednak jednoznaczne potwierdzenie występowania odpływów treści żołądkowej powyżej górnego zwieracza przełyku wymaga zastosowania trudno dostępnych i uciążliwych metod diagnostycznych (24-godzinna pH-metria z użyciem specjalnego pH-metru dedykowanego do badania tego typu refluksu). W części badań dotyczących wpływu refluksu treści żołądkowej na wyniki leczenia chorych z PZZP uwzględniono w związku z tym jedynie obecność objawów sugerujących GERD, stwierdzając istotnie gorsze wyniki leczenia w tej grupie pacjentów [30]. W pracy DelGaudio i wsp. potwierdzono, że u chorych, u których objawy PZZP utrzymywały się mimo leczenia operacyjnego, stwierdzano zarówno większe nasilenie objawów refluksu żołądkowo-przełykowego, jak i wyniki pH-metrii potwierdzające NPR [73]. Niektórzy autorzy uważają jednak, że refluks niekoniecznie przyczynia się do rozwoju PZZP, ale jedynie powoduje podobne objawy kliniczne, czym można tłumaczyć uporczywe utrzymywanie się dolegliwości w tej grupie chorych mimo właściwego leczenia zatok przynosowych [74].

### **3.5.11 Palenie tytoniu**

Udokumentowany negatywny wpływ dymu tytoniowego na funkcjonowanie śluzówki dróg oddechowych oraz transport śluzowo rzęskowy skłonił licznych badaczy do poszukiwania zależności między paleniem papierosów a efektem leczenia chorych z PZZP. W części badań zaobserwowano negatywny wpływ palenia na wyniki operacji endoskopowej zatok [44, 75, 76]. Gorsze wyniki ESS w ocenie endoskopowej stwierdzono również u dzieci narażonych na dym tytoniowy w środowisku domowym [28]. Nie wszyscy autorzy potwierdzają jednak istnienie tego typu zależności. W większości prac nie stwierdzono różnic między osobami palącymi i niepalącymi ani w badaniu endoskopowym i tomografii komputerowej, ani w subiektywnie ocenianym nasileniu objawów choroby zarówno w okresie przedoperacyjnym, jak i pooperacyjnym [25, 77]. W badaniach, które przeprowadzili Das i wsp. stwierdzono nawet większą subiektywną poprawę bezpośrednio po leczeniu operacyjnym u osób palących niż u niepalących [78], jednak różnica ta nie utrzymywała się w trakcie 4-letniej obserwacji po zabiegu [79].

### **3.5.12 Depresja i lęk**

W obiektywnej ocenie zaawansowania PZZP (endoskopia, tomografia komputerowa) nie stwierdzono różnic między pacjentami z depresją a pozostałymi chorymi, jednak u pacjentów cierpiących na depresję obserwowano tendencję do zgłaszania większego nasilenia objawów klinicznych zarówno przed, jak i po zabiegu operacyjnym [25, 80]. Należy jednak zaznaczyć, że niektóre powszechnie stosowane kwestionariusze oceniające nasilenie objawów PZZP, takie jak *Rhinosinusitis Disability Index* (RSDI) lub *Sinonasal Outcome Test* (SNOT-20) zawierają również pytania dotyczące samopoczucia emocjonalnego, a zatem depresja mogła wpływać na ostateczny wynik testu niezależnie od objawów bezpośrednio związanych z przewlekłym zapaleniem zatok przynosowych [81]. W przypadku kwestionariusza RSI (*Rhinosinusitis Symptom Inventory*) nie zanotowano statystycznie istotnej korelacji rezultatów tego badania z wynikami szpitalnej skali lęku i depresji (*Hospital Anxiety and Depression Scale*) [82].

### **3.5.13 Nasilenie krwawienia podczas zabiegu**

W pracy Albu i wsp. analizowano przyczyny niepowodzeń endoskopowej chirurgii zatok szczękowych z wykorzystaniem modelu proporcjonalnego hazardu Coxa (modelu statystycznego wykorzystywanego w analizie przeżycia) [83]. Wynik tej analizy wskazywał, że w badanej grupie chorych jedynym niezależnym czynnikiem prognostycznym było nasilenie krwawienia podczas zabiegu operacyjnego.

### **3.5.14 Eozynofilia obwodowa i tkankowa**

We współczesnej diagnostyce rynologicznej dużą wagę przywiązuje się do oceny charakteru nacieku zapalnego w błonie śluzowej zatok przynosowych, zwłaszcza obecności i proporcji eozynofili i neutrofilii. Oceny rodzaju komórek napływowych można dokonać metodami nieinwazyjnymi (cytologia złuszczeniowa) lub w badaniu histopatologicznym wycinków z błony śluzowej (np. materiału usuniętego w czasie operacji endoskopowej zatok). W zależności od wyniku badania, nacieki dzieli się na eozynofilowe i neutrofilowe (lub eozynofilowe i nieeozynofilowe). Charakter nacieku pozwala pośrednio wnioskować na temat patomechanizmu PZZP. Nacieki neutrofilowe spotyka się głównie u chorych bez polipów nosa, m.in. z zaburzeniami wentylacji i drenażu zatok, infekcją, refluksiem żołądkowo-przełykowym. Nacieki eozynofilowe częściej występują u pacjentów z polipami nosa, alergią, a także astmą aspirynową oraz korelują z większym nasileniem zmian w przedoperacyjnej tomografii komputerowej. [10, 19]

Sugeruje się w związku z tym, że eozynofilowy charakter nacieków może być negatywnym czynnikiem prognostycznym w leczeniu operacyjnym chorych z PZZP. Podwyższona eozynofilia obwodowa oraz eozynofilowy charakter nacieków zapalnych w śluzówce zatok przynosowych związany był z większą częstością nawrotów PZZP w ciągu 5-letniego okresu obserwacji m.in. w badaniu Matsuwaki i wsp. [84]. Podobną zależność między gęstością nacieku eozynofili w polipach nosa a ryzykiem wznowy potwierdzili również Tosun i wsp. [85].

## **3.6 Uwarunkowania nasilenia poszczególnych objawów PZZP oraz możliwości prognozowania ich zmian po leczeniu operacyjnym**

W poprzednim podrozdziale omówiono wpływ najważniejszych czynników klinicznych na wynik operacji endoskopowej zatok przynosowych w odniesieniu do wyników pooperacyjnego badania endoskopowego i tomografii komputerowej oraz sumarycznego nasilenia dolegliwości związanych z PZZP. Wystąpienie i stopień uciążliwości **poszczególnych objawów rozważanych niezależnie** wynika jednak z rozmaitych uwarunkowań, a ponadto różne dolegliwości w niejednakowym stopniu ustępują po leczeniu operacyjnym [3, 86]. Poniżej zostanie przedstawiona patogeneza najważniejszych symptomów PZZP oraz czynniki wpływające na ich subiektywne nasilenie i szanse uzyskania poprawy po operacji endoskopowej zatok przynosowych, ponieważ dane te są niezbędne do prawidłowego prognozowania wyników leczenia.

### **3.6.1 Upośledzenie drożności nosa**

Upośledzenie drożności nosa stanowi na ogół najbardziej nasilony objaw PZZP zgłaszany przez pacjentów kwalifikowanych do operacji z powodu nieskuteczności leczenia zachowawczego [87-89]. Jednocześnie dolegliwość ta jest najbardziej podatna na leczenie

chirurgiczne– poprawę w tym właśnie zakresie obserwuje się u największego odsetka chorych, a stopień notowanej poprawy zwykle bywa znaczniejszy niż w przypadku innych objawów PZZP [3, 88, 90, 91]. Wysoką skuteczność ESS w odniesieniu do upośledzonej drożności nosa potwierdzono w licznych badaniach [86, 87, 89, 92]. Co więcej Senior i wsp. wykazali, że ten efekt zabiegu jest trwały i u 97% chorych utrzymuje się przez ponad 7 lat po leczeniu operacyjnym [93].

Wykazano, że na zaburzenia drożności nosa u pacjentów z PZZP wpływają następujące (często wzajemnie powiązane) czynniki:

- Polipy nosa. Pacjenci z polipami nosa zwykle zgłaszają większe upośledzenie drożności nosa niż chorzy bez polipów, jednak mogą również oczekiwać bardziej zauważalnej poprawy po leczeniu operacyjnym. W związku z tym ostatecznie po zabiegu oceniana subiektywnie drożność nosa w obu grupach chorych jest porównywalna. [46, 94-97]
- Nasilenie zmian zapalnych w obrazie tomografii komputerowej. Stopień nasilenia zaburzeń drożności nosa u pacjentów przed operacją koreluje z zaawansowaniem choroby stwierdzanym w przedoperacyjnej tomografii komputerowej [98-100], co jednak nie musi wiązać się z gorszym rokowaniem u pacjentów z bardziej rozległymi zmianami w obrębie zatok.
- Nadwrażliwość na niesteroidowe leki przeciwzapalne. Pacjentów z nadwrażliwością na niesteroidowe leki przeciwzapalne cechuje na ogół znacznego stopnia niedrożność nosa, która zwykle ulega znacznej poprawie po ESS. Mimo to u około 40% chorych dolegliwości ponownie nawracają po leczeniu operacyjnym, co wiąże się ze wznową polipów nosa. [19, 53, 66, 67, 92]
- Astma. Przegląd literatury wykazuje, że pacjenci z astmą na ogół zgłaszają gorszą drożność nosa [92, 101] i istotnie mniejszą poprawę w tym zakresie po ESS niż chorzy bez współistniejącej astmy [63, 64].
- Płeć. W badaniu, które przeprowadzili Busaba i wsp., mężczyźni częściej niż kobiety zgłaszali upośledzenie drożności nosa. U mężczyzn notowano również większą poprawę po ESS w odniesieniu do tego objawu. Zaobserwowaną prawidłowość można jednak tłumaczyć większą częstością występowania polipów nosa u mężczyzn w badanej grupie. [102]
- GERD. W niewielkim badaniu pilotażowym DiBaise i wsp. notowali niewielką poprawę drożności nosa po leczeniu farmakologicznym inhibitorem pompy protonowej, co mogłoby sugerować rolę GERD w patogenezie tego objawu, jednak jak dotąd kwestia ta nie została dostatecznie wyjaśniona [103].

### **3.6.2 Wydzielina z nosa i sływanie wydzieliny po tylnej ścianie gardła**

Uzyskanie poprawy w odniesieniu do takich objawów PZZP jak wydzielina z nosa lub sływanie wydzieliny po tylnej ścianie gardła jest trudniejsze i mniej przewidywalne niż w przypadku upośledzenia drożności nosa. Jakkolwiek w większości badań notowano

istotne zmniejszenie ilości wydzieliny po operacji endoskopowej zatok [86, 89, 92], to u około 25% chorych nie udawało się uzyskać w tym zakresie poprawy [88, 91]. Senior i wsp. dowodzili, że ograniczenie nasilenia dolegliwości może utrzymywać się przez wiele lat po zabiegu [93], jednak badania Solera i wsp. nie potwierdziły tej tezy – w badanej przez nich grupie chorych już po 6 miesiącach dochodziło na ogół do ponownego pogorszenia [87]. Marks i Shamsa uznali spływanie wydzieliny po tylnej ścianie za objaw PZZP najtrudniejszy do zlikwidowania na drodze operacyjnej [32].

Obserwowane u części pacjentów utrzymywanie się tej dolegliwości mimo prawidłowo przeprowadzonego ESS można tłumaczyć ograniczonym wpływem operacji na nasilenie stanu zapalnego w obrębie błony śluzowej oraz obecnością chorób towarzyszących, które niezależnie od zmian zapalnych w zatokach przynosowych wpływają na zwiększenie produkcji śluzu w górnych drogach oddechowych.

Analiza piśmiennictwa dokonana pod kątem poszukiwania czynników wpływających na ilość wydzieliny z nosa oraz jej spływanie po tylnej ścianie gardła wykazała, że dolegliwości te mogą zależeć od następujących indywidualnych cech pacjenta:

- GERD. Odplwy treści żołądkowej mogą powodować zwiększenie produkcji wydzieliny w nosie nie tylko na zasadzie bezpośredniego drażnienia śluzówki (w przypadku refluksu do nosogardła i jamy nosowej), ale także na drodze odruchowej za pośrednictwem nerwu błędnego [104]. Spływanie wydzieliny po tylnej ścianie gardła jest bardzo charakterystyczną dolegliwością pacjentów z GERD i refluksem krtaniowo-gardłowym (również bez PZZP), a leczenie farmakologiczne refluksu przy pomocy inhibitorów pompy protonowej zmniejsza nasilenie tej dolegliwości [105, 106]. Można na tej podstawie wnioskować, że u pacjentów cierpiących na GERD w przypadku braku odpowiedniego leczenia farmakologicznego uporczywe spływanie wydzieliny po tylnej ścianie gardła może utrzymywać się niezależnie od poprawy funkcji zatok uzyskanej w wyniku operacji endoskopowej. Hipoteza ta jest obecnie przedmiotem intensywnych badań, jednak brak w tym aspekcie jednoznacznej konkluzji, a co za tym idzie – zaleceń terapeutycznych [10].
- Nasilenie zmian zapalnych w obrazie tomografii komputerowej. Wiele badań potwierdza, że rozległość zmian zapalnych stwierdzanych w przedoperacyjnym badaniu TK koreluje z ilością wydzieliny z nosa i uciążliwością jej spływania po tylnej ścianie gardła [98-100]. Podobnie jak w przypadku omawianej wcześniej niedrożności nosa, nie stwierdzono bezpośredniej zależności między przedoperacyjnym wynikiem TK a nasileniem dolegliwości po zabiegu operacyjnym.
- Astma jest jednym z nielicznych czynników, który wydaje się wpływać na wyniki leczenia operacyjnego w odniesieniu do ilości wydzieliny z nosa oraz spływania wydzieliny po tylnej ścianie gardła. U chorych z astmą zanotowano w tym zakresie nieco gorsze wyniki niż u pozostałych pacjentów [63, 64].
- Nadwrażliwość na niesteroidowe leki przeciwzapalne. Uciążliwe spływanie wydzieliny po tylnej ścianie gardła i wydzielina z nosa są częstymi objawami u pacjentów z nadwrażliwością na niesteroidowe leki przeciwzapalne i cechują się u nich

większym nasileniem niż w pozostałych grupach chorych [67, 92]. Po leczeniu operacyjnym często stwierdza się u tych chorych nawrót omawianych dolegliwości [53].

- Płeć. Analizując zależność wyników leczenia operacyjnego od płci, Marks i Shamsa zauważyli, że po ESS kobiety częściej od mężczyźni skarżą się na uporczywe spływanie wydzieliny po tylnej ścianie gardła [32].
- Polipy nosa. W pracy de Loos i wsp. stwierdzono, że pacjenci z polipami nosa częściej zgłaszali obecność wydzieliny z nosa, nieustanną konieczność noszenia chusteczek i oczyszczania nosa w porównaniu z pacjentami bez polipów [107]. Inne badania wykazały jednak, że dolegliwości te w znacznym stopniu dotyczą również pacjentów bez polipów nosa [94]. Obecności polipów nosa nie można zatem jednoznacznie uznać za czynnik warunkujący ilość wydzieliny u pacjentów z PZZP.

### **3.6.3 Zaburzenia węchu**

Upośledzenie węchu należy do najbardziej uciążliwych dolegliwości związanych z PZZP. Węch odpowiada nie tylko za doznania estetyczne związane z odczuwaniem zapachów, lecz także odgrywa istotną rolę w wyzwalaniu emocji i wspomnień oraz pełni funkcję ochronną, ostrzegając np. przed spożyciem nieświeżej żywności, ulatniającym się gazem lub pożarem [108-110].

Szacuje się, że PZZP odpowiada za około 15-30% wszystkich przypadków hyposmii (osłabienia węchu) i anosmii (utruty węchu) [111]. Patogeneza upośledzenia funkcji powonienia w PZZP jest niezwykle złożona, a przy tym nie do końca poznana. Z tego względu prognozowanie wyników leczenia w tym aspekcie wydaje się stanowić wyjątkowo trudne wyzwanie. Zaburzenia funkcji powonienia i ich uwarunkowania budzą szczególne zainteresowanie badaczy, dzięki czemu potencjalne czynniki kliniczne wpływające na zaburzenia węchu zostały zdecydowanie dokładniej przeanalizowane niż w przypadku innych objawów PZZP. Dlatego też zagadnienie to zostanie omówione nieco szerzej w niniejszej pracy.

Przyczyn upośledzenia odczuwania zapachów przez chorych z PZZP upatruje się w **dwóch mechanizmach: przewodzeniowym i czuciowo-nerwowym** [111]. Przewodzeniowe zaburzenia węchu wynikają z ograniczenia dopływu powietrza do nabłonka węchowego z powodu mechanicznej blokady spowodowanej obrzękiem śluzówki, obecnością polipów itp. Hyposmia tego rodzaju jest na ogół podatna na terapię i ma tendencję do ustępowania po leczeniu przeciwobrzękowym lub chirurgicznym. Wystąpienie zaburzeń węchu w drugim z wymienionych mechanizmów – tj. na podłożu czuciowo-nerwowym – ma bardziej skomplikowaną patogenezę. Wykazano, że stan zapalny toczący się w jamie nosowej i zatokach przynosowych obejmuje również nabłonek węchowy (*neuroepithelium*), który w przeszłości uważany był za niezdolny do indukcji odpowiedzi immunologicznej [112]. Nabłonek węchowy w warunkach fizjologicznych znajduje się w stanie tzw. „neuralnej homeostazy”: na skutek działania czynników szkodliwych (do których zalicza się również

stan zapalny) pewna liczba komórek węchowych ginie najprawdopodobniej w procesie apoptozy, ale na ich miejsce odbudowywane są nowe komórki węchowe. Uważa się, że procesy chorobowe mogą wywoływać nieskompensowany wzrost odsetka ginących komórek *neuroepithelium* [113]. Ponadto mediatory procesu zapalnego pobudzają gruczoły śluzowe do wzmożonego wydzielania. Śluz nabłonka węchowego produkowany przez gruczoły Bowmana jest wysoko swoistą substancją, istotnie różniącą się od śluzu oddechowego. Nadmierne wydzielanie może prowadzić do zmian w jego składzie jonowym i zmiany mikrośrodowiska komórek węchowych [114]. Ze względu na opisany powyżej złożony patomechanizm zaburzeń węchu w PZZP, optymalnym sposobem postępowania wydaje się połączenie chirurgii (celem poprawy dostępu powietrza do nabłonka węchowego) i glikokortykoterapii (w celu zmniejszenia nasilenia stanu zapalnego).

Większość badań potwierdza korzystny **wpływ zastosowania chirurgii endoskopowej na funkcję zmysłu powonienia** u chorych z PZZP [89, 109, 115-118]. Tym niemniej nie u każdego z operowanych pacjentów można się spodziewać dobrych wyników leczenia [115, 118, 119]. Podawany przez różnych autorów odsetek chorych, u których zanotowano poprawę węchu po ESS waha się w szerokich granicach od 23% do 97% w zależności od zastosowanej metody oceny zmysłu powonienia (metody subiektywne lub obiektywne) oraz charakterystyki badanej grupy [108, 119-121]. Należy jednak zaznaczyć, że obserwowana poprawa zazwyczaj jest niepełna, co oznacza, że mimo leczenia węch na ogół nie wraca do stanu prawidłowego, nawet w przypadku całkowitego usunięcia zmian mogących upośledzać dostęp powietrza do nabłonka węchowego [109, 116, 122, 123]. Ponadto wydaje się, że u części pacjentów uzyskane wyniki terapii mogą nie być trwałe [124]. Co więcej, do 20% operowanych chorych zgłasza pogorszenie odczuwania zapachów po ESS [108, 121, 125], co może być związane z uszkodzeniem *neuroepithelium* [122]. W metaanalizie przeprowadzonej przez Chestera i wsp., obejmującej 2070 pacjentów, hyposmia została uznana za jeden z najmniej podatnych na leczenie objawów PZZP [3].

Poniżej omówiono najważniejsze czynniki wpływające na funkcję zmysłu powonienia u chorych z PZZP, poczynając od najistotniejszych.

- **Polipy nosa.** Większość autorów uważa, że obecność polipów nosa stanowi najważniejszy czynnik wpływający na dysfunkcję węchu w PZZP [96, 107, 108, 111, 117, 119, 121, 125-127]. Upośledzenie powonienia u chorych z polipami nosa może wynikać nie tylko z mechanicznej blokady, ale także z nawracających infekcji i bliznowacenia [111]. Pacjenci, u których przed leczeniem operacyjnym stwierdzono polipy nosa, na ogół mają także gorszy węch po zabiegu [108]. Tym niemniej, ponieważ u tych chorych występują na ogół szczególnie nasilone zaburzenia powonienia, ich satysfakcja z poprawy uzyskanej dzięki leczeniu operacyjnemu bywa większa niż w przypadku pacjentów bez polipów nosa [117, 121, 125].
- **Wynik badania histopatologicznego.** W badaniu Solera i wsp. liczba eozynofili w nacieku zapalnym w obrębie błony śluzowej zatok była wyższa u chorych z anosmią i hyposmią niż u pacjentów bez zaburzeń węchu, jednak po uwzględnieniu wpływu obecności polipów nosa na wynik analizy, wyniki badania



histopatologicznego okazały się nie korelować z pooperacyjną poprawą węchu [128]. Pade i wsp. stwierdzili, że mimo gorszych wyników przedoperacyjnych pacjenci z eozynofilią wykazywali większą poprawę węchu niż pozostali, podobnie jak miało to miejsce w omówionym wcześniej przypadku chorych z polipami nosa [121].

- Nadwrażliwość na niesteroidowe leki przeciwzapalne wydaje się wpływać na zmysł powonienia podobnie jak polipowatość nosa i eozynofilia tkankowa, z którymi bardzo często współistnieje [129, 130]. Pacjentów wykazujących nadwrażliwość na aspirynę zazwyczaj cechuje duże nasilenie zaburzeń węchu przed zabiegiem oraz możliwość uzyskania znacznej poprawy dzięki leczeniu operacyjnemu [117, 121, 125]. U wielu spośród tych chorych anosmia utrzymuje się jednak nawet mimo leczenia [67, 131], a u znacznej części tych chorych, u których uzyskano poprawę na drodze operacyjnej, należy spodziewać się ponownego pogorszenia związanego z nawrotem polipów nosa [53].
- Wiek. Liczne badania dotyczące funkcji zmysłu węchu potwierdzają, z wiekiem ulega ona istotnemu pogorszeniu, co jest szczególnie zauważalne zwłaszcza u osób po 65 roku życia [109, 132, 133]. Podobną zależność obserwowano także u pacjentów z PZZP przed leczeniem operacyjnym [111, 125, 134]. Mimo to w wielu badaniach nie potwierdzono negatywnego wpływu wieku na wyniki ESS w odniesieniu do dysfunkcji zmysłu powonienia [117, 121, 125].
- Astma. Negatywny wpływ współistniejącej astmy na funkcję zmysłu powonienia u chorych z PZZP był opisywany przez licznych autorów [63, 101, 111, 117, 125, 135]. Pacjenci cierpiący na astmę na ogół uzyskują poprawę węchu po operacji, jednak wyniki badań porównujących ostateczne efekty leczenia w tej grupie chorych z pozostałymi pacjentami nie są jednoznaczne [64, 108, 117]. U chorych z astmą notowano również większą tendencję do nawrotu hyposmii po leczeniu operacyjnym niż u pozostałych pacjentów [53].
- Nasilenie zmian zapalnych w obrazie tomografii komputerowej. Jakkolwiek zakres zmian stwierdzony w przedoperacyjnym badaniu TK koreluje z nasileniem zaburzeń węchu występujących przed zabiegiem, jednak nie musi to implikować gorszego rokowania po leczeniu operacyjnym [39, 98, 100, 108, 117, 118, 123, 125]. W badaniu, które Sil i wsp. przeprowadzili 5 lat po operacji endoskopowej zatok odsetek pacjentów wykazujących poprawę węchu był tym większy, im większe zmiany obserwowano wyjściowo w tomografii komputerowej [37]. Young i wsp. stwierdzili jednak, że pacjenci ci mają mniejszą szansę na długotrwałe utrzymywanie się efektów operacji w odniesieniu do zaburzeń węchu [53].
- Nasilenie i czas trwania zaburzeń węchu przed podjęciem leczenia operacyjnego. Pacjenci, u których przed leczeniem operacyjnym stwierdzano bardzo nasiloną dysfunkcję zmysłu powonienia, na ogół cechowali się gorszym węchem również po operacji [121]. Chorzy ci zgłaszali jednak najbardziej zauważalny zysk z przeprowadzonego leczenia [111]. W większości badań nie stwierdzono natomiast korelacji pomiędzy czasem trwania objawów, a stopniem utraty węchu lub jego pooperacyjnej poprawy [118, 136].

- Wcześniejsze leczenie operacyjne. W wielu badaniach stwierdzono, że pacjenci, którzy przebyli liczne polipektomie lub operacje endoskopowe zatok przynosowych, mają gorszy węch od pozostałych chorych [111, 134]. Przebyte leczenie może jednak nie być niezależnym czynnikiem warunkującym dysfunkcję węchu, ale zarówno zaburzenia powonienia, jak i konieczność licznych reoperacji mogą wynikać z analogicznych przyczyn. Przebyte leczenie operacyjne nie musi również warunkować gorszego rokowania po leczeniu operacyjnym [117, 125].
- Płeć. Uważa się, że kobiety potrafią rozpoznawać zapachy lepiej niż mężczyźni, a sprawność zmysłu węchu w mniejszym stopniu słabnie u nich z wiekiem. [109]. W niektórych badaniach dotyczących pacjentów z PZZP u mężczyzn notowano gorsze wyniki w przedoperacyjnych testach węchu [108], jednak większość autorów nie stwierdziła zależności między płcią a funkcją zmysłu powonienia [111, 134]. Jedynie nieliczni badacze obserwowali niewielkie różnice między kobietami a mężczyznami w odniesieniu do poprawy węchu po ESS. Pade i wsp. zanotowali większe prawdopodobieństwo zmniejszenia dysfunkcji zmysłu powonienia po zabiegu u kobiet, jednak tylko w grupie chorych bez polipów nosa [121]. W badaniu Minovi i wsp. kobiety uzyskały istotnie lepsze od mężczyzn wyniki w testach węchu po operacji zatoki czołowej metodą Draf III [117]. Autorzy zaznaczyli jednak, że ocena węchu polegała na werbalnym określeniu odczuwanego zapachu przed i po operacji, a zatem kobiety – jako bardziej zainteresowane tą dziedziną – mogły lepiej od mężczyzn zapamiętać zapachy prezentowane przed leczeniem.
- Alergia. Dostępne dane dotyczące wpływu alergii na zmysł powonienia nie są jednoznaczne. U alergików obserwuje się często zaburzenia węchu, na ogół o charakterze fluktuacyjnym z istotnym pogorszeniem po ekspozycji na alergen [109, 137, 138]. Badania dotyczące pacjentów z PZZP nie potwierdziły jednak istotnego wpływu alergii na dysfunkcję zmysłu powonienia ani jej poprawę po leczeniu operacyjnym w tej grupie chorych [111, 121, 125]. Część badań sugeruje nawet, że chorych z niealergicznym nieżytem nosa mogą cechować bardziej nasilone zaburzenia węchu niż alergików [134, 139].
- Palenie papierosów. Związek między paleniem papierosów a pogorszeniem zmysłu węchu jest dobrze udokumentowany [109, 140]. W przypadku pacjentów z PZZP niektóre badania potwierdziły wstępowanie znacznie większych dysfunkcji węchu u palaczy [111], podczas gdy w innych nie zanotowano takiej zależności [134]. Wpływ palenia papierosów na wyniki ESS pozostaje zatem niepewny.
- Skrzywienie przegrody nosa. Wpływ skrzywienia przegrody nosa i innych nieprawidłowości anatomicznych w obrębie jamy nosowej oraz ich leczenia chirurgicznego na zmysł węchu również pozostaje niejasny, ponieważ wyniki badań są w tym zakresie wzajemnie sprzeczne [121-123, 132, 141]. Niektórzy autorzy uważają nawet, że oddychanie przy zwiększonych oporach nosowych może prowadzić do lepszego przewodzenia zapachów do nabłonka węchowego, a co za tym idzie – mniejszego ryzyka zaburzeń powonienia [138].

### **3.6.4 Bóle głowy i twarzy**

Ból głowy i twarzy jest powszechnie uważany za jeden z najbardziej charakterystycznych objawów zapalenia zatok. **Jego związek z PZZP jest jednak aktualnie podawany w wątpliwość.** W ostatnich latach dowiedziono, że u większości chorych zgłaszających bóle głowy i twarzy, przypisywane przez nich zapaleniu zatok, w rzeczywistości można rozpoznać inne – najczęściej neurologiczne - podłoże dolegliwości [10, 142]. W związku z powyższym wiele uwagi poświęca się możliwościom odróżnienia tych chorych, u których ból rzeczywiście ma podłoże rynologiczne (tj. mogących wynieść korzyść z leczenia ukierunkowanego na PZZP) od pacjentów, u których poprawę może przynieść raczej właściwe leczenie neurologiczne [143].

Zgodnie z przytoczoną na początku niniejszego rozdziału definicją EPOS, do rozpoznania PZZP i kwalifikacji do ESS nie wystarcza jedynie obecność „charakterystycznego” bólu twarzy, ale muszą być obecne również inne objawy rynologiczne oraz zmiany patologiczne widoczne w endoskopii i/lub tomografii komputerowej [10]. Niestety kryteria te nie zawsze wystarczająco jasno różnicują przyczynę bólów głowy. Jak wiadomo, zmiany w TK bywają incydentalnie stwierdzane u osób bez jakichkolwiek objawów klinicznych ze strony zatok i mogą przypadkowo współistnieć z bólem głowy pochodzenia neurologicznego u pacjenta bez PZZP [143]. Co więcej, elementem obrazu klinicznego wielu postaci naczyniopochodnych bólów głowy (np. migreny dotyczącej twarzy, bólu klasterowego, połowiczego bólu głowy) mogą być objawy rynologiczne takie jak upośledzenie drożności nosa czy wydzielina z nosa, wynikające z pobudzenia autonomicznego układu nerwowego [10, 144]. Opisana w ostatnich latach neuralgia twarzy, mająca cechy napięciowego bólu głowy, tzw. „ból środkowego segmentu twarzy”, charakteryzuje się lokalizacją uważaną dotychczas za typową dla dolegliwości zatokopochodnych, a ponadto podobnie jak ból pochodzenia rynologicznego może nasilać się przy pochylaniu do przodu i współwystępować z zaburzeniami drożności nosa [10, 142]. Ponadto należy pamiętać, że nawet u pacjentów rzeczywiście cierpiących z powodu PZZP mogą niezależnie dodatkowo występować również bóle głowy innego pochodzenia.

Trudne zagadnienie prawidłowego **odróżnienia bólu „zatokopochodnego” od dolegliwości na innym podłożu** ma również istotne znaczenie dla prognozowania wyników operacji endoskopowych zatok przynosowych, ponieważ pacjenci faktycznie cierpiący z powodu PZZP z zasady powinni mieć większą szansę na uzyskanie poprawy po zabiegu operacyjnym. Ból wynikający z innych przyczyn teoretycznie jest na takie leczenie niepodatny. (Niestety pewne obserwacje kliniczne sugerują, że nawet ta intuicyjna reguła nie w każdym przypadku okazuje się słuszna. Zagadnienie to dodatkowo komplikuje poprawne prognozowanie, co zostanie szerzej wyjaśnione poniżej.)

Kwestia właściwej kwalifikacji pacjentów z bólami głowy i twarzy do operacji endoskopowej zatok i przewidywania skuteczności ESS w odniesieniu do likwidacji tych objawów jest zagadnieniem niezwykle istotnym także dla samych chorych. Pacjenci decydujący się na zabieg operacyjny często oczekują przede wszystkim efektu w postaci wyeliminowania tej właśnie dolegliwości. W badaniu przeprowadzonym przez Solera i wsp. chorzy

kwalifikowani do ESS ocenili ból głowy jako objaw PZZP najsilniej zaburzający ich codzienne funkcjonowanie, na drugim miejscu natomiast plasował się ból lub rozpieranie w obrębie twarzy [87].

Ocena **wpływu operacji endoskopowej zatok na uciążliwość bólu głowy i twarzy** budzi liczne kontrowersje. Wielu autorów podaje, że nasilenie zarówno bólu twarzy, jak i bólu głowy, ulega istotnemu zmniejszeniu po zabiegu endoskopowym [86, 89, 92, 95]. Dowiedziono, że pozytywny efekt leczenia może utrzymywać się przez długi czas [145], co w badaniu Senora i wsp. potwierdzono w trakcie niemal 8-letniej obserwacji [93]. We wzmiankowanym wcześniej systematycznym przeglądzie badań przeprowadzonym przez Chestera w 2009 roku wykazano, że operacja endoskopowa wywiera większy pozytywny wpływ w odniesieniu do bólów twarzy niż bólów głowy [3]. W badaniu Solera i wsp. wykazano natomiast wyłącznie zmniejszenie nasilenia bólu twarzy, podczas gdy ból głowy nie uległ zmniejszeniu po leczeniu operacyjnym [87]. Tarabichi i wsp. oceniali wyniki ESS pod kątem możliwości całkowitego wyeliminowania bólu głowy, a nie tylko jego złagodzenia. Jak wynika z ich obserwacji, dolegliwości wprawdzie prawie zawsze ustępowały bezpośrednio po zabiegu, ale po upływie roku od operacji zaledwie około 60% chorych pozostawało wolnych od bólu [143]. Należy przy tym zaznaczyć, że we wszystkich wymienionych wyżej badaniach chorzy byli kwalifikowani do zabiegu dopiero po potwierdzeniu w tomografii komputerowej obecności zmian typowych dla zapalenia zatok. Analizując przyczyny niepowodzeń ESS w odniesieniu do bólów głowy u wielu pacjentów ostatecznie potwierdzono neurologiczne podłoże dolegliwości [146, 147].

Co ciekawe, nawet chorzy cierpiący na bóle głowy pochodzenia innego niż zatokowe, mylnie zakwalifikowani do ESS, mogą również przejściowo odczuwać poprawę po zabiegu operacyjnym. Po kilku miesiącach dochodzi u nich jednak na ogół do nawrotu dolegliwości. Sądzi się, że operacja może powodować uraz aferentnych włókien czuciowych, zmiany wynikające z neuroplastyczności jąder czuciowych w obrębie pnia mózgu lub dysonans poznawczy prowadzący do przejściowego zmniejszenia odczuwania bólu [142].

Analizując dostępną literaturę, zidentyfikowano następujące czynniki wpływające na nasilenie bólu głowy i twarzy u pacjentów z PZZP oraz mogące stanowić podstawę predykcji wyników operacji endoskopowej zatok w odniesieniu do tych objawów:

- Polipy nosa. Bóle głowy, twarzy i czoła notuje się zdecydowanie częściej u chorych z PZZP bez polipów nosa niż u pacjentów z polipami. Prawidłowość tę potwierdzono jednoznacznie w licznych badaniach [46, 95, 96, 107, 127, 145, 146, 148-150]. W pracy Gigera i wsp. dotyczącej pacjentów bez polipów nosa zanotowano, że ból twarzy był najczęściej zgłaszanym przez nich objawem, a zarazem najczęściej notowano zmniejszenie jego uciążliwości na skutek operacji. Poprawa ta utrzymywała się w ciągu co najmniej 3-letniej obserwacji [91]. Mimo to ból na ogół nie ustępuje całkowicie i również po zabiegu u pacjentów bez polipów nosa jest bardziej nasilony niż u chorych z polipami nosa [46].
- Nasilenie zmian w tomografii komputerowej. Badając uwarunkowania bólu głowy i twarzy w PZZP wielu autorów stwierdziło, że zaawansowanie zmian o cechach stanu

zapalnego w tomografii komputerowej nie koreluje z nasileniem dolegliwości [100, 143]. Co więcej, również lokalizacja bólu często wydaje się nie wynikać z umiejscowienia zmian stwierdzanych w tomografii komputerowej [143, 151]. Bonfils i wsp. zauważyli jednak, że ból twarzy częściej występuje u pacjentów, u których w tomografii komputerowej stwierdza się zajęcie głównie zatoki czołowej, szczękowej i sitowia przedniego, niż u chorych z masywnymi zmianami obejmującymi wszystkie zatoki [39].

- Interesujące obserwacje dotyczące bólu w okolicy czołowej poczynili DelGaudio i wsp. [150] Wyniki ich badań mogą częściowo wyjaśniać, dlaczego w wielu pracach nie udało się potwierdzić korelacji między zaawansowaniem zmian zapalnych w tomografii komputerowej a nasileniem bólu. Zauważona przez autorów cytowanej pracy **zależność wydaje się nieliniowa**: największe nasilenie bólu czoła notowano u tych pacjentów, u których występowało umiarkowane pogrubienie śluzówki w zatoce czołowej, a mniejsze zarówno u tych, u których nie było zmian w zatoce czołowej, jak i u tych, u których była ona całkowicie zacieniona w badaniu TK. Autorzy zaproponowali hipotezę, że obserwowana prawidłowość jest związana z możliwością wyrównywania ciśnienia w zatoce. Gdy ujście jest drożne, ciśnienie wyrównywane jest bez przeszkód. Problem pojawia się w miarę narastania zmian zapalnych zewężających ujście, jednak gdy zatoka jest całkowicie wypełniona przez zmiany zapalne, nie ma w niej powietrza i stąd brak reakcji na różnice ciśnień. Tego rodzaju nieliniowa zależność jest na ogół trudna do zauważenia przy zastosowaniu tradycyjnych liniowych metod statystycznych.
- Drugim potencjalnym wyjaśnieniem braku korelacji między zmianami obserwowanymi w badaniu TK a nasileniem dolegliwości jest założenie, że w przypadku części chorych ból nie ma pochodzenie zatokowego. Zauważono, że u pacjentów zgłaszających ból twarzy jako dominującą dolegliwość bardzo często w obrazie tomografii komputerowej i badania endoskopowego zatok nie stwierdza się nieprawidłowości [94, 98, 99]. W grupie pacjentów z bólem twarzy poddanych operacji zatok mimo niestwierdzenia zmian zapalnych w obrazie TK, dolegliwości często utrzymują się mimo operacji, a dopiero wnikliwa diagnostyka pozwala wykazać neurologiczne pochodzenie bólu [147, 152].
- Nadwrażliwość na niesteroidowe leki przeciwzapalne. Ból głowy i twarzy jest objawem rzadko stwierdzanym u pacjentów z nadwrażliwością na niesteroidowe leki przeciwzapalne. Obserwacja ta znajduje uzasadnienie w kontekście opisywanych powyżej zależności między bólem a obecnością polipów nosa i rozległością zmian w tomografii komputerowej. U chorych z astmą aspirynową na ogół stwierdza się masywne polipy nosa i całkowitą bezpowietrzność zatok, co wiąże się z rzadszym występowaniem dolegliwości bólowych. [19]
- Wcześniejsze leczenie operacyjne. Związek między nasileniem bólu głowy i twarzy a przebytymi zabiegami operacyjnymi rzadko bywał przedmiotem wnikliwych analiz. Marks i Shamsa zauważyli, że pacjenci wielokrotnie operowani zgłaszali w tym zakresie nieco mniejszą poprawę po ESS niż chorzy operowani po raz pierwszy [32]. Autorzy cytowanej pracy dopatrywali się przyczyny obserwowanej zależności w

chirurgicznych uszkodzeniach struktur nerwowych lub kostnych bądź przewlekłej reakcji okostnowej.

- Płeć. Ci sami autorzy zanotowali również mniejszą poprawę w zakresie dolegliwości bólowych u kobiet niż u mężczyzn [32]. Różnicę tę tłumaczyli większą skłonnością mężczyzn do bagatelizowania swoich objawów lub różnicami anatomicznymi między kobietami a mężczyznami (mniejsze wymiary jamy nosowej kobiet mogą utrudniać kontrolę endoskopową). Nie ma jednak pewności, czy zgłaszane przez ich pacjentów bóle głowy były wyłącznie pochodzenia zatokowego – w przeciwnym razie zaobserwowaną prawidłowość można tłumaczyć np. różnicami w częstości występowania migrenowych bólów głowy w populacji zależnie od płci. Busaba i wsp. również zauważyli, że kobiety chorujące na PZZP wykazują większą tendencję do skarg na ból twarzy lub ból głowy, lecz także częściej występuje u nich PZZP bez polipów nosa (tj. postać choroby ściśle związana z tego rodzaju dolegliwościami) [102].
- Astma. Niewielu badaczy obserwowało związek astmy z nasileniem dolegliwości bólowych w PZZP, a wyniki dostępnych badań nie są jednoznaczne. Seybt i wsp. zanotowali mniejsze nasilenie bólu głowy u pacjentów z PZZP i współistniejącą astmą [101], co może wynikać z częstszego występowania polipów nosa w tej podgrupie chorych. Z kolei Kountakis i Bradley zauważyli bardziej nasilony ból głowy u astmatyków badanych 12 miesięcy po operacji endoskopowej zatok niż u pacjentów bez współistniejącej astmy [63].
- GERD. Wpływ refluksu żołądkowo-przełykowego na dolegliwości bólowe w PZZP nie jest dotychczas znany. O jego potencjalnym znaczeniu może jednak świadczyć cytowane uprzednio badanie przeprowadzone przez DiBaise i wsp., w którym stwierdzono pewną poprawę w zakresie bólu twarzy u pacjentów leczonych inhibitorem pompy protonowej [103].

### **3.6.5 Kaszel**

Schorzenia nosa i zatok przynosowych należą do najczęstszych przyczyn przewlekłego kaszlu. Znaczenie tej dolegliwości jest na tyle istotne, że w najnowszej edycji EPOS został on uwzględniony jako jeden z czterech głównych objawów definiujących PZZP u dzieci [10].

Mechanizmy odpowiedzialne za pobudzenie kaszlowe w PZZP nie są w pełni wyjaśnione. Mogą one obejmować spływanie wydzieliny po tylnej ścianie gardła, aspirację wydzieliny do dolnych dróg oddechowych lub obniżenie progu odruchu kaszlowego. Sugeruje się, że bodźce pochodzące z jamy nosowej nie są w stanie bezpośrednio wywoływać kaszlu, ale powodują zwiększenie wrażliwości na bodźce kaszlowe. Nadmierna wrażliwość na bodźce kaszlowe oznacza wywoływanie odruchu kaszlowego przez czynniki, które u osób zdrowych go nie wywołują lub zwiększoną reakcją kaszlową na bodźce o określonym nasileniu. Według tej hipotezy PZZP obniża próg reakcji na pobudzenie wynikające z innych przyczyn (np. GERD), dlatego też leczenie PZZP może zapewnić poprawę, ale niekoniecznie całkowicie wyeliminować kaszel. [153]

Hipoteza akcentująca rolę splywania wydzieliny po tylnej ścianie gardła mówi, że przyczyną kaszlu może być pobudzenie zakończeń nerwowych w obrębie gardła dolnego i krtani przez treść ściekającą z jamy nosowej przez nozdrza tylne do nosogardła [154]. Wydzielina z zatok może również być również potencjalnie aspirowana do drzewa oskrzelowego [155], powodując podrażnienie receptorów kaszlowych lub stymulację stanu zapalnego. Znaczenie tych mechanizmów jest jednak kwestionowane, ponieważ większość pacjentów, u których stwierdza się splywanie wydzieliny, nie skarży się na przewlekły kaszel [156].

Po leczeniu operacyjnym chorych z PZZP notuje się na ogół zmniejszenie nasilenia przewlekłego kaszlu [89], jednak stopień zgłaszanej przez pacjentów poprawy zwykle bywa umiarkowany w porównaniu ze zmniejszeniem niedrożności nosa czy ilości wydzieliny z nosa [86].

Zidentyfikowano następujące czynniki mogące wpływać na nasilenie kaszlu u pacjentów z PZZP:

- GERD jest uważany za jedną z najczęstszych przyczyn przewlekłego kaszlu, również u pacjentów bez towarzyszącego PZZP [157]. W patogenezie tej dolegliwości przypuszczalnie biorą udział takie mechanizmy jak stymulacja nerwu błędnego przy podrażnieniu przetyku treścią żołądkową, aspiracja kwaśnej treści lub nawet wdychanie zakwaszonego powietrza [155]. Kaszel obserwowany u chorych z PZZP i współistniejącym GERD może zatem być skutkiem głównie tego drugiego schorzenia, dlatego przypuszczalnie przy braku odpowiedniego leczenia przeciwrefluksowego po operacji zatok można u tych pacjentów oczekiwać utrzymywania się uporczywego kaszlu lub co najwyżej niewielkiej poprawy wynikającej z podwyższenia progu odruchu kaszlowego.
- Astma. Zgodnie z oczekiwaniami, większe nasilenie kaszlu obserwuje się również u pacjentów ze współistniejącą astmą [107]. Podobnie jak w przypadku GERD, również w odniesieniu do astmy można zakładać, że operacja zatok może wywierać jedynie niewielki wpływ na pobudzenie kaszlowe w tej grupie chorych, gdyż wynika ono głównie z przyczyn pozazatokowych.
- Polipy nosa. W większości badań nie stwierdzano różnic w nasileniu kaszlu między chorymi z polipami nosa i bez nich. Sahlstrand-Johnson zaobserwował jednak większą uciążliwość tego objawu u pacjentów z polipami nosa [149].

### **3.6.6 Dolegliwości ze strony ucha środkowego**

Dolegliwości ze strony ucha środkowego nie są zaliczane do głównych objawów PZZP, jednak ich występowanie u chorych z zapaleniem zatok jest na tyle typowe, że zostały wymienione w amerykańskiej definicji opracowanej przez Rhinosinusitis Task Force w 1996 roku jako jeden z tzw. objawów dodatkowych uwzględnianych przy rozpoznaniu PZZP [12]. Około 40% pacjentów z PZZP zauważa uczucie pełności w uchu lub uszach, a u 15% występuje ból ucha. Dolegliwości te są najprawdopodobniej związane z dysfunkcją trąbki

sluchowej, która może wynikać z zaburzeń transportu śluzowo-rzęskowego lub obrzęku okolicy wału trąbkowego [158]. Związek między PZZP a upośledzeniem drożności trąbki słuchowej jest szczególnie zauważalny u dzieci, u których istotną rolę często odgrywa przerost migdałka gardłowego [159-162]. PZZP stanowi również najczęstszą przyczyną występowania wysiękowego zapalenia ucha środkowego w wieku dorosłym [163].

Wykazano, że uczucie pełności w uchu i ból ucha może zmniejszać się po operacji endoskopowej zatok przynosowych [89, 158], jednak zwykle obserwuje się w tym zakresie zaledwie umiarkowaną poprawę [86].

Zagadnienie wpływu różnych czynników klinicznych na funkcję trąbki słuchowej było przedmiotem bardzo licznych badań, najczęściej dotyczących przewlekłego wysiękowego zapalenia ucha. Literatura dotycząca tej dziedziny jest w związku z tym niezwykle obszerna i jej wyczerpujący przegląd znacząco przekraczałby ramy niniejszej pracy. Dostępne piśmiennictwo podaje natomiast bardzo niewiele informacji na temat wpływu zidentyfikowanych czynników ryzyka na wyniki operacji endoskopowych zatok w odniesieniu do dolegliwości usznych. Z tego względu poniżej zostaną jedynie skrótowo zasygnalizowane pewne podstawowe uwarunkowania dysfunkcji trąbki słuchowej. Zrezygnowano natomiast ze szczegółowego omówienia zagadnień niemających bezpośredniego związku z istotnym dla niniejszej pracy problemem prognozowania wyników ESS.

- Polipy nosa. Obecność polipów w nosie wiąże się z kilkoma czynnikami mogącymi powodować dysfunkcję trąbki słuchowej, takimi jak mechaniczna blokada jej ujścia i stan zapalny błony śluzowej nosa i nosogardła. U pacjentów z zaawansowaną polipowatością nosa często stwierdza się dolegliwości uszne, takie jak ból i uczucie pełności w uchu, które mogą utrzymywać się mimo skutecznego leczenia objawów ze strony nosa [164]. Ponadto stwierdzono, że pewne formy wysiękowego zapalenia ucha (tzw. eozynofilowe zapalenie ucha środkowego) typowo współistnieją z polipami nosa, jakkolwiek patogenezą w tym przypadku nie została dotychczas wyjaśniona [165].
- Nadwrażliwość na niesteroidowe leki przeciwzapalne. Parietti-Winkler i wsp. stwierdzili, że spośród pacjentów z polipami nosa, osoby z nadwrażliwością na niesteroidowe leki przeciwzapalne 5-krotnie częściej chorowały również na wysiękowe zapalenie ucha [166]. Autorzy sugerują, że patologia ucha środkowego powinna zostać włączona do koncepcji „choroby połączonych dróg oddechowych” (*united airway disease*), dotychczas uwzględniającej głównie wzajemne zależności między obszarami zainteresowań pulmonologii i rynologii.
- Przebyte zabiegi operacyjne. U pacjentów reoperowanych z polipami nosa stwierdzono większe ryzyko rozwoju wysiękowego zapalenia ucha niż u pacjentów wcześniej nieoperowanych [166].
- GERD. Niektóre badania sugerują, że dolegliwości uszne mogą być powodowane przez refluks nadprzełykowy, również u pacjentów bez współistniejącego PZZP [167]. Wydaje się, że treść żołądkowa może przedostawać się przez trąbkę słuchową do ucha



środkowego [168]. Znaczenie refluku i jego leczenia dla patologii ucha pozostaje jednak kwestią dyskusyjną [169].

- Palenie papierosów. W badaniach eksperymentalnych na modelu zwierzęcym wykazano, że narażenie na dym tytoniowy powoduje zmiany histologiczne oraz zwiększoną produkcję śluzu w trąbce słuchowej [170]. Dzieci palących rodziców są bardziej narażone na rozwój przewlekłego wysiękowego zapalenia ucha [171]. Również u osób dorosłych palenie papierosów może prowadzić do przerostu tkanki chłonnej w nosogardle, a co za tym idzie do rozwoju wysiękowego zapalenia ucha [163]. Wpływ dymu tytoniowego na dolegliwości uszne u chorych z PZZP a także na wyniki ESS nie jest jednak jednoznacznie określony.
- Skrzywienie przegrody nosa. Deron i wsp. potwierdzili, że korekta skrzywienia przegrody może poprawiać nie tylko drożność nosa, lecz również funkcję trąbki słuchowej [172].
- Alergia wydaje się odgrywać istotną rolę w dysfunkcji trąbki słuchowej [173], jakkolwiek jej znaczenie dla patologii ucha środkowego nie zostało dotychczas w pełni udowodnione [174, 175]. W przypadku PZZP wpływ współistniejącej alergii na dolegliwości ze strony ucha środkowego wydaje się wątpliwy. W badaniu [166] dotyczącym pacjentów z polipami nosa nie stwierdzono zależności między alergią a nasileniem dolegliwości usznych w PZZP, podczas gdy w pracy de Loos i wsp. pacjenci z alergią zgłaszali mniejsze zaburzenia słuchu i uczucie pełności w uchu niż chorzy bez alergii.

### **3.6.7 Ból zębów**

Zgodnie z definicją podaną przez Rhinosinusitis Task Force, ból zębów, podobnie jak dolegliwości uszne, jest zaliczany do objawów dodatkowych PZZP [12]. W badaniu Solera i wsp. dolegliwość tę zgłaszało aż 67% chorych, jakkolwiek miała ona najmniejsze nasilenie ze wszystkich objawów PZZP [87]. W tej samej pracy zanotowano zmniejszenie nasilenie bólu zębów po operacji endoskopowej zatok, jednak poprawa ta nie była trwała i po 6-miesięcznej obserwacji doszło do ponownego nasilenia dolegliwości [87]. Ling i Kountakis stwierdzili złagodzenie bólu zębów po ESS w stopniu porównywalnym do innych objawów [89], podczas gdy Bhattacharyya uznał, że jest to objaw w najmniejszym stopniu ulegający poprawie na skutek leczenia operacyjnego zatok przynosowych [86].

Poszukiwanie czynników wpływających na nasilenie izolowanego objawu, jakim jest ból zębów, u chorych operowanych z powodu PZZP, wydaje się nie budzić znaczącego zainteresowania badaczy. Wnikliwy przegląd literatury pozwolił wyodrębnić tylko jedno schorzenie towarzyszące, którego wpływ na nasilenie bólu zębów po operacji endoskopowej zatok został udokumentowany:

- Astma. Kountakis i Bradley zaobserwowali, że u pacjentów z astmą uciążliwość bólu zębów ulega zmniejszeniu po ESS w mniejszym stopniu niż u pozostałych chorych [63].

### **3.6.8 Nieprzyjemny zapach z ust**

Kolejnym objawem dodatkowym występującym w PZZP wg Rhinosinusitis Task Force jest nieprzyjemny zapach z ust [12]. Jego wystąpienie może się wiązać z rozwojem bakterii w wytwarzanej w nadmiarze wydzielinie spływającej po tylnej ścianie gardła, a także z upośledzeniem drożności nosa i nadmiernym wysychaniem śluzówek jamy ustnej prowadzącym do zaburzeń ich oczyszczania [176]. Jak wykazują badania, również ta dolegliwość na ogół staje się mniej uciążliwa po operacji endoskopowej zatok [89], jednak stopień obserwowanej poprawy jest umiarkowany [86].

Spośród nielicznych zidentyfikowanych czynników wpływających na przebieg PZZP i jednocześnie potencjalnie modyfikujących nasilenie nieprzyjemnego zapachu z ust można wymienić:

- GERD i palenie papierosów. Nieprzyjemny zapach z ust należy do częstych problemów nękających pacjentów z chorobą refluksową przełyku oraz palaczy, również bez towarzyszącego PZZP [177].
- Astma. Kountakis i Bradley zauważyli większą poprawę w odniesieniu do omawianej dolegliwości po ESS u pacjentów z astmą. Autorzy nie zaproponowali jednak żadnej hipotezy mogącej tłumaczyć zaobserwowany fakt [63].
- Ponadto nieprzyjemny zapach z ust może być powodowany przez rozmaite przyczyny nie mające bezpośredniego związku z przebiegiem PZZP, takie jak niewłaściwa dieta i higiena jamy ustnej, choroby zębów i przyzębia, zaburzenia produkcji śliny, schorzenia gardła i migdałków podniebiennych, zaburzenia metaboliczne oraz przewlekłe choroby płuc, nerek i wątroby [176].

### **3.6.9 Zmęczenie**

Obecność PZZP w wielu przypadkach wiąże się z uczuciem przewlekłego zmęczenia i osłabienia, które również jest traktowane jako jeden z dodatkowych objawów tego schorzenia wg definicji Rhinosinusitis Task Force [12]. Ponad połowa pacjentów z PZZP skarży się na przewlekłe zmęczenie, jednak przyczyny występowania tej dolegliwości nie są jasne [178]. Po operacji endoskopowej zatok na ogół obserwuje się poprawę w odniesieniu do omawianego objawu [86, 87, 89, 178, 179]. Analizując dostępną literaturę zidentyfikowano następujące czynniki potencjalnie modyfikujące nasilenie przewlekłego zmęczenia u chorych z PZZP:

- Płeć. U kobiet zanotowano wyższy poziom przewlekłego zmęczenia zgłaszanego przed operacją [179].
- Polipy nosa. Pacjenci z polipami nosa w mniejszym stopniu skarżyli się na przewlekłe zmęczenie niż pacjenci bez polipów nosa [179].
- Nasilenie zmian zapalnych w obrazie tomografii komputerowej wykonanej przed operacją korelowało ze zgłaszaną przez pacjentów uciążliwością przewlekłego

zmęczenia towarzyszącego PZZP w badaniu, które przeprowadzili Kenny i wsp. [100], jednak badania Sauttera i wsp. nie potwierdziły tej zależności [179].

- Nasilenie dolegliwości przed zabiegiem. Pacjenci zgłaszający większą uciążliwość przewlekłego zmęczenia przed operacją podawali bardziej zadowalającą poprawę w tym zakresie po zabiegu niż pozostali chorzy [179].

### 3.7 Modele predykcyjne i sztuczne sieci neuronowe w PZZP

Poszukiwanie możliwości przewidywania wyników leczenia operacyjnego chorych z PZZP w odniesieniu do indywidualnych pacjentów przy pomocy wielowymiarowych modeli predykcyjnych jest – na ile się to udało ustalić na podstawie dostępnego piśmiennictwa - tematem rzadko poruszonym w literaturze. Modele regresyjne i model proporcjonalnego hazardu Coxa bywają używane raczej do identyfikacji i oceny wpływu poszczególnych czynników ryzyka na przebieg choroby niż do celów prognostycznych, w związku z czym nie podlegają ewaluacji jako narzędzia prognostyczne [29, 50, 83, 111, 180, 181].

**Opisane w dostępnym piśmiennictwie statystyczne modele budowane z założeniem potencjalnej roli predykcyjnej są bardzo nieliczne.** Sil, Mackay i Rowe-Jones zaproponowali model regresyjny służący do przewidywania, którzy z operowanych pacjentów w ciągu 5 lat po ESS mogą wymagać włączenia leczenia sterydami doustnymi i antybiotykami z powodu nawrotu dolegliwości o intensywności podobnej jak przed operacją [37]. W badanej grupie liczącej 109 pacjentów zidentyfikowano 12 chorych spełniających tak zdefiniowane kryterium niepowodzenia leczenia. Podstawą predykcji było nasilenie zmian w tomografii komputerowej i wielkość polipów nosa. Prawdopodobieństwo prawidłowej prognozy dla grupy badanej wynosiło 81.7%, nie podano jednak danych dotyczących próby zastosowania modelu w celu dokonania predykcji u nowych pacjentów. Katotomichelakis i wsp. [182] posłużyli się wieloczynnikową regresją liniową w celu przewidywania wyników obiektywnych badań węchu oraz i regresją logistyczną dla predykcji prawdopodobieństwa osiągnięcia prawidłowej funkcji węchu 6 miesięcy po ESS. Pierwszy z zaproponowanych modeli prawidłowo wyjaśniał około 70% obserwowanej zmienności, drugi natomiast właściwie klasyfikował 76% pacjentów. Podobnie jak w poprzedniej pracy, podano wyłącznie wyniki uzyskane dla danych wykorzystanych przy tworzeniu modelu. Zgodnie z ogólnie przyjętymi zasadami, ewaluacja jakości modelu powinna być przeprowadzona na podstawie analizy nowych przypadków [183], bo tylko taka walidacja modelu pozwala ocenić jego zdolność do generalizacji i wykluczyć ewentualne nadmierne dopasowanie do danych uczących.

W dostępnej literaturze nie odnaleziono natomiast informacji o modelach wieloczynnikowych służących do przewidywania subiektywnej poprawy po leczeniu operacyjnym chorych z PZZP w odniesieniu do innych niż zaburzenia węchu objawów, będących przedmiotem zainteresowania niniejszej pracy.

Z kolei analizując dostępne piśmiennictwo pod kątem **zastosowania sztucznych sieci neuronowych do diagnostyki i terapii PZZP** znaleziono tylko jedno streszczenie [184], nie mające jednak związku z podejmowanym w tej rozprawie zagadnieniem przewidywania wyników leczenia. Cytowane badanie dotyczyło próby wykorzystania sztucznej sieci neuronowej w automatycznej diagnostyce PZZP na podstawie analizy chemicznej próbek powietrza wydychanego przez pacjentów.

**Wyniki przeglądu dostępnego piśmiennictwa sugerują, że temat prognozowania wyników leczenia operacyjnego chorych z PZZP z wykorzystaniem sztucznych sieci neuronowych nie był dotychczas przedmiotem żadnych badań.**

### 3.8 Podsumowanie

W przedstawionym powyżej przeglądzie piśmiennictwa zwraca uwagę fakt, że czynniki, dla jakich w niektórych badaniach stwierdzano statystycznie istotną korelację z wynikiem operacji, nie miały znaczącego wpływu na wyniki leczenia dla innych grup chorych. Obserwowane niezgodności mogą wynikać z następujących przyczyn:

- Poszczególne badania dotyczą niejednakowych grup chorych, zróżnicowanych pod względem zaawansowania choroby, częstości występowania chorób towarzyszących itd. Można założyć, że w przypadku badań, w których liczebność pewnych podgrup pacjentów była niewielka, analizy nie osiągały progu istotności statystycznej mimo występowania określonych zależności przyczynowo-skutkowych w populacji.
- Brak jednolitej definicji „sukcesu leczenia”. Efekt leczenia oceniany na podstawie ustępowania objawów subiektywnych mierzonych przy pomocy różnych narzędzi pomiarowych może zależeć od innych czynników, niż wyniki badań obiektywnych lub kwalifikacja do reoperacji.
- Wyniki badań naukowych przeprowadzonych przez różnych autorów dotyczą różnie długiego okresu obserwacji po zabiegu operacyjnym.
- W większości badań analiza statystyczna dotyczyła każdego z czynników prognostycznych niezależnie, natomiast nie uwzględniała ich synergii - łącznego oddziaływania czynników występujących w określonych „konstelacjach” w poszczególnych grupach chorych. Jedynie nieliczne prace wykorzystujące wieloczynnikowe modele statystyczne (np. regresję wieloczynnikową, model proporcjonalnego hazardu Coxa) uwzględniają ten problem.

Analiza dostępnego piśmiennictwa sugeruje, że wynik leczenia chorych z PZZP u każdego z nich indywidualnie zależy od wzajemnych relacji wielu działających równocześnie czynników. Przykładem może być opisany powyżej wpływ alergii na wynik leczenia. Wpływ ten miał znaczenie rokownicze u chorych z masywnymi polipami nosa [48], natomiast w licznych innych badaniach dotyczących bardziej zróżnicowanych grup chorych nie okazał się

istotny. Wyniki te sugerują, że predykcji wyników leczenia korzystniej jest dokonywać analizując wpływ obu tych zmiennych (tj. obecności alergii i wielkości polipów nosa) łącznie, niż każdej z nich niezależnie. Ponadto wydaje się, że część zależności między czynnikami ryzyka a nasileniem dolegliwości (np. między zakresem zmian w tomografii komputerowej a intensywnością bólu czoła) ma charakter nieliniowy. **Wieloczynnikowa analiza uwzględniająca potencjalnie nieliniowe zależności między zmiennymi jest trudna przy użyciu standardowych metod statystycznych, ale jest możliwa do przeprowadzenia z zastosowaniem sztucznych sieci neuronowych.** Jednak w dostępnej literaturze nie natrafiono na żadną wzmiankę dotyczącą tego rodzaju badań prowadzonych przez innych autorów.

Ponadto należy podkreślić, że większość cytowanych prac dotyczy oceny *ex post* wpływu poszczególnych czynników ryzyka na ostateczny wynik leczenia operacyjnego, natomiast stosunkowo nieliczne są informacje o próbach antycypacji - przewidywania z wyprzedzeniem wyniku leczenia u poszczególnych chorych na podstawie obecności lub braku określonych czynników ryzyka. W szczególności stwierdzono brak doniesień o zastosowaniu w tym celu sztucznych sieci neuronowych. Ustalenie tych faktów stanowi podstawę do twierdzenia, że **temat wybrany do zbadania i naukowego opracowania w ramach tej rozprawy jest oryginalny, a wybrana metoda jego rozwiązania w postaci modelu wykorzystującego uczące się sieci neuronowe nie była dotychczas stosowana w odniesieniu do analizowanego zagadnienia.** Z kolei potwierdzona przydatność sieci neuronowych dla wielu innych problemów związanych z koniecznością prognozowania w różnych dziedzinach, w tym również w medycynie, skłania do zbadania użyteczności właśnie tego narzędzia informatycznego w kontekście rozważanego w tej pracy problemu prognozowania skutków leczenia chorych z PZZP.

## 4. Sztuczne sieci neuronowe – wprowadzenie i przegląd piśmiennictwa

### 4.1 Wprowadzenie

Sztuczne sieci neuronowe są systemami informatycznymi o unikatowych możliwościach, służącymi do przetwarzania danych na podobnych zasadach, jak ma to miejsce w układach nerwowych istot żywych [185]. Bywają wykonywane w formie specjalistycznych systemów elektronicznych jako tak zwane neurokomputery, najczęściej jednak mają formę programów komputerowych symulujących działanie neurokomputera na zwykłym komputerze. Powstały one w wyniku badań w dziedzinie **sztucznej inteligencji**, dotyczących modelowania podstawowych struktur występujących w mózgu oraz ich funkcji [186]. Mimo, że są to modele bardzo uproszczone (w mózgu człowieka występuje około  $10^{15}$  połączeń między komórkami nerwowymi, podczas gdy na sztuczną sieć neuronową składa się na ogół od kilkuset do kilkudziesięciu tysięcy połączeń) [187], pod wieloma względami posiadają potencjał inteligentnej analizy informacji nieosiągalny dla tradycyjnych metod matematycznych i informatycznych. Do cech wyróżniających sztuczne sieci neuronowe należy między innymi zdolność samodzielnego odnajdywania reguł i zależności metodą uczenia się na podstawie przykładów.

### 4.2 Zastosowania sztucznych sieci neuronowych – przegląd piśmiennictwa

Sieci neuronowe są wykorzystywane do realizacji zadań takich jak przetwarzanie i analiza danych, ich predykcja i klasyfikacja oraz sterowanie [186, 187]. Zakres praktycznych zastosowań sieci neuronowych jest niezwykle szeroki. Aktualnie udane aplikacje modeli neuronowych obejmują nie tylko obszar zainteresowań inżynierii, finansów czy fizyki, jak miało to miejsce w początkowej fazie ich rozwoju, ale także medycyny i biologii. Po ponad pięciu dziesięcioleciach nieustannego wzrostu popularności tej techniki przypuszczalnie nie istnieje obecnie żadna dziedzina, w której nie zostałyby one dotychczas z mniejszym lub większym sukcesem wykorzystane [187].

Wyczerpujący przegląd literatury dotyczącej zastosowań sieci neuronowych (nawet wyłącznie w dziedzinach medycznych) musiałby przypuszczalnie liczyć setki tysięcy pozycji. Ze względu na tematykę niniejszej pracy poniżej zostaną krótko wspomniane jedynie te przykłady aplikacji, które bezpośrednio wiążą się z zagadnieniem prognozowania (tj. przewidywania przyszłych zdarzeń na podstawie faktów z przeszłości) bądź dotyczą obszaru zainteresowań otolaryngologii.

**Prognozowanie z wykorzystaniem modeli neuronowych** wykazało skuteczność w tak zróżnicowanych obszarach jak przewidywanie przyszłych cen akcji na podstawie wcześniejszych notowań giełdowych [188], prognozowanie zapotrzebowania na energię elektryczną dostarczaną przez elektrownię [189], prognozowanie rozwoju przedsiębiorstw, poszukiwania ropy naftowej czy nawet typowanie na wyścigach konnych [187]. Sztuczne sieci neuronowe były również szeroko wykorzystywane w medycynie, między innymi do prognozowania przebiegu choroby i wyników leczenia w odniesieniu do nowotworów [190-195] chorób zakaźnych [196], chorób serca [197, 198] i wielu innych.

Omawiając dotychczasowe **zastosowania sieci neuronowych w otolaryngologii** można wyróżnić trzy zasadnicze grupy zagadnień. Poza wspomnianym wyżej przewidywaniem efektów terapii należy do nich wspomaganie diagnostyki oraz automatyczna interpretacja sygnałów medycznych [199].

Podobnie jak w innych dziedzinach medycyny, również w otolaryngologii szczególne zainteresowanie budzą możliwości **przewidywania wyników leczenia** chorych onkologicznych. Prace dotyczące tej dziedziny są jednak jak dotąd stosunkowo nieliczne.

Jones i wsp. [200] podjęli próbę przewidywania czasu przeżycia pacjentów leczonych z powodu raka płaskonabłonkowego krtani. Podstawę dla predykcji stanowiły informacje dotyczące wieku i płci pacjentów, ich stanu ogólnego, lokalizacji guza, jego cech histopatologicznych i stopnia zaawansowania klinicznego. Wyniki uzyskane przez model neuronowy okazały się w tym przypadku porównywalne z wynikami otrzymanymi z zastosowaniem tradycyjnych metod statystycznych (model Coxa, krzywa Kaplana-Meiera), jednak sieć neuronowa wykazała większą wrażliwość na różnice wieku chorych i stopień zaawansowania regionalnego nowotworu.

Bryce i wsp. [201] posłużyli się sztuczną siecią neuronową do przewidywania skuteczności radioterapii i radiochemioterapii zaawansowanych nowotworów głowy i szyi. Celem ich modelu było zidentyfikowanie grupy pacjentów, u których sama radioterapia nie zapewnia akceptowalnego poziomu szans przeżycia, w związku z czym uzasadnione jest zastosowanie bardziej agresywnych metod leczenia. W przypadku tego zagadnienia wykazano przewagę sieci neuronowych nad regresją logistyczną.

Jak dotąd bardzo niewiele badań odnosiło się do zastosowań predykcyjnych modeli neuronowych w otolaryngologii do przewidywania wyników leczenia chorób nienowotworowych, w tym zwłaszcza niezwykle istotnych dla tej dziedziny przewlekłych schorzeń zapalnych. Przegląd dostępnej literatury wskazuje, że pierwsze prace dotyczące tego zagadnienia zostały przeprowadzone przez autorkę niniejszej pracy wraz z zespołem Kliniki Otolaryngologii w Krakowie. Ich celem było zaprojektowanie sztucznych sieci neuronowych służących do prognozowania poprawy słuchu po leczeniu operacyjnym u chorych z przewlekłym zapaleniem ucha środkowego [202, 203]. Na podstawie danych dotyczących stanu ucha przed zabiegiem operacyjnym oraz przebiegu zabiegu możliwe okazało się przewidywanie stanu słuchu po upływie roku od operacji.

Analizując literaturę przedmiotu nie odnaleziono natomiast żadnych publikacji odnoszących się do prognozowania wyników leczenia schorzeń zapalnych zatok przynosowych.

Liczne zastosowania sieci neuronowych w otolaryngologii dotyczyły natomiast **wspomagania diagnostyki klinicznej**. W 2009 roku Williams opatentował urządzenie bazujące na modelu neuronowym, służące do badań przesiewowych w kierunku alergii [204]. Wstępne rozpoznanie schorzenia stawiane było na podstawie wywiadu, wyników skórnych testów punktowych dla wybranych alergenów oraz poziomu specyficznych IgE. Uczenie sieci przeprowadzono w oparciu o dane kilkudziesięciu pacjentów, dla których znane były diagnozy postawione przez doświadczonego lekarza. Zadaniem urządzenia było wytypowanie chorych mogących wymagać konsultacji alergologa oraz identyfikacja przypadków nie wymagających zaangażowania specjalisty.

Wspomaganiem diagnostyki otolaryngologicznej przy pomocy modeli neuronowych zajmowali się również Juhola i wsp. [205]. Zbudowane przez nich modele samoczynnie dokonywały klasyfikacji chorych cierpiących na zawroty głowy na podstawie charakteru zgłaszanych dolegliwości. Sieci neuronowe w procesie samouczenia nabyły zdolności podawania prawidłowych wstępnych rozpoznań dla schorzeń najczęściej występujących w badanej grupie pacjentów (łagodne położeniowe zawroty głowy, choroba Meniere'a, *neuronitis vestibularis* i nerwiak nerwu przedsionkowego), natomiast nie były w stanie prawidłowo zaklasyfikować chorych manifestujących rzadziej występujące jednostki chorobowe.

Sztuczne sieci neuronowe zostały również zastosowane w diagnostyce bezdechu sennego: el-Solh i współpracownicy zaproponowali metodę przewidywania niezwykle istotnego dla rozpoznania i oceny stopnia ciężkości tego schorzenia wskaźnika AHI (bezdech/spływanie oddechu). Predykcji dokonywano na podstawie danych klinicznych i pomiarów antropometrycznych pacjenta, co mogłoby stanowić alternatywę dla uciążliwego i kosztownego badania polisomnograficznego [206].

Szeroko eksplorowanym obszarem zastosowań sieci neuronowych w otolaryngologii była także **automatyczna interpretacja sygnałów medycznych**. Modele neuronowe okazały się użyteczne w automatycznej klasyfikacji sygnałów ABR (słuchowych potencjałów wywołanych pnia mózgu) [207], TEOEA (otoemisji akustycznych wywołanych trzaskiem) w ramach przesiewowych badań słuchu u noworodków [208] i DPOEA (produktów zniekształceń nieliniowych ślimaka) w diagnostyce niedosłuchu odbiorczego [209]. Ponadto sztuczne sieci neuronowe wykorzystywano do rozpoznawania dysfonii na podstawie wyników badania elektrolottograficznego [210] lub analizy akustycznej sygnału mowy [211, 212] (tego ostatniego tematu dotyczyły między innymi badania prowadzone przez autorkę niniejszej pracy w ramach działalności Studenckiego Koła Naukowego przy Klinice Otolaryngologii w Krakowie [213]).



### 4.3 Modelowanie z wykorzystaniem sieci neuronowych

Sieci neuronowe są zaawansowaną techniką modelowania komputerowego [214]. Mają one zdolność odwzorowywania bardzo złożonych zależności między sygnałami wejściowymi i wyjściowymi (również nieliniowych). Spośród innych metod modelowania komputerowego wyróżnia je między innymi zdolność stosunkowo łatwego odwzorowywania funkcji nieliniowych z dużą liczbą zmiennych niezależnych (tzw. funkcji wektorowych) [186].

Jedną z unikatowych cech sieci neuronowych jest możliwość **modelowania bez konieczności formułowania wstępnych hipotez**. [186] Oznacza to, że można je stosować w przypadku takich zagadnień, gdzie istnieją przesłanki przemawiające za istnieniem zależności między pewnymi sygnałami wejściowymi a wartościami wynikowymi, ale charakter tej zależności nie jest badaczowi znany [189]. W przypadku modelowania zjawisk medycznych ta właściwość ma o tyle istotne znaczenie, że patogenezę większości schorzeń nie jest na tyle poznana, aby umożliwić jednoznaczne wnioskowanie na temat charakteru (a co za tym idzie - matematycznego opisu) zależności przyczynowo-skutkowych np. pomiędzy czynnikami modulującymi przebieg choroby a nasileniem objawów czy wynikiem leczenia.

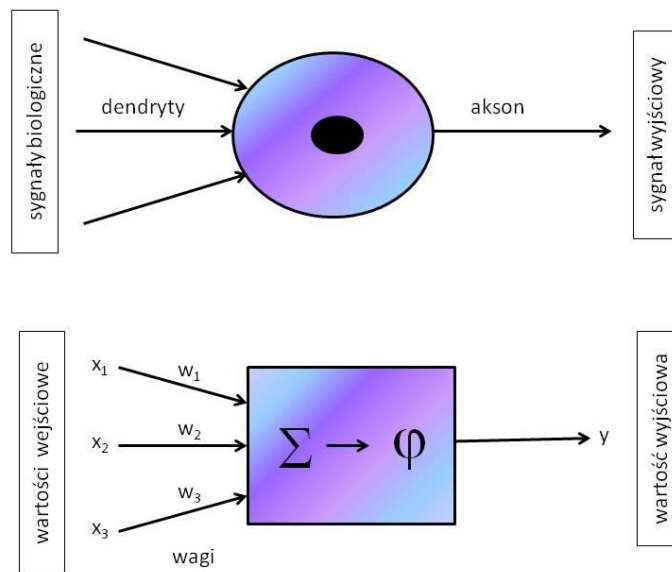
Sieci neuronowe **samoczynnie** konstruują modele, automatycznie ucząc się na podstawie podanych przez użytkownika przykładów. Zadaniem badacza jest zatem głównie zgromadzenie reprezentatywnych danych, zawierających informacje, jak manifestuje się zależność, która ma być modelowana [186]. Ponadto do badacza należy wybór danych, które mogą posłużyć do modelowania jako dane wejściowe lub sygnały wyjściowe modelu, oraz ich matematyczne zakodowanie w sposób umożliwiający wykorzystanie przez sieci neuronowe, kontrola prawidłowości procesu uczenia i krytyczna ocena uzyskanych wyników.

W związku z łatwością i wygodą stosowania oraz dzięki dostępności odpowiedniego oprogramowania sieci neuronowe stały się narzędziem dostępnym dla wielu badaczy. Dotyczy to użytkowników posiadających odpowiednią wiedzę dotyczącą modelowanych problemów (np. zagadnień medycznych) oraz w zakresie specyfiki wykorzystania sieci neuronowych. Ważna jest także pełna dostępność tej techniki dla badaczy nie posiadających formalnego wykształcenia w zakresie matematyki czy informatyki. Takie wykształcenie (lub stała konsultacja specjalisty) byłoby niezbędne w przypadku konieczności samodzielnego budowania modeli od podstaw. Sieci neuronowe są tu więc czynnikiem znacząco ułatwiającym przejście od obserwacji medycznych do komputerowego modelu procesów leżących u podstaw tych obserwacji, a także przejście od wyników komputerowych obliczeń do ich interpretacji i wykorzystania w konkretnych potrzebach medycznych (w zakresie diagnostyki, prognozy, optymalizacji terapii, kontroli postępów procesu leczenia i ewentualnego wspomagania rehabilitacji).

## 4.4 Budowa i zasada działania sieci neuronowych

Podstawową jednostką budującą sieć neuronową jest **sztuczny neuron**, stanowiący bardzo uproszczony model neuronu biologicznego – rys. 11 [185-187, 189].

**Rysunek 11. Porównanie neuronu biologicznego i sztucznego neuronu.**



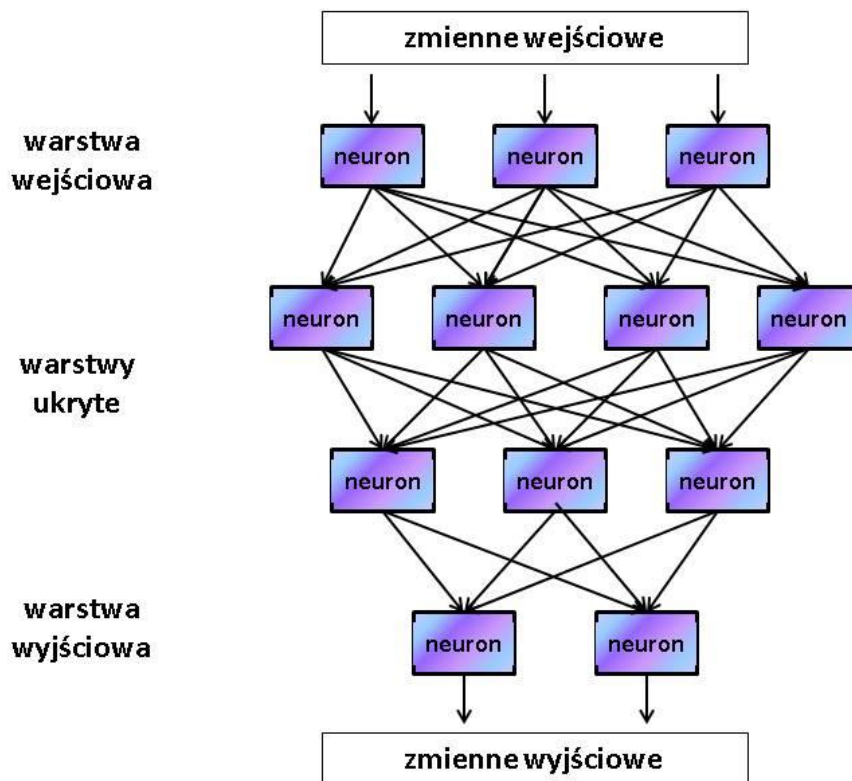
Zasada działania sztucznego neuronu jest następująca [186]:

- wartości wejściowe ( $x_1, x_2, \dots, x_n$ ) docierają do sztucznego neuronu podobnie jak sygnały biologiczne docierają do neuronu biologicznego poprzez dendryty;
- każda z wartości wejściowych wprowadzana do neuronu mnożona jest przez liczbę charakteryzującą dane połączenie (są to tzw. wagi -  $w_1, w_2, \dots, w_n$ ) o wartości dodatniej lub ujemnej, co stanowi pewną analogię do silniejszego lub słabszego oddziaływania na neuron poszczególnych sygnałów wejściowych oraz do obecności synaps pobudzających lub hamujących w układzie nerwowym;
- w neuronie obliczana jest ważona suma wejść (tj. suma wartości wejściowych przemnożonych przez odpowiadające im wagi);
- wyliczone w wyżej opisany sposób tak zwane łączne pobudzenie neuronu (odpowiednik potencjału postsynaptycznego w neurobiologii), przekształcane następnie przez odpowiednio dobraną funkcję aktywacji (funkcję przejścia  $\phi$ ) neuronu,
- wartość otrzymana w wyniku powyższych przekształceń stanowi wartość wyjściową (sygnał wyjściowy), opuszczający neuron przez jego wyjście, stanowiące odpowiednik biologicznego aksonu.

## 4.5 Architektura sieci

W obrębie sieci sztuczne neurony są ze sobą połączone [185]. W najczęściej stosowanych sieciach neuronowych (tzw. perceptronach wielowarstwowych – MLP, *multilayer perceptrone*) są one ułożone w **warstwy**. Pierwsza z nich, tzw. warstwa wejściowa, służy do wprowadzania danych do sieci. Kolejne warstwy biorące udział w przetwarzaniu danych noszą nazwę ukrytych. Ostatecznie sygnały przekazywane są do warstwy wyjściowej, której neurony generują sygnały interpretowane jako wyniki działania sieci i wykorzystywane jako rozwiązania zadania postawionego sieci przez użytkownika [187, 189]. Neurony wejściowe można zatem porównać do neuronów odbierających sygnały od receptorów (na przykład komórki zwoju spiralnego w uchu wewnętrznym), a neurony wyjściowe do neuronów sterujących aktywnością efektorów (na przykład motoneuronów sterujących pracą mięśni). Neurony ukryte pełnią rolę analogiczną do struktur nerwowych przetwarzających sygnał z receptorów na decyzje uruchamiające odpowiednią odpowiedź efektorów. Na rysunku 12 przedstawiono przykładową sieć neuronową, posiadającą dwie warstwy ukryte.

**Rysunek 12. Przykład sieci neuronowej – perceptron wielowarstwowy typu *feedforward* posiadający trzy neurony w warstwie wejściowej, dwie warstwy ukryte (zawierające odpowiednio cztery i trzy neurony) i dwa neurony w warstwie wyjściowej.**



Zazwyczaj neurony łączone są ze sobą w ten sposób, że w obrębie sieci sygnał przekazywany jest tylko w jednym kierunku, tj. od neuronów wejściowych, poprzez kolejne warstwy ukryte, do neuronów wyjściowych (jak na rysunku 12 Rysunek 12). Są to tzw. sieci jednokierunkowe (*feedforward*). Tego rodzaju sieci ze względu na stabilne zachowanie najczęściej znajdują zastosowania praktyczne, między innymi w medycynie. Połączenia między warstwami tworzone są zwykle wg zasady „każdy z każdym”, tj. tak, aby każdy neuron następnej warstwy był połączony z każdym z neuronów poprzedniej warstwy [186, 189].

W licznych badaniach dotyczących sieci neuronowych wykorzystywane są jednak również sieci o bardziej skomplikowanej strukturze, w szczególności posiadające sprzężenia zwrotne, czyli wsteczne połączenia biegnące od neuronów „dalszych” do „bliższych”. Szczególnego rodzaju sieci, w których wszystkie połączenia mają charakter sprzężeń zwrotnych, tzw. sieci Hopfielda, wykorzystywane są między innymi jako techniczne odpowiedniki pamięci asocjacyjnej. [186, 187]

Warto wspomnieć, że istnieją różne typy sieci neuronowych różniące się między innymi pod względem zastosowanych funkcji aktywacji neuronów. W warstwach ukrytych opisanego powyżej perceptronu wielowarstwowego (MLP, *multilayer perceptrone*) znajdują się na ogół neurony posiadające funkcje aktywacji o charakterze sigmoidalnym. W związku z tym działanie sieci polega na podziale przestrzeni sygnałów wejściowych za pomocą hiperpłaszczyzn - wielowymiarowych odpowiedników linii prostych dzielących przestrzeń dwuwymiarową lub płaszczyzn dzielących przestrzeń trójwymiarową na obszary odpowiadające np. poszczególnym wyodrębnianym przez sieć klasom. Istnieją jednak również sieci mające zdolność dokonywania podziału przestrzeni przy użyciu innych powierzchni granicznych, na przykład hipersfer (wielowymiarowych odpowiedników okręgów na płaszczyźnie lub sfer w trójwymiarowej przestrzeni). Powierzchnie takie uzyskuje się najczęściej dzięki zastosowaniu radialnych funkcji bazowych. Odpowiednie sieci określane są skrótem RBF – *Radial Basis Function*. Różnice w działaniu tych dwóch typów sieci zobrazowano na rysunku 13. Neurony radialne w warstwie ukrytej posiadają również sieci neuronowe specyficznego rodzaju, których zadaniem jest wyznaczenie prawdopodobieństw przynależności sygnałów wejściowych do określonych klas – tzw. sieci probabilistyczne (PNN – *probabilistic neural networks*) oraz analogicznie funkcjonujące sieci realizujące regresję uogólnioną (GRNN – *generalized regression neural networks*). [215]

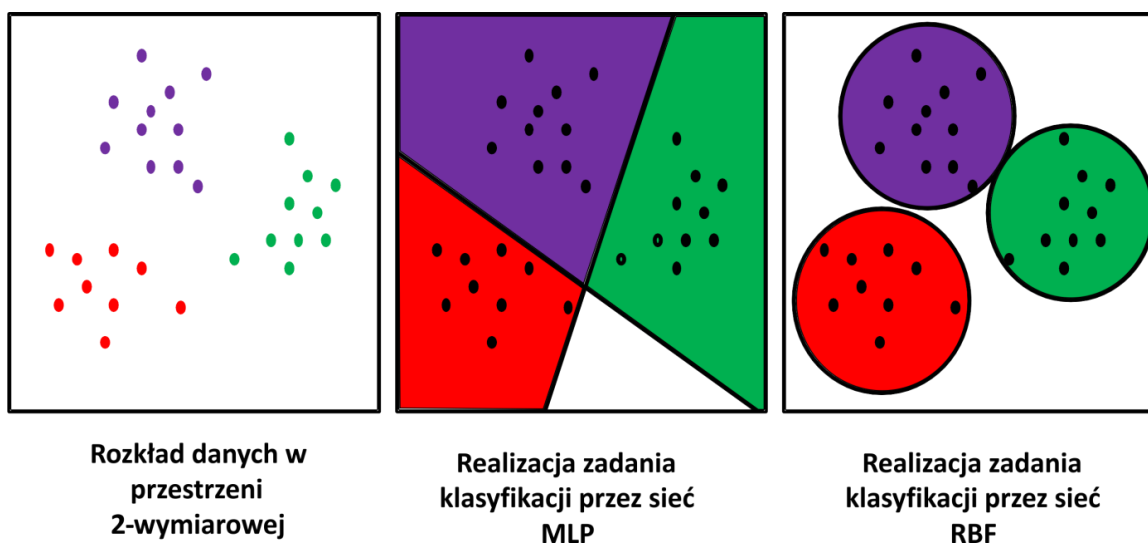
W dalszej części rozdziału omówione będzie działanie najczęściej stosowanego typu sieci, tj. perceptronu wielowarstwowego (MLP).

Odpowiedni wybór typu i architektury sieci dla danego zadania, w tym **liczby warstw i liczby neuronów** w poszczególnych warstwach – należy do użytkownika i najczęściej przeprowadzany jest metodą empiryczną, ponieważ nie istnieją w tym względzie ścisłe reguły. Trzeba przy tym zważać na fakt, że zbyt prosta sieć może nie być w stanie modelować skomplikowanego zadania, podczas gdy sieć o zbyt dużych rozmiarach traci zdolność generalizacji (wykorzystywania wiedzy zdobytej podczas uczenia do rozwiązywania nowych zagadnień, pojawiających się podczas późniejszej eksploatacji modelu), ponieważ ma skłonność do zbyt dokładnego dopasowania się do zbioru uczącego [189]. Zagadnienie to jest

szerzej omówione w podrozdziale dotyczącym wielkości zbioru uczącego i liczby zmiennych wejściowych, warunkującej rozmiar pierwszej warstwy sieci, a także w podrozdziale dotyczącym generalizacji.

Liczba warstw ukrytych w sieci neuronowej może być w zasadzie dowolna, jednak można udowodnić, że perceptron wielowarstwowy z dwoma warstwami ukrytymi teoretycznie jest w stanie modelować dowolny problem [216]. Z tego względu sieci neuronowe o większej liczbie warstw w praktyce nie są używane. Jak się okazuje, większość praktycznych problemów można rozwiązać z pomocą sieci korzystającej z wyłącznie jednej warstwy ukrytej.

**Rysunek 13. Porównanie zasady działania perceptronu wielowarstwowego (MLP) i sieci radialnej (RBF).**



## 4.6 Uczenie sieci neuronowej

### 4.6.1 Założenia i cele uczenia sieci

W trakcie uczenia sieć neuronowa na podstawie przykładów zawartych w zbiorze uczącym **samoczynnie odnajduje występujące w tych danych prawidłowości, zależności, tendencje i trendy**. W efekcie zmieniając swoje parametry sieć stopniowo dopasowuje się do danych ze zbioru uczącego [187].

W procesie uczenia sieć neuronowa powinna nabyć zdolności rozwiązywania **zadań podobnych** do tych, które były jej stawiane w trakcie uczenia, choć niekoniecznie identycznych. Oznacza to, że odpowiednio nauczony model tego rodzaju musi posiadać zdolność generalizacji (uogólniania) zdobytej wiedzy na nowe przypadki. [189]

#### **4.6.2 Metody uczenia sieci neuronowych**

Istnieje wiele **algorytmów uczenia** sieci neuronowych. Ich celem jest systematyczne zmniejszanie błędu popełnianego przez sieć w trakcie uczenia poprzez odpowiednio dokonywane automatyczne zmiany parametrów wszystkich neuronów występujących w sieci [186]. Istnieją dwie zasadnicze metody uczenia sieci neuronowych. Są to: **uczenie z nauczycielem** oraz **uczenie bez nauczyciela**, zwane również samouczeniem [189]. W niniejszej pracy stosowano wyłącznie tę pierwszą metodę, z związku z czym zostanie ona szerzej omówiona poniżej.

**Uczenie z nauczycielem** przebiega w następujących etapach [189]:

- przygotowanie zbioru danych uczących, zawierających dla każdego przypadku zarówno zmienne wejściowe, jak i odpowiadające im zmienne wyjściowe (jest to analogia do przykładów zadań podawanych wraz z rozwiązaniami) – ich źródłem mogą być np. dane o operowanych w przeszłości pacjentach wraz z informacją o uzyskanych u nich wynikach leczenia,
- wielokrotne uruchamianie sieci dla wszystkich przypadków ze zbioru danych uczących i sprawdzanie zgodności generowanych przez nią wartości wyjściowych z wartościami prawidłowymi (każdorazowe zaprezentowanie całego zbioru uczącego to tzw. „epoka”),
- przeprowadzanie w trakcie każdego wyżej wymienionego uruchomienia takich zmian współczynników wagowych (przy użyciu wybranego algorytmu uczenia), aby stopniowo minimalizować sumaryczny błąd popełniany przez sieć.

Jak opisano powyżej, proces uczenia polega na automatycznym **ustaleniu takich współczynników wagowych** we wszystkich neuronach sieci, by wartość sumarycznego błędu popełnianego przez sieć była najmniejsza [186]. W ten sposób parametry modelu dopasowywane są do dostępnych danych uczących. Przed rozpoczęciem uczenia wagom i wartościom progowym nadawane są wartości losowe.

Zadanie automatycznej modyfikacji wag i progów może być realizowane przy pomocy jednego z licznych **algorytmów uczenia** sieci. Najbardziej znaną metodą uczenia z nauczycielem jest algorytm wstecznej propagacji błędów (BP – *backpropagation*). Jest to metoda stosunkowo powolna, ale uniwersalna, tj. można przy jej pomocy rozwiązać praktycznie każde zadanie [186]. Polega ona na matematycznym wyznaczeniu po każdym pokazie przypadku uczącego wektora gradientu na powierzchni błędu, tj. kierunku najszybszego wzrostu funkcji błędu - i zmianie wag w sieci w kierunku przeciwnym. Jest to mechanizm skutecznie zmierzający do osiągnięcia minimum błędu [187]. Współcześnie dysponujemy również innymi algorytmami uczenia sieci neuronowych, takimi jak nowoczesne algorytmy drugiego rzędu (metoda gradientów sprzężonych, metoda Levenberga-Marquardta). Są one znacznie szybsze od algorytmu BP, jednak nie tak uniwersalne. Do

dyspozycji badacza są także wypracowane empirycznie modyfikacje algorytmu BP (metoda szybkiej propagacji, metoda Delta-Bar-Delta) [215]. **Zakończenie uczenia** następuje – zależnie od decyzji użytkownika – po określonej liczbie epok, po osiągnięciu zadanego poziomu błędu lub w przypadku braku dalszego zmniejszania wartości błędu lub po wykryciu (przy zastosowaniu tzw. zbioru walidacyjnego) znamion „przeuczenia”, o czym będzie mowa w dalszej części niniejszego rozdziału.

## 4.7 Kodowanie danych na potrzeby sieci neuronowych

Sieci neuronowe przetwarzają wyłącznie **dane numeryczne**. Z tego względu zarówno dane wejściowe, jak i wyjściowe, muszą zostać zakodowane w formie liczbowej.

W przypadku danych występujących *a priori* w formie **liczbowej** (np. wiek wyrażony w latach, liczba przebytych zabiegów) wartości odpowiednich zmiennych mogą być bezpośrednio wykorzystane przez sieć, niekiedy wymagając jedynie przeskalowania do odpowiedniego przedziału. W przypadku zmiennych nominalnych mających charakter **dwustanowy** (np. płeć) stosowane jest proste kodowanie binarne (np. 0 – mężczyzna, 1 – kobieta). Zmienne nominalne o charakterze **wielostanowym** na ogół koduje się metodą *jeden-z-N* polegającą na tym, że do reprezentacji pojedynczej zmiennej nominalnej używa się kilku zmiennych numerycznych. Ich liczba jest równa liczbie możliwych wartości zmiennej nominalnej, przy czym kodowanie polega na tym, że jedna z tych zmiennych numerycznych przyjmuje wartość 1, pozostałe zaś wartość 0. Na przykład lokalizację raka krtani (nadgłośnia, głośnia lub podgłośnia) można zakodować następująco:  $nadgłośnia = \{1,0,0\}$ ,  $głośnia = \{0,1,0\}$ ,  $podgłośnia = \{0,0,1\}$  [199]. Skutkiem takiego kodowania jest zwiększenie liczby neuronów (w przytoczonym przypadku – jedna zmienna jest kodowana przez 3 neurony), a co za tym idzie – zwiększenie rozmiarów całej sieci.

## 4.8 Typy problemów rozwiązywanych przez sieci neuronowe

Problemy rozwiązywane przez sieci neuronowe można zasadniczo podzielić na dwie kategorie: **klasyfikację i regresję** [215].

**Klasyfikacja** polega na przydzieleniu rozpatrywanego przypadku do jednej z zadanych klas. Przykładowe zadania dla modelu prognostycznego tego rodzaju to np. określenie, czy po operacji nastąpi poprawa w nasileniu występujących dolegliwości (dwie klasy: poprawa lub brak poprawy). Również wybór właściwej diagnozy z listy ustalonych wcześniej możliwości można traktować jako problem klasyfikacji.

**Regresja** polega na prognozowaniu z pomocą sieci konkretnej wartości wybranej zmiennej wyjściowej, na przykład nasilenia objawów wyrażonego w skali 0-10.

## 4.9 Zbiór uczący i zmienne wejściowe

Skuteczność działania modelu w postaci sieci neuronowej bardzo silnie zależy od zbioru danych wykorzystanych do jego uczenia, ponieważ sieć neuronowa w procesie uczenia nabywa zdolności modelowania **jedynie tych zależności, które były możliwe do zaobserwowania w zbiorze uczącym** [215]. Nie można zatem oczekiwać prawidłowych rozwiązań dla zagadnień zasadniczo różnych od tych wykorzystanych w trakcie uczenia. Jeżeli na przykład sieć neuronowa ma prognozować skuteczność leczenia operacyjnego, a w zbiorze uczącym znaleźli się jedynie pacjenci zoperowanie zgodnie z regułami sztuki, trudno wymagać, aby model prawidłowo przewidział niepowodzenie zabiegu u kolejnej osoby, w przypadku której operację przeprowadzono nieprawidłowo.

Zadaniem badacza korzystającego z sieci neuronowych jest określenie, jakie dane powinien zawierać zbiór uczący, aby umożliwić zbudowanie skutecznego modelu neuronowego. W tym celu należy określić przede wszystkim, na podstawie **jakich zmiennych** (danych wejściowych) jest osiągalne prawidłowe działanie modelu, czyli generowanie pożądanych danych wyjściowych. Na tym etapie niezbędna jest znajomość specyfiki modelowanego zagadnienia. W przypadku prognozowania wyników leczenia konieczna będzie na przykład wiedza dotycząca czynników potencjalnie mogących wpływać na przebieg choroby i ostateczne efekty terapii.

Określenie właściwego **rozmiaru zbioru uczącego** jest kwestią dyskusyjną. Istnieją w tym zakresie pewne reguły heurystyczne, jednak trudno uznać je za uniwersalne, ponieważ w rzeczywistości liczba przypadków niezbędnych do nauczenia konkretnej sieci zależy w dużej mierze od złożoności poddawanej modelowaniu zależności funkcyjnej [215].

Najprostsza ze wspomnianych reguł mówi, że liczba przypadków uczących powinna być 5-10 razy większa od liczby połączeń występujących w sieci [217]. Nieco bardziej skomplikowaną regułą empiryczną, która jednak w przejrzysty sposób ilustruje istotę problemu, zastosowali Andrea i wsp. [218]. Wprowadzili oni **parametr  $\rho$** , określający stosunek liczby przypadków uczących do całkowitej liczby parametrów sieci dostosowywanych w trakcie uczenia, wyrażony następującym równaniem:

$$\rho = \frac{U}{(I + 1)H + (H + 1)O}$$

gdzie  $U$  oznacza liczbę przypadków uczących, natomiast  $I$ ,  $H$  i  $O$  odpowiednio ilość neuronów w warstwie wejściowej, ukrytej i wyjściowej. Aby działanie sieci było prawidłowe, zdaniem autorów metody wartości  $\rho$  powinny mieścić się w przedziale od 1.8 do 2.2. Modele, dla których wartość  $\rho$  jest większa od 3 na ogół nie mają zdolności nauczenia się wszystkich istotnych cech zbioru uczącego, podczas gdy wartość  $\rho$  bliska 1 świadczy o nadmiernej złożoności modelu, co może dawać mu możliwość zbyt dokładnego dopasowania do zbioru uczącego i skutkuje utratą zdolności do generalizacji wiedzy.



Wspólną cechą przedstawionych powyżej reguł jest uzależnienie wielkości zbioru uczącego od liczby połączeń występujących w sieci. Należy w tym miejscu zauważyć, że im większa liczba zmiennych wejściowych zostanie uwzględniona, tym więcej neuronów pojawi się w sieci. W efekcie liczba połączeń między neuronami odpowiednio wzrośnie, a co za tym idzie – niezbędny zbiór uczący będzie musiał być liczniejszy. Jest to logiczna konkluzja, ponieważ narzucając wykorzystanie większej ilości zmiennych wejściowych, użytkownik dąży do modelowania bardziej wielowymiarowych zależności. Na ogół wymaga to od modelu większej elastyczności, czyli zwykle większej ilości elementów. Właściwe ustalenie znacznej liczby wag oraz odkrycie wpływu każdej zmiennej na ostateczny wynik wymaga nieodzownie zapoznania sieci ze znaczną liczbą „przykładów” ze zbioru uczącego. Ilustrując to zagadnienie można posłużyć się następującym przykładem: Aby zorientować się, że wielkość guza nowotworowego wpływa na wynik jego leczenia, wystarczy poznać kilku chorych z nowotworami o różnym zaawansowaniu. Jednak do odkrycia subtelnego wpływu rozmaitych chorób towarzyszących czy cech histopatologicznych guza na długość przeżycia, niezbędne jest długoletnie doświadczenie kliniczne i obserwacja losów wielu kolejnych pacjentów. Łatwo jednak także zauważyć, że gdy jest to możliwe, to zwiększanie liczby danych wejściowych jest celowe, ponieważ predykcje oparte na większej liczbie danych będą niewątpliwie dokładniejsze niż wnioskowanie w oparciu o pojedynczą zmienną, czy kilka danych wybranych z dużej liczby informacji dostępnych.

W związku z powyższym, a także ze względu na istotną kwestię zależności zdolności do generalizacji od wielkości sieci, dysponując ograniczonym zbiorem danych uczących użytkownik jest często zmuszony do rezygnacji z wykorzystania pewnej liczby zmiennych wejściowych.

Do zadań użytkownika należy więc często takie przeprowadzenie **redukcji wektora wejściowego** (tj. zmniejszenie liczby zmiennych wejściowych), aby uzyskać sieć o dobrej zdolności do generalizacji w miarę możliwości bez strat istotnych informacji.

Zadane to może być realizowane na przykład następującymi metodami [215]:

- odrzucenie zmiennych uznanych przez użytkownika za mniej przydatne (wywierające mniejszy wpływ na badane zjawisko) na podstawie znajomości modelowanego zagadnienia lub analizy statystycznej wzajemnych zależności między zmiennymi wejściowymi a zmiennymi wyjściowymi,
- identyfikacja i usunięcie zmiennych skorelowanych,
- ocena wpływu poszczególnych zmiennych na działanie modelu (tzw. analiza wrażliwości) i usunięcie potencjalnie zbędnych wejść,
- określenie wektora wejściowego metodą eksperymentalną poprzez przetestowanie licznych modeli posługujących się różnymi wektorami wejściowymi,
- metody krokowe (postępująca i wsteczna), tj. stopniowe zwiększanie bądź zmniejszanie rozmiarów wektora wejściowego [219],
- zastosowanie algorytmów genetycznych, tj. metody wywodzącej się z teorii ewolucji i doboru naturalnego („krzyżowanie” modeli i selekcja najlepszych „osobników”

spośród „potomstwa” celem dalszego krzyżowania - z dopuszczeniem losowych mutacji - w celu uzyskania kolejnych „pokoleń” coraz skuteczniejszych sieci),

- redukcja wymiaru przestrzeni wejść metodą odpowiedniej transformacji, np. analiza składowych głównych (PCA – *principal component analysis*), polegająca na „rzutowaniu” danych pierwotnych do przestrzeni o mniejszym rozmiarze, co umożliwia zastąpienie dużej liczby zmiennych wejściowych mniejszą ich liczbą, zawierającą jednak (prawie) tę samą informację, co oryginalny wektor wejściowy,
- nieliniowy odpowiednik PCA, realizowany za pomocą autoasocjacyjnej sieci neuronowej (sieć dokonująca „kompresji” sygnałów wejściowych w warstwie ukrytej tak, aby w warstwie wyjściowej dało się je ponownie dokładnie odtworzyć, co gwarantuje, że uzyskane w warstwie ukrytej „skompresowane” dane są równo cenne jak dane oryginalne).

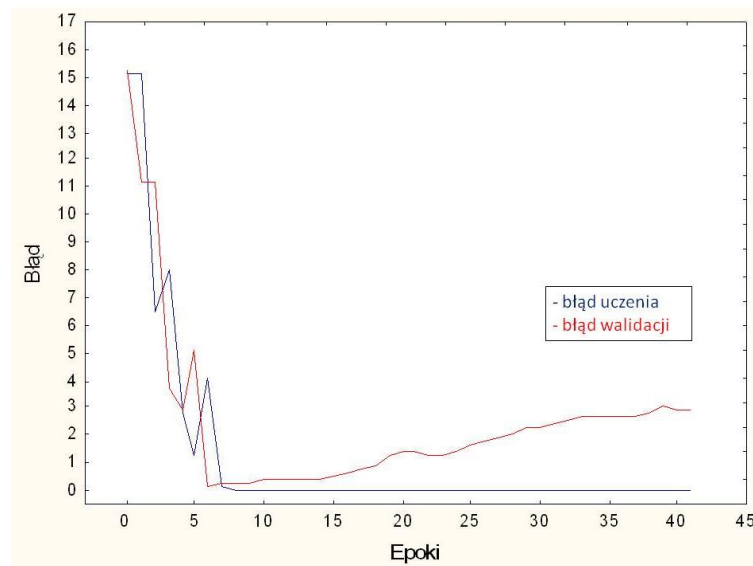
#### 4.10 Zdolność sieci do generalizacji, walidacja

Jak wyjaśniono powyżej, w trakcie uczenia sieci neuronowej celem działania algorytmu uczącego jest stopniowa minimalizacja błędu popełnianego przez sieć przy rozwiązywaniu zadań należących do zbioru uczącego. Docelowym zadaniem modelu nie jest jednak podawanie prawidłowych wyników dla przypadków ze zbioru uczącego (dla których właściwe odpowiedzi z założenia są już znane), ale wyznaczanie tych wyników dla nowych przypadków. Sieć neuronowa powinna zatem posiadać zdolność **generalizacji** wiedzy zdobytej w procesie uczenia. Uzyskanie prawidłowych wyników nie tylko dla przypadków uczących, ale również dla wszelkich innych przypadków jest możliwe pod warunkiem, że w procesie uczenia sieci uda się prawidłowo wymodelować rzeczywistą zależność istniejącą między zmiennymi wejściowymi i wyjściowymi. Tymczasem w uczącej się sieci neuronowej modelowane są zależności występujące w zbiorze uczącym (niekiedy nawet przypadkowe), a niekoniecznie charakterystyczne dla całej populacji, tj. dla wszystkich zadań, które w przyszłości model powinien rozwiązywać.

Jednym z problemów napotykanym w procesie uczenia jest wspomniane uprzednio tzw. nadmierne dopasowanie (**przeuczenie**). Dochodzi do niego, jeśli model pod koniec procesu uczenia nie dopasowuje się już do zasadniczego kształtu aproksymowanej zależności, ale do drobnych szczegółów reprezentowanych przez pojedyncze dane ze zbioru uczącego [220]. Zauważono, że sieci neuronowe posiadające dużą liczbę wag mają na tyle dużą elastyczność, że wykazują tendencję do takich właśnie nadmiernie szczegółowych dopasowań. Mniejsza złożoność sieci wymusza poszukiwanie funkcji modelującej dane uczące w sposób bardziej ogólny, a co za tym idzie – na ogół bliższy ilustrowanej przez nie realnie istniejącej zależności. Jednak nadmierna redukcja wielkości sieci może prowadzić do zupełnej utraty zdolności modelowania zjawiska o określonej złożoności. Stąd dla każdego zadania praktycznego konieczne jest empiryczne poszukiwanie właściwych rozmiarów sieci.

Aby uniknąć nadmiernego dopasowania do zbioru uczącego w trakcie uczenia sieci, wykorzystuje się proces **walidacji** [215]. W tym celu niezbędny jest dodatkowy zbiór danych (zadań wraz ze znanymi poprawnymi rozwiązaniami), dla których w trakcie uczenia sprawdza się wartości błędu niezależnie od obserwacji jego zmniejszania się dla zbioru uczącego. Równoczesne zmniejszanie się błędu uczenia i błędu walidacji świadczy o prawidłowym przebiegu procesu uczenia. Z kolei zmniejszanie błędu uczenia przy niezmiennym lub rosnącym błędzie walidacji sugeruje przeuczenie, ponieważ oznacza to utratę zdolności do generalizacji i osiągnięcie w efekcie coraz gorszych wyników dla nowych przypadków przy poprawie dopasowania jedynie do zbioru uczącego (rys. 14). Obserwując błąd uczenia i błąd walidacji w trakcie uczenia użytkownik może wnioskować, czy wybrana przez niego architektura sieci jest prawidłowa, czy też wymaga poszerzenia lub redukcji.

**Rysunek 14. Wykres przedstawiający spadek błędu uczenia oraz początkowy spadek, a następnie wzrost błędu walidacji w trakcie uczenia sieci neuronowej, co świadczy o przeuczeniu modelu.**



## 4.11 Testowanie

W trakcie uczenia sieć nie dopasowuje się do zbioru walidacyjnego w taki sposób, jak do zbioru uczącego, jednak pośrednio wpływa on na ostateczny wybór modelu. Z tego względu jakość odpowiedzi udzielanych przez sieć dla zbioru walidacyjnego również nie pozwala na całkowicie pewne wnioskowanie o jej efektywności w odniesieniu do nowych danych, nie wykorzystanych w procesie uczenia. Można bowiem założyć, że uzyskany model przypadkowo będzie działał prawidłowo w odniesieniu do zbioru uczącego i zbioru walidacyjnego, a jednak nadal jego zdolność do generalizacji będzie niezadowalająca.

Z tego względu wskazane jest stworzenie trzeciego zbioru danych (zadań ze znanymi rozwiązaniami), służącego wyłącznie do oceny jakości modelu już po zakończeniu jego

uczenia. Jest to tzw. **zbiór testowy**. Uważa się, że raportując wyniki uzyskiwane przez model należy odnosić się przede wszystkim właśnie do tego zbioru, ponieważ jako jedyny pozwala on w sposób obiektywny wnioskować, jak model będzie sobie radził z nowymi, zupełnie nieznanymi przypadkami [219, 221].

#### 4.12 Problem wyboru zbioru uczącego, walidacyjnego i testowego

Jak wyjaśniono powyżej, aby prawidłowo nauczyć sieć neuronową i prawidłowo przetestować jej działanie, potrzebne są trzy zbiory: uczący, walidacyjny i testowy. Każdy z tych zbiorów powinien zawierać **dane reprezentatywne dla opisywanej zależności** [215]. Wybranie wadliwego zbioru uczącego, który nie odzwierciedla w pełni wszystkich aspektów modelowanego zjawiska, musi prowadzić do niewłaściwego odwzorowania badanej zależności przez sieć. Podobnie niewłaściwie dobrany zbiór walidacyjny zaburza proces uczenia. Jeśli zawiera „zbyt proste” zadania, użytkownik może dokonać wyboru modelu, który w rzeczywistości nie poradzi sobie z bardziej skomplikowanymi problemami. Z kolei umieszczenie w zbiorze walidacyjnym przypadków istotnie różniących się od tych w zbiorze uczącym (przypadki nietypowe) spowoduje, że osiągnięcie dobrej jakości walidacji porównywalnej z jakością uczenia może się okazać niemożliwe. Z analogicznego powodu istotny jest wybór **reprezentatywnego** zbioru testowego.

Optymalnym rozwiązaniem byłoby wykorzystanie zarówno do uczenia, jak i do walidacji i testowania bardzo dużych zbiorów danych, z których każdy zawierałby wszystkie możliwe „typy przypadków” mogące wystąpić w populacji. Na ogół jednak te zbiory danych, jakimi dysponuje badacz, nie są dostatecznie liczne, a ponadto nieuchronnie pojawiają się w nich pojedyncze przypadki nietypowe, które zakłócają przebieg uczenia, walidacji i testowania modelu.

W odniesieniu do baz danych o ograniczonej liczebności istotnym problemem staje się więc **właściwe rozdzielenie przypadków** pomiędzy zbiór uczący, walidacyjny i testowy. Najpopularniejszym podejściem jest **podział losowy**. Doświadczenie wskazuje jednak, że w przypadku niedostatecznie licznych zbiorów, zawierających przypadki o dużej różnorodności, takie podejście może prowadzić do przypadkowego uzyskania niereprezentatywnych zbiorów [203]. Zwłaszcza dostępna w niektórych programach metoda rozlosowywania zbiorów na nowo dla każdej kolejnej tworzonej sieci może skutkować wyborem modelu, dla którego przez przypadek wylosowano szczególnie korzystny zbiór walidacyjny i testowy. Przekonanie o jego dobrej jakości podważa dopiero zakończona fiaskiem próba zastosowania go w praktyce dla zupełnie nowych przypadków.

Jedną z alternatywnych metod jest wstępny podział zbioru danych na **skupienia** (klastry), tj. niewielkie podzbiory podobnych do siebie przypadków. Następnie z każdego skupienia przypadki przydzielane są odpowiednio do zbioru uczącego, testowego i walidacyjnego. W ten sposób można założyć, że w każdym ze zbiorów znajdą się przypadki odzwierciedlające wszystkie „typy problemów” występujące w całej bazie danych.

### 4.13 Poszukiwanie optymalnego modelu

Ze względu na specyfikę metod uczenia sieci neuronowych, w zasadzie nigdy nie można mieć pewności, czy uzyskany model jest optymalny, tj. czy poziom błędu osiągnął najniższą możliwą wartość. Nie jest możliwe jednoznaczne określenie, czy w procesie uczenia zostało osiągnięte tak zwane minimum globalne, czy też jedynie minimum lokalne funkcji błędu, ponieważ z założenia błąd zmniejszany jest zgodnie z „lokalnym gradientem”, a punkt początkowy „eksploracji powierzchni błędu” (która może mieć bardzo złożony charakter) ustalany jest losowo [215].

Jednym z rozwiązań tego problemu jest wielokrotne rozpoczynanie procesu uczenia z wykorzystaniem różnych (losowych) początkowych wartości wag. Wymaga to przeprowadzenia wielu eksperymentów w celu uzyskania sieci o możliwie najmniejszej wartości błędu.

### 4.14 Interpretacja działania modelu - ograniczenia

Ponieważ proces przetwarzania danych w obrębie sieci neuronowej jest bardzo skomplikowany, użytkownik na ogół nie ma możliwości bezpośredniej analizy przekształceń, jakim poddawane są zmienne w kolejnych warstwach sieci. Z tego powodu modele neuronowe są znacznie mniej „przejrzyste” niż tradycyjne modele statystyczne [221]. Jest to istotna i niekorzystna różnica, która powoduje, że sieci neuronowe nie poddają się prostej interpretacji. Sceptycy twierdzą, że sieć neuronowa stanowi rodzaj „czarnej skrzynki”, do której podaje się dane, otrzymując na wyjściu określone wyniki, jednak procesy zachodzące „we wnętrzu” pozostają niewyjaśnione. Cecha ta powoduje, że model neuronowy zazwyczaj **nie daje możliwości wnioskowania na temat mechanizmu** modelowanego zjawiska. Korzystając z sieci neuronowych należy zatem zdawać sobie sprawę, że ich celem **nie jest** wyjaśnianie zależności przyczynowo-skutkowych. Ich działanie jest czysto pragmatyczne: mają dobrze przewidywać wartości wyjściowe dla podawanych na wejściu danych opisujących rozważane przypadki.

Również wnioskowanie o istotności lub nieistotności pewnych zmiennych dla przebiegu zjawiska wyłącznie na podstawie faktu wykorzystania ich przez skuteczny model neuronowy (lub automatycznego odrzucenia w trakcie optymalizacji działania sieci) powinno być przeprowadzane z daleko idącą ostrożnością. Niejednokrotnie zdarza się, że równie dobre wyniki uzyskuje się przy pomocy dwóch różnych sieci, bazujących na całkowicie odmiennych zestawach danych wejściowych i odnajdujących prawidłowe rozwiązania zadanego problemu dwiema całkowicie niezależnymi metodami.

Z tego względu sieci neuronowe można traktować wyłącznie jako skuteczne narzędzie np. do prognozowania przebiegu pewnych zjawisk, jednak zdecydowanie należy zachować daleko idącą ostrożność próbując posługiwać się nimi w celu wyjaśniania tych zjawisk lub ich interpretacji.

#### **4.15 Tworzenie modeli neuronowych - zadania badacza (podsumowanie)**

Jak omówiono powyżej, mimo dostępności oprogramowania znacznie ułatwiającego korzystanie z sieci neuronowych, automatycznie przeprowadzającego dla użytkownika cały szereg operacji matematycznych, znaczna część zadań związanych z tworzeniem modeli neuronowych spoczywa nadal na badaczu i wymaga od niego zarówno znajomości modelowanego zagadnienia, jak i specyfiki działania sieci.

Podsumowując, **zadania badacza** można streścić w następujących punktach:

- wybór zagadnienia, które może być modelowane przy użyciu sieci neuronowych,
- skompletowanie bazy danych zawierającej wystarczającą liczbę przykładów, aby dostatecznie zobrazować modelowaną zależność,
- określenie, jakie zmienne wejściowe mogą być użyteczne w celu wyznaczenia pożądanых zmiennych wyjściowych oraz ewentualnie przeprowadzenie redukcji wektora wejściowego,
- właściwe zakodowanie zmiennych wejściowych i wyjściowych,
- wyodrębnienie reprezentatywnego zbioru uczącego, walidacyjnego i testowego,
- wybór odpowiedniej architektury sieci, określenie jej cech (typ sieci, liczba warstw i neuronów w poszczególnych warstwach, funkcje aktywacji neuronów) – najczęściej na drodze licznych eksperymentów,
- właściwe przeprowadzenie procesu uczenia,
- krytyczna interpretacja uzyskanych wyników i prawidłowy wybór najskuteczniejszego modelu.

## 5. Materiał i metoda

### 5.1 Grupa badana i jej charakterystyka kliniczna

Grupę badaną stanowiło 183 dorosłych chorych operowanych metodą endoskopową z powodu PZZP w Oddziale Klinicznym Kliniki Otolaryngologii Szpitala Uniwersyteckiego w Krakowie.

**Kryteria włączenia** do badania obejmowały:

- rozpoznanie PZZP zgodnie z definicją EPOS2012 [10],
- kwalifikację do operacji endoskopowej zatok przynosowych wobec nieskuteczności leczenia zachowawczego oraz
- świadomą zgodę pacjenta na udział w badaniu.

Przyjęte **kryteria wyłączenia** również były zgodne z EPOS2012 i obejmowały:

- mukowiscydozę,
- niedobory odporności,
- wrodzone zaburzenia czynności rzęsek,
- nieinwazyjne grzybniaki zatok lub inwazyjną grzybicę zatok,
- zapalenia naczyń i choroby przebiegające z wytwarzaniem ziarniny (w tym zwłaszcza ziarniniak Wegenera i zespół Churga-Strauss) oraz
- nowotwory zatok.

Ponadto z badania wyłączono chorych, którzy nie zgłosili się na badanie kontrolne oraz tych, u których przeprowadzono kolejną operację zatok w okresie pomiędzy analizowanym zabiegiem operacyjnym, a badaniem kontrolnym.

**Przedoperacyjna ocena pacjentów** obejmowała:

- wywiad:
  - wiek i płeć,
  - choroby towarzyszące (0 – obecne, 1 – nieobecne, 0.5 – niepewne, np. nie w pełni potwierdzone podejrzenie choroby):
    - astma oskrzelowa,
    - nadwrażliwość na niesteroidowe leki przeciwzapalne (AIA),
    - alergia,
    - refluks żołądkowo-przełykowy,
  - przebyte zabiegi operacyjne dotyczące nosa i zatok przynosowych:
    - polipektomie (liczba - niezależnie po stronie prawej i lewej),

- operacje endoskopowe zatok przynosowych (liczba – niezależnie po stronie prawej i lewej),
- badanie laryngologiczne (uzupełnione o śródoperacyjną endoskopową ocenę jamy nosowej i zatok przynosowych przeprowadzoną po obkurczeniu śluzówki):
  - ocena obecności polipów nosa
    - wielkość polipów po każdej ze stron była oceniana jak w klasyfikacji Lund i Kennedy’ego [17]: 0 – brak, 1 – występują w przewodzie nosowym środkowym, 2 – wykraczają poza przewód nosowy środkowy,
    - zgodnie z EPOS2012 chory był zaliczany do grupy „z polipami nosa” jeśli polipy występowały obustronnie w przewodach nosowych środkowych [10],
    - polipy były określane jako „masywne” jeśli obustronnie wykraczały poza przewód środkowy,
  - ocena obecności skrzywienia przegrody nosa:
    - 0 – przegroda określana jako prosta – maksymalnie stopień I w skali Levine i May, tj. widoczna boczna i przyśrodkowa powierzchnia małżowiny nosowej środkowej,
    - 1 – przegroda skrzywiona – stopień II-IV w skali Levine i May, tj. przedni przyczep małżowiny nosowej środkowej częściowo lub całkowicie zasłonięty przez przegrodę nosa bądź przegroda przemieszczona w stronę bocznej ściany jamy nosa [19],
- kwestionariusz samooceny objawów:
  - czas trwania objawów (1 – od 3 miesięcy do 1 roku, 2 – od 1 do 3 lat, 3 – od 3 do 5 lat, 4 – powyżej 5 lat),
  - ogólne nasilenie dolegliwości związanych z występowaniem PZZP w zmodyfikowanej skali VAS (szczegółowy opis poniżej),
  - nasilenie poszczególnych dolegliwości w skali od 0 do 3 (szczegółowy opis poniżej):
    - upośledzenie drożności nosa (trudności z oddychaniem przez nos),
    - wydzielina w nosie,
    - spływanie wydzieliny po tylnej ścianie gardła,
    - kaszel,
    - kichanie,
    - ból lub ucisk w obrębie twarzy,
    - ból lub ucisk w obrębie czoła,
    - ból głowy (w innym miejscu niż twarz lub czoło),
    - ból lub uczucie pełności w uchu,
    - upośledzenie węchu (brak odczuwania zapachów),
    - długotrwałe zmęczenie, rozdrażnienie, brak koncentracji,
    - nieprzyjemny zapach z ust,
    - ból górnych zębów,



- tomografię komputerową zatok przynosowych, ocenianą zgodnie ze skalą Lund-Mackay (0-24) [14],
- przebieg zabiegu operacyjnego:
  - zakres zabiegu (zabieg jednostronny czy obustronny, liczba otwartych zatok) i jego radykalność (tj. czy zostały otwarte wszystkie zatoki zajęte przez zmiany zapalne),
- po zabiegu - wynik pooperacyjnego badania histopatologicznego materiału z zatok przynosowych:
  - rodzaj komórek dominujących w nacieku zapalnym (0 - nacieki nieeozynofilowe, tj. najczęściej złożone z limfocytów i plazmacytów, 1- nacieki głównie eozynofilowe, 0.5 – rodzaj komórek dominujących w nacieku zapalnym niepewny).

Wszyscy pacjenci zostali poddani operacji endoskopowej zatok przynosowych w znieczuleniu ogólnym. Zakres zabiegu był zależny od lokalizacji i rozległości zmian w zatokach przynosowych. W niektórych przypadkach został on ograniczony ze względu na warunki anatomiczne lub nadmierne krwawienie. Leczenie pooperacyjne było dostosowane do indywidualnych potrzeb pacjenta zależnie od nasilenia objawów, chorób towarzyszących i wyników kontrolnych badań endoskopowych. U wszystkich pacjentów w okresie pooperacyjnym stosowano oczyszczanie zatok ze strupów i zalegającej wydzieliny pod kontrolą endoskopu, płukanie nosa oraz zależnie od okoliczności sterydy donosowo lub doustnie, antybiotyki, mukolityki, leki obkurczające błonę śluzową, leki antyhistaminowe i inhibitory pompy protonowej.

**Pooperacyjna ocena nasilenia objawów** przeprowadzana była w trakcie rutynowych kontroli: (I) 3-6 miesięcy po zabiegu operacyjnym, (II) około 12 miesięcy (10-15 miesięcy) po zabiegu operacyjnym. W trakcie kontroli pacjenci wypełniali kwestionariusze samooceny objawów analogiczne jak przed operacją.

Ponieważ nie wszyscy pacjenci zgłosili się na oba badania kontrolne, analizy przeprowadzono niezależnie dla **dwóch grup chorych**:

- **grupa I** - pacjenci badani przed zabiegiem i 3-6 miesięcy po zabiegu (140 pacjentów),
- **grupa II** - pacjenci badani przed zabiegiem i około 12 miesięcy po zabiegu (72 pacjentów).

29 pacjentów zostało zbadanych zarówno po 3-6 miesiącach, jak i po roku od zabiegu i w związku z tym zostali oni niezależnie włączeni zarówno do grupy I, jak i do grupy II. Szczegółową charakterystykę poszczególnych grup przedstawiono w tabelach 2 i 3.

Przebieg badania został zatwierdzony przez Komisję Bioetyczną Uniwersytetu Jagiellońskiego (opinia nr KBET/219/B/2012).

Tabela 2. Charakterystyka grupy I (140 pacjentów zbadanych przed zabiegiem i 3-6 miesięcy po zabiegu).

Cecha	Wartości				
wiek	zakres: 19-75 lat; mediana 45 lat				
dolegliwości przed zabiegiem (VAS)	zakres: 1-10; mediana 8; brak danych: 2 (1%)				
ocena TK (skala Lund-Mackay)	zakres: 1-24; mediana 13; brak danych: 11 (8%)				
	liczba pacjentów (%)				
pleć	kobiety: 72 (51%)			mężczyźni: 68 (49%)	
czas trwania objawów	3 -12 mies.: 11 (8%)	1-3 lata: 42 (30%)	3-5 lat: 23 (16%)	> 5 lat: 61 (44%)	brak danych: 3 (2%)
	tak / obecne	nie / nieobecne	niewpewne	brak danych	
polipy nosa	43 (31%)	95 (68%)	-	2 (1%)	
masywne polipy	19 (14%)	119 (85%)	-	2 (1%)	
skrzywienie przegrody nosa	78 (56%)	59 (42%)	-	3 (2%)	
astma	46 (33%)	88 (63%)	6 (4%)	-	
AIA	20 (14%)	116 (83%)	4 (3%)	-	
alergia	58 (42%)	79 (56%)	3 (2%)	-	
refluks	28 (20%)	109 (78%)	3 (2%)	-	
przebyte polipektomie	28 (20%) liczba: 1-20	112 (90%)	-	-	
przebyte ESS	24 (17%) liczba: 1-12	116 (83%)	-	-	
radykałność aktualnego zabiegu	114 (82%)	20 (14%)	-	6 (4%)	
eozynofile dominują w pooperacyjnym badaniu histopatologicznym	82 (59%)	44 (31%)	12 (8%)	2 (1%)	

**Tabela 3. Charakterystyka grupy II (72 pacjentów zbadanych przed zabiegiem i około 12 miesięcy po zabiegu).**

<b>Cecha</b>	<b>Wartości</b>				
wiek	zakres: 22-73 lat; mediana 48 lat				
dolegliwości przed zabiegiem (VAS)	zakres: 2-10; mediana 8				
ocena TK (skala Lund-Mackay)	zakres: 1-24; mediana 12; brak danych: 4 (6%)				
	liczba pacjentów (%)				
pleć	kobiety: 40 (56%)			mężczyźni: 32 (44%)	
czas trwania objawów	3 -12 mies.: 8 (11%)	1-3 lata: 22 (31%)	3-5 lat: 10 (14%)	> 5 lat: 32 (44%)	brak danych: -
	<b>tak / obecne</b>	<b>nie / nieobecne</b>	<b>niepewne</b>	<b>brak danych</b>	
polipy nosa	18 (25%)	53 (74%)	-	1 (1%)	
masywne polipy	8 (11%)	63 (88%)	-	1 (1%)	
skrzywienie przegrody nosa	33 (46%)	38 (53%)		1 (1%)	
astma	18 (25%)	53 (74%)	1 (1%)	-	
AIA	6 (8%)	65 (91%)	1 (1%)	-	
alergia	24 (33%)	45 (63%)	3 (4%)	-	
reflaks	19 (26%)	51 (71%)	2 (3%)	-	
przebyte polipektomie	13 (18%) liczba: 1-10	59 (82%)	-	-	
przebyte ESS	11 (15%) liczba: 1-10	61 (85%)	-	-	
radykalność aktualnego zabiegu	59 (82%)	12 (17%)	-	1 (1%)	
eozynofile dominują w pooperacyjnym badaniu histopatologicznym	39 (54%)	26 (36%)	4 (6%)	3 (4%)	

## 5.2 Kwestionariusz samooceny objawów

Jako miarę sukcesu leczenia w niniejszej pracy przyjęto ustępowanie objawów subiektywnych PZZP po zabiegu operacyjnym. Założenie to wynika z definicji kontroli choroby, która zakłada, że najważniejszym celem terapii jest uwolnienie chorego od dolegliwości [10]. Jak podkreślano w rozdziale „Przewlekłe zapalenie zatok przynosowych – wprowadzenie i przegląd piśmiennictwa”, ustępowanie objawów nie zawsze idzie w parze z normalizacją obrazu w badaniu endoskopowym i tomografii komputerowej. Ponadto tomografia komputerowa jako badanie związane z ekspozycją na promieniowanie jonizujące nie może być rutynowo powtarzana u każdego pacjenta w celu oceny wyników leczenia.

Również w odniesieniu do poszczególnych objawów PZZP rozważanych niezależnie dostępne są niejednokrotnie obiektywne metody pomiarowe (np. rynomanometria służąca do pomiaru drożności nosa, testy olfaktometryczne do oceny węchu), jednak ich wyniki często nie korelują w prosty sposób z subiektywnym nasileniem dolegliwości. Uważa się, że nasilenie objawów odczuwanych przez pacjenta jest wynikiem interakcji między różnymi czynnikami, gdzie uwarunkowania biologiczne i fizjologiczne stanowią tylko element całości. Z tego powodu obecnie zakłada się, że pomiary oparte o subiektywną ocenę pacjenta są bardziej użyteczne w ewaluacji wyników leczenia. [10]

Zastosowany w niniejszej pracy kwestionariusz samooceny objawów opracowano opierając się na definicjach PZZP wg EPOS2012 [10] oraz Rhinosinusitis Task Force [12], a także na stosowanych powszechnie kwestionariuszach, głównie SNOT-20 [22]. Dokonane modyfikacje miały na celu uzyskanie możliwie najbardziej adekwatnych odpowiedzi pacjentów samodzielnie wypełniających kwestionariusz oraz umożliwienie oceny każdego z objawów niezależnie (a nie sumarycznie lub w kontekście jakości życia – jak np. w SNOT-20).

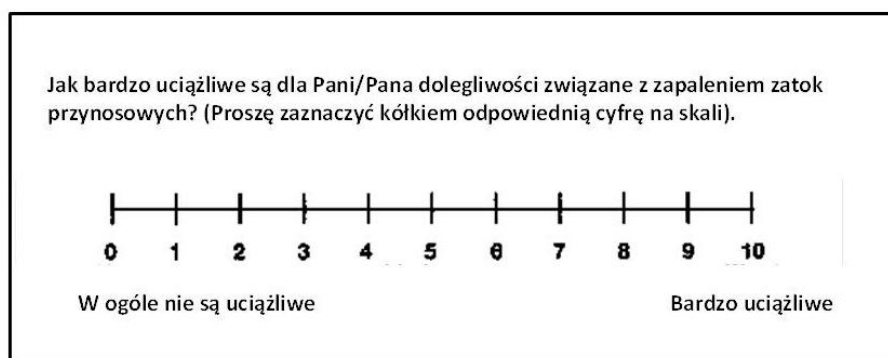
W celu oceny ogólnego nasilenia dolegliwości posłużono się **zmodyfikowaną skalą wizualno-analogową (VAS)** – rys. 15. Modyfikacja względem skali rekomendowanej przez konsensus EPOS2012 polegała na zaznaczeniu poszczególnych wartości (od 0 do 10) na skali. W efekcie narzędzie utraciło wprawdzie właściwości typowej skali analogowej, gdzie możliwe byłoby zaznaczanie również wartości niecałkowitych, jednak we wstępnym etapie badań przeprowadzonych na potrzeby niniejszej pracy wykazano, że taka postać skali była bardziej zrozumiała dla pacjentów i nie wymagała dodatkowych objaśnień w celu uzyskania miarodajnych odpowiedzi. W dalszej części pracy jako „skala VAS” jest skrótowo określana właśnie opisana powyżej skala zmodyfikowana. W analizach statystycznych jest ona traktowana jako skala interwałowa (przedziałowa).

Ewaluacji poszczególnych objawów dokonano w **skali 4-stopniowej** (0 - „brak objawu”, 1 – „objaw mało uciążliwy”, 2 – „średnio uciążliwy”, 3 – „bardzo uciążliwy”). Skala ta jest traktowana jako skala porządkowa (rangowa).

Obie skale – tj. zarówno skala VAS, jak i skala 4-stopniowa, są rekomendowane przez EPOS2012 do oceny nasilenia objawów PZZP.

Opracowując kwestionariusz ograniczono do niezbędnego minimum liczbę pytań w celu zwiększenia czytelności i skrócenia czasu wypełniania kwestionariusza. W przypadku bardziej rozbudowanego kwestionariusza podczas wstępnej ewaluacji narzędzia stwierdzono częstsze pomijanie pewnej liczby pytań przez pacjentów. Dodatkowo zastosowano cieniowanie pól w tabeli wypełnianej przez pacjenta – od barwy białej w polu oznaczającym brak objawu do najciemniejszej w polu oznaczającym znaczne nasilenie objawu (rys. 16).

**Rysunek 15. Zmodyfikowana skala wizualno-analogowa (VAS) zastosowana w kwestionariuszu samooceny objawów.**



**Rysunek 16. Tabela służąca do samooceny nasilenia poszczególnych objawów PZZP w skali 4-punktowej .**

Proszę zaznaczyć krzyżykami, jak bardzo uciążliwe są dla Pani/Pana poszczególne dolegliwości związane z przewlekłym zapaleniem zatok przynosowych:

	wcale	mało	średnio	bardzo
Upośledzona drożność nosa (trudności z oddychaniem przez nos)				
Wydzielina w nosie				
Spływanie wydzieliny po tylnej ścianie gardła				
Kaszel				
Kichanie				
Ból lub ucisk w obrębie twarzy				
Ból lub ucisk w obrębie czoła				
Ból głowy (w innym miejscu niż twarz lub czoło)				
Ból lub uczucie pełności w uchu				
Upośledzenie węchu (brak odczuwania zapachów)				
Długotrwałe zmęczenie, rozdrażnienie, brak koncentracji				
Nieprzyjemny zapach z ust				
Ból górnych zębów				

### 5.3 Zakres i wiarygodność analizowanych danych

Celem niniejszej pracy było opracowanie narzędzi prognostycznych **mogących potencjalnie znaleźć zastosowanie w praktyce klinicznej**. Analizy użyteczności tworzonych dotychczas modeli predykcyjnych wykazały, że na możliwość ich praktycznego wykorzystania wpływa między innymi dostępność niezbędnych danych wejściowych [222]. Modele wymagające przeprowadzania dodatkowych, potencjalnie kosztownych i czasochłonnych badań diagnostycznych, na ogół nie znajdują szerokiego zastosowania. Z tego powodu w niniejszej pracy **dopuszczono do wykorzystania w modelowaniu jedynie takich danych, jakie są rutynowo dostępne** w przypadku każdego pacjenta kwalifikowanego do operacji endoskopowej zatok przynosowych.

Rutynowe postępowanie przedoperacyjne w każdym ośrodku przeprowadzającym tego rodzaju zabiegi obejmuje zawsze wywiad i badanie laryngologiczne, badanie endoskopowe i tomografię komputerową zatok. W wielu przypadkach jednak nie są dostępne wyniki szczegółowych badań dotyczących chorób towarzyszących – w różnych ośrodkach (oraz zależnie od indywidualnych potrzeb i możliwości) ich diagnostyka bywa przeprowadzana już po leczeniu operacyjnym bądź nawet pomijana. Ze względu na te uwarunkowania zdecydowano oceniać obecność chorób towarzyszących wyłącznie na podstawie danych z wywiadu, nie wymagając dodatkowych testów potwierdzających lub wykluczających występowanie danego schorzenia (np. spirometrii w diagnostyce astmy lub pH-metrii w diagnostyce refluksu żołądkowo-przełykowego). Uzyskane w ten sposób dane mogą oczywiście być obciążone błędem, jednak celem pracy była ocena, na ile wiarygodnych predykcji można oczekiwać dysponując takimi właśnie informacjami, na jakich w obecnych – odległych od ideału - warunkach na ogół muszą bazować lekarze w swojej codziennej praktyce.

### 5.4 Analiza statystyczna

Analizę statystyczną przeprowadzono przy pomocy programu Statistica 7.0. W celu sprawdzenia normalności rozkładu posłużono się testem Shapiro-Wilka. Żadna ze zmiennych ilorazowych (wiek, skala Lund-Mackay) ani przedziałowych (skala VAS) nie miała rozkładu normalnego. Wykluczyło to możliwość zastosowania typowych testów parametrycznych. W przypadku zmiennych dotyczących obecności chorób towarzyszących lub nacieków eozynofilowych w badaniu histopatologicznym, przyjmujących wartości 0 (brak), 1 (obecne) lub 0.5 (niepewne), w analizach nie brano pod uwagę przypadków, dla których wynik był niepewny. Analizując zależności między poszczególnymi zmiennymi a nasileniem objawów po operacji posłużono się następującymi testami:

- do oceny korelacji dwóch zmiennych ilorazowych lub przedziałowych – współczynnik korelacji rang Spearmana,

- do oceny zależności między zmiennymi ilorazowymi lub przedziałowymi a zmiennymi binarnymi – test Manna-Whitney’a,
- do oceny zależności między dwiema zmiennymi binarnymi - test  $\chi^2$  (ewentualnie test  $\chi^2$  z poprawką Yatesa).

Przyjęto poziom istotności statystycznej  $\alpha=0.05$ .

## 5.5 Sztuczne sieci neuronowe

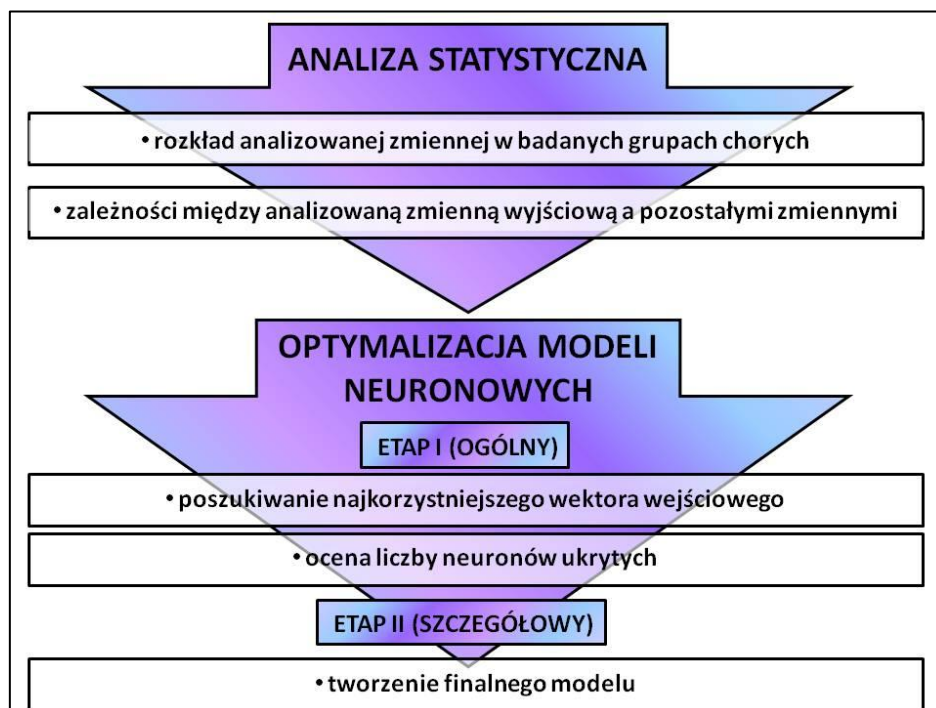
W niniejszej pracy omówiono stosunkowo obszernie metody stosowane do budowy sieci neuronowych, ponieważ – wobec braku jednolitego „obowiązującego” schematu postępowania w tym zakresie - wiele z przyjętych rozwiązań ma charakter autorski i stanowi wkład własny w badania nad modelami neuronowymi, a przyjęte założenia mają istotne znaczenie dla przebiegu eksperymentów opisanych w rozdziale „Wyniki”.

Sztuczne sieci neuronowe na potrzeby niniejszej pracy tworzono w programie Statistica 7.0. Budowane modele miały charakter **sieci klasyfikacyjnych**, tj. ich zadaniem było odpowiednie przyporządkowanie przypadków do jednej z dwóch zadanych klas. Tworzenie modeli przebiegało metodą „**uczenia z nauczycielem**”. W pracy wykorzystano wyłącznie **trójwarstwowe sieci typu MLP**, czyli perceptrony wielowarstwowe posiadające jedną warstwę ukrytą. **Przyjęto zapisywać architekturę sieci podając kolejno liczbę neuronów w warstwach: wejściowej, ukrytej i wyjściowej** (np. zapis MLP 6-5-1 oznacza perceptron trójwarstwowy o 6 neuronach wejściowych, 5 neuronach ukrytych i 1 neuronie wyjściowym).

### 5.5.1 Optymalizacja modeli

**Nie istnieje uniwersalna metoda prowadząca do uzyskania optymalnego modelu neuronowego dla konkretnego problemu.** Poszukiwania najlepszej sieci opierają się na ogół na licznych pracochłonnych eksperymentach, których celem jest znalezienie najwłaściwszych parametrów modelu. Podejmowane decyzje dotyczą stosowanego typu sieci, liczby zmiennych wejściowych, liczby neuronów w warstwie ukrytej, stosowanych funkcji aktywacji, algorytmów uczących itd. Duża liczba „stopni swobody” przy poszukiwaniu optymalnego modelu powoduje, że eksperymenty te bywają często prowadzone metodą prób i błędów, a zatem mogą mieć nieuporządkowany, trudny do prześledzenia i pozornie chaotyczny charakter. Z tego względu na potrzeby niniejszej pracy, celem usystematyzowania postępowania i zwiększenia przejrzystości opisu, opracowano jednolity schemat prowadzenia eksperymentów dla każdej z przewidywanych zmiennych (rys. 17).

**Rysunek 17. Schemat prowadzenia badań zmierzających do stworzenia optymalnego modelu neuronowego dla każdej z przewidywanych zmiennych.**



W każdym przypadku poszukiwania optymalnego modelu obejmowały dwa etapy – ogólny i szczegółowy, poprzedzone analizą statystyczną rozkładu rozważanej zmiennej wyjściowej w badanych grupach chorych i jej zależności od dostępnych zmiennych wejściowych.

Pierwszy etap eksperymentów miał na celu wstępne, orientacyjne ustalenie potencjalnie najbardziej obiecujących parametrów sieci dla rozważanego zagadnienia. W tym etapie dokonywano przede wszystkim systematycznej analizy zmian efektywności działania sieci w zależności od liczby i rodzaju zmiennych wejściowych oraz liczby neuronów ukrytych.

Drugi etap stanowiła próba zbudowania konkretnej sieci możliwie najlepiej modelującej analizowane zjawisko. Eksperymenty prowadzone w drugim etapie opierały się na wynikach uzyskanych w etapie pierwszym, ale ze swej natury miały mniej uporządkowany charakter, ponieważ w przypadku każdego z finalnie wybranych modeli w ostatniej fazie jego optymalizacji na bieżąco podejmowano decyzje o kierunku jego dalszych modyfikacji zależnie od empirycznie ustalanych indywidualnych właściwości sieci.

### **5.5.3 Automatyczny projektant sieci (APS)**

W eksperymentach dotyczących poszukiwania optymalnych modeli wykorzystywano wielokrotnie specjalne narzędzie informatyczne wspomagające tworzenie i testowanie sieci neuronowych, dostępne w pakiecie Statistica 7.0 - tzw. *Automatycznego projektanta sieci* (APS). APS wykorzystuje **zaawansowane algorytmy optymalizacji wieloparametrycznej**, pozwalające na przebadanie znacznej liczby sieci w krótkim czasie oraz ułatwiające automatyczny wybór najlepszych modeli. Wykorzystywane są przy tym zaimplementowane



w programie nowoczesne techniki doboru zmiennych wejściowych, selekcji liczby neuronów ukrytych i optymalizacji innych istotnych parametrów.

Narzędzie to jest bardzo efektywne, ale doświadczenie własne uczy, że w celu uzyskania rzeczywiście optymalnego modelu, poszukiwania prowadzone przez *Automatycznego projektanta* wymagają ścisłego ukierunkowania przez użytkownika. Ponadto tworzone przez APS modele wymagają niejednokrotnie dalszej „ręcznej” optymalizacji, jednak analiza uzyskiwanych przez nie wyników stanowi istotną wskazówkę co do potencjalnie obiecujących kierunków dalszych poszukiwań oraz dostarcza dogodnego punktu wyjścia przy budowaniu ostatecznego modelu.

Wykorzystanie APS ma tę przewagę nad samodzielnym projektowaniem poszczególnych sieci, że daje możliwość **przetestowania i porównania bardzo licznych modeli**, co w przypadku „ręcznego” uczenia sieci byłoby niezwykle mozolne i czasochłonne. Dzięki APS, w przypadku każdego zagadnienia poruszanego w niniejszej pracy możliwe było przeprowadzenie odpowiednio ukierunkowanych eksperymentów, w których każdorazowo testowano setki różnych sieci neuronowych o rozmaitych parametrach i architekturze oraz – co istotne – różnych wstępnych wartościach wag synaptycznych. Ten sposób postępowania miał zwiększyć szansę na odnalezienie minimum globalnego na powierzchni błędu utworzonej dla każdego z analizowanych zagadnień.

#### **5.5.4 Zastosowanie *Automatycznego projektanta sieci* w niniejszej pracy**

**Możliwości *Automatycznego projektanta* wykorzystano w niniejszej pracy na dwa różne sposoby** w kolejnych etapach pracy nad budowaniem modeli.

Jak wspomniano powyżej, w pierwszym (ogólnym) etapie badań prowadzono systematyczne poszukiwania najbardziej obiecującej architektury sieci. Ze względu na znaczny wpływ **efektów losowych** (zwłaszcza wstępnego rozkładu wag synaptycznych) na ostateczną efektywność uzyskiwanej sieci, porównywanie jakości pojedynczych modeli o różnych architekturach nie może stanowić podstawy do uznania konkretnej architektury za „bardziej obiecującą” od innej. Często się zdarza, że sieć o architekturze gorzej nadającej się do modelowania rozważanego zjawiska uzyskuje przypadkiem lepsze wyniki niż model o potencjalnie właściwszej architekturze, ponieważ przypadkiem zostanie dla niej wylosowany bardziej korzystny początkowy zestaw wag. Na ogół jednak sieci o podobnej architekturze (poza przypadkami skrajnie korzystnego lub skrajnie niekorzystnego losowania) cechują się zbliżoną efektywnością w rozwiązywaniu konkretnego zadania.

Z tego względu w pierwszym etapie **dla każdej zadanej architektury tworzono przy pomocy *Automatycznego projektanta* grupę 100 różnych sieci**. Następnie w sposób globalny porównywano ich efektywność z grupami zawierającymi także po 100 uzyskanych analogicznie sieci o innych architekturach. Na tej podstawie dokonywano wyboru parametrów modeli potencjalnie najkorzystniej wpływających na działanie sieci.

Przebieg eksperymentów w etapie pierwszym był zatem następujący:

- Poszukując najbardziej obiecującego **wektora wejściowego**, dla każdego z testowanych zestawów zmiennych wejściowych tworzono 100 sieci o 1-12 neuronach ukrytych (dla grupy I) lub 1-6 neuronach ukrytych (dla dwukrotnie mniej licznej grupy II), przy czym *Automatyczny projektant* miał możliwość redukcji liczby zmiennych wejściowych. Wartości brakujących danych zastępowano wartością średnią ze wszystkich dostępnych obserwacji.
- Poszukując najkorzystniejszego **rozmiaru drugiej warstwy sieci**, zmieniano stopniowo ilość neuronów ukrytych, każdorazowo tworząc grupę liczącą 100 sieci o zadanej z góry liczbie neuronów ukrytych. Następnie dokonywano porównań efektywności uzyskanych grup sieci. Również na tym etapie *Automatyczny projektant* miał możliwość redukcji liczby zmiennych wejściowych, a braki danych zastępowano średnią.
- Ponadto dla modeli neuronowych służących do predykcji pierwszej z analizowanych zmiennych przeprowadzono w tym etapie dodatkowe analizy dotyczące podziału całego zbioru danych o leczonych i badanych pacjentach na podzbiory **uczący, walidacyjny i testowy**. Posłużono się w tym celu analogiczną metodą jak w przypadku poszukiwań najkorzystniejszego wektora wejściowego.
- Podobną metodą posługiwano się za każdym razem, gdy oceniano wpływ zmiany architektury na efektywność sieci.

W drugim etapie (szczegółowym) dążono natomiast do zbudowania jednego konkretnego modelu, możliwie najlepiej nadającego się do rozwiązania rozważanego zadania. W tym przypadku – w przeciwieństwie do sytuacji omawianej powyżej – znalezienie najkorzystniejszego wstępnego rozlosowania wag było szczególnie pożądane. W tej fazie badań, jeśli uznano to za celowe, **przy pomocy *Automatycznego projektanta* budowano 1000 różnych sieci o narzuconej z góry architekturze**, ustalonej na podstawie rozważań przeprowadzonych w etapie pierwszym. Spośród tych 1000 modeli analizowano 50 o najniższych błędach końcowych po procesie uczenia ocenianych na podstawie walidacji i wybierano jeden model o największej efektywności. Sposób oceny efektywności sieci omówiono poniżej.

Opisany powyżej **algorytm postępowania** został opracowany przez autorkę na potrzeby niniejszej pracy i opiera się w głównej mierze na dotychczasowym własnym doświadczeniu w budowaniu modeli neuronowych. Może on jednak być rekomendowany jako zweryfikowany doświadczalnie sposób wyboru architektury sieci neuronowej przy rozwiązywaniu także innych zagadnień z wykorzystaniem rozważanej tu techniki informatycznej.

#### **5.5.5 Narzędzia wykorzystywane w optymalizacji wektora wejściowego (zestawu zmiennych wejściowych)**

Ponieważ *Automatyczny projektant* jest w stanie dokonać optymalizacji wektora wejściowego sieci tylko w ograniczonym zakresie, w celu doboru zbioru zmiennych wejściowych posługiwano się następującymi metodami dostępnymi w pakiecie *Statistica 7.0*:

- **Analiza statystyczna wpływu poszczególnych zmiennych wejściowych na rozważane zmienne wyjściowe.** Każdy z podrozdziałów niniejszej pracy, dotyczący zastosowania sieci neuronowych dla prognozowania wartości poszczególnych zmiennych wyjściowych, jest poprzedzony wynikami tego rodzaju analizy.
- **Analiza wzajemnych korelacji** zmiennych wejściowych.
- **Analiza wrażliwości modelu dla poszczególnych wejść sieci neuronowej.** Jest to metoda pozwalająca zidentyfikować te zmienne, które niewiele wnoszą do wyniku działania sieci. Polega ona na ocenie straty ponoszonej przy odrzuceniu konkretnej zmiennej ze zbioru uczącego. Miarą wrażliwości jest iloraz błędu uzyskanego przez sieć bez danej zmiennej wejściowej i błędu uzyskanego z kompletem zmiennych wejściowych. Iloraz wynoszący  $\leq 1$  sugeruje, że oceniana zmienna jest zbędna. Na podstawie analizy wrażliwości można także uporządkować zmienne wejściowe pod względem ich ważności, przypisując im kolejne rangi. Wyniki tej prostej analizy należy jednak traktować z rozwagą, ponieważ ze względu na potencjalnie istotne dla działania sieci efekty synergii między zmiennymi wejściowymi, analiza oparta na badaniu skutków usuwania pojedynczych zmiennych może nie odzwierciedlać rzeczywistego znaczenia poszczególnych wejść sieci.
- **Metody doboru cech wykorzystujące sieci neuronowe probabilistyczne:**
  - **algorytmy krokowe** (metoda **krokowa postępująca** i **krokowa wsteczna**, polegające na stopniowym dodawaniu lub usuwaniu po jednej zmiennej z następującą oceną wpływu tych zmian na jakość predykcji dokonywanych przez sieci neuronowe),
  - **algorytmy genetyczne** (strategia optymalizacyjna opisana nieco szerzej we wstępie do niniejszej pracy, bazująca na teorii ewolucji i doboru naturalnego, z reguły prowadząca do uzyskiwania kolejnych „pokoleń” coraz skuteczniejszych sieci neuronowych).

W przypadku algorytmów krokowych i algorytmów genetycznych jako punkt wyjścia przyjmowano wektor złożony ze wszystkich dostępnych zmiennych, a brakujące dane usuwano wraz z zawierającymi je przypadkami uczącymi. Doboru cech (tj. użytych w modelu zmiennych wejściowych) dokonywano za każdym razem kilkakrotnie, modyfikując przy tym tzw. „karę za cechę” decydującą o ostatecznym rozmiarze wektora wejściowego. Ostatecznie do dalszych analiz wybierano taki wektor, dla którego uzyskano najmniejszą wartość błędu (pojęcie błędu sieci neuronowej wyjaśniono dokładniej poniżej).

### **5.5.6 Ocena efektywności sieci neuronowych**

Oceniając i porównując efektywność działania uzyskiwanych sieci neuronowych posługiwano się dwoma pojęciami:

- **Jakość uczenia, walidacji i testowania.** W przypadku sieci klasyfikujących jakość sieci ustala się w sposób ilościowy oznaczając odsetek prawidłowo

zaklasyfikowanych przypadków odpowiednio w podzbiorach: uczącym, walidacyjnym i testowym. Jakość testowania wynosząca 70% oznacza zatem, że dla 70% przypadków ze zbioru testowego sieć neuronowa podaje prawidłowe wyniki.

- **Błąd (dla podzbioru uczącego, walidacyjnego i testowego).** W zadaniach odwołujących się do stosowania sieci neuronowych najczęściej stosuje się miarę błędu określaną jako SSE lub MSE. Jest to miara oparta na sumie kwadratów różnic między otrzymaną i rzeczywistą wartością wyjścia sieci dla wszystkich indywidualnych przypadków wchodzących w skład danego podzbioru. Błędy podczas sumowania podnosi się do kwadratu żeby uniknąć efektu kompensowania błędów dodatnich przez błędy ujemne (po podniesieniu do kwadratu każdy błąd jest dodatni), a ponadto podnoszenie do kwadratu uwrażliwia używaną miarę na duże błędy (na przykład trzy razy większy błąd ma wtedy dziewięć razy większy udział w sumie). Uzyskaną sumę kwadratów błędów SSE (*Sum Square Errors*) dzieli się przez liczbę elementów rozważanego zbioru przypadków w celu uzyskanie błędu uśrednionego (przypadającego na pojedynczy przypadek) MSE (*Mean Square Errors*) oraz często wyciąga się pierwiastek z MSE w celu uzyskania wartości błędów w tych samych jednostkach, w jakich mierzona jest wartość wyjściowa sieci, co ułatwia ich interpretację. Uczenie sieci polega na minimalizacji błędu MSE dla podzbioru uczącego. Błąd MSE jest bardzo intuicyjny, stosowany jest jednak głównie w zadaniach regresyjnych, gdy sieć po zakończeniu nauki ma podawać na wyjściu wartości liczbowe będące rozwiązaniem analizowanego problemu. Błąd ten jest jednak mniej przydatny w zadaniach klasyfikacyjnych, gdy sieć powinna na wyjściu wskazywać właściwą kategorię, do której można zaliczyć obiekt scharakteryzowany przez dane wejściowe. Taki właśnie charakter miały zadania rozwiązywane w tej pracy i z tego powodu w opisywanych dalej eksperymentach wykorzystywano jako miarę błędu tak zwaną entropię wzajemną. Entropia, wprowadzona do teorii informacji przez Claude'a E. Shannona, jest ogólnie miarą niepewności informacyjnej [223]. Wzrost entropii oznacza zawsze wdzieranie się chaosu do rozważanych procesów lub zjawisk. Jest więc logiczne, że proces wprowadzania porządku poprzez automatycznie generowaną klasyfikację, oceniać będziemy właśnie stopniem zmniejszenia entropii. Błąd jest wtedy sumą iloczynów zadanych wartości oraz logarytmów błędów wyznaczanych dla każdego neuronu wyjściowego. Funkcja ta została specjalnie zaprojektowana dla problemów klasyfikacyjnych, gdzie stosowana jest łącznie z wykorzystywaną w warstwie wyjściowej logistyczną funkcją aktywacji. Malejąca wartość tak zdefiniowanego błędu oznacza malejącą niepewność podejmowanych decyzji i wskazuje na rosnącą stabilność budowanej w trakcie uczenia klasyfikacji.

W przypadku sieci neuronowych praktycznie nigdy nie da się stwierdzić z całą pewnością, że opracowany model jest najlepszym z możliwych do rozwiązania rozważanego zagadnienia. **Dokonując wyboru finalnych modeli prezentowanych w niniejszej pracy kierowano się**

**następującymi kryteriami** (bazującymi na cytowanej wcześniej literaturze [186, 187, 189] i doświadczeniu własnym):

- najniższy **błąd walidacji i testowania**,
- najwyższa **jakość walidacji i testowania**,
- błędy walidacji i testowania nie są znacząco wyższe, a jakości walidacji i testowania znacząco niższe niż (odpowiednio) błąd i jakość uczenia (brak spełnienia tego kryterium oznaczałoby niedostateczną **zdolność modelu do generalizacji** i sygnalizowałoby potencjalne przeuczenie),
- brak ewidentnie zbędnych zmiennych wejściowych potwierdzony w **analizie wrażliwości** (brak spełnienia tego kryterium sugerowałoby, że istnieje możliwość dalszych modyfikacji i poprawy działania modelu),
- wartość **parametru  $\rho$**  (określającego stosunek liczby przypadków uczących do całkowitej liczby połączeń występujących w sieci jak opisano w rozdziale 4.9) zawarta w przedziale sugerującym odpowiednią złożoność modelu,
- brak ewidentnych „**artefaktów**” w dokonywanych predykcjach (na przykład przewidywania zawsze takich samych wyników u wszystkich pacjentów).

### **5.5.7 Podział zbioru danych na podzbiory: uczący, walidacyjny i testowy**

Jak zaznaczono w podrozdziale 4.12 „Sztuczne sieci neuronowe – wprowadzenie i przegląd piśmiennictwa”, do prawidłowego uczenia i testowania sieci neuronowej zbiorów dostępnych przypadków, na których przeprowadza się formowanie neuronowego modelu, należy podzielić na podzbiory: uczący, walidacyjny i testowy. **Każdy z tych podzbiorów powinien zawierać dane reprezentatywne dla całej opisywanej zależności.** Niewłaściwy podział posiadanej bazy danych na podzbiory, a w szczególności włączenie obserwacji prezentujących mniej typowe dane do zbioru testowego, może prowadzić do błędów w procesie uczenia (informacja o właściwościach rozważanych danych dostępna podczas uczenia pozbawiona tych danych będzie niepełna). Ponadto obecność obserwacji nietypowych w zbiorze testowym skutkuje nieprawidłową interpretacją wyników ewaluacji modelu (niemal nieuchronnie błędne wyniki uzyskiwane podczas testowania dla obserwacji istotnie różniących się od dostępnych w zbiorze uczącym mogą wprowadzać w błąd eksperymentatora).

W celu wyodrębnienia podzbiorów: uczącego, testowego i walidacyjnego, z których każdy byłby możliwie najbardziej reprezentatywny dla całego zbioru danych, w niniejszej pracy posłużono się wspomnianą w podrozdziale 4.12 metodą wstępnego orientacyjnego podziału dostępnych przypadków na skupienia. Szczegóły metodyki podziału opracowano na potrzeby niniejszej pracy, bazując na wcześniejszych własnych doświadczeniach, ale można je potraktować jako potencjalnie przydatne także i dla innych badaczy.

**Podział na skupienia** przeprowadzono w programie Statistica 7.0, wykorzystując jeden z najpopularniejszych algorytmów służących do analizy skupień, tj. grupowanie metodą k-średnich. Algorytm ten polega na podziale przypadków na k skupień, z których każde jest

zwarte wewnętrznie i w możliwie największym stopniu różni się od pozostałych. Program przyporządkowuje przypadki do odpowiednich skupień w taki sposób, aby ostatecznie uzyskać najbardziej istotne wyniki analizy wariancji. [224]

Na potrzeby niniejszej pracy grupowanie metodą k-średnich przeprowadzono wykorzystując wszystkie dostępne zmienne, zarówno dotyczące okresu przedoperacyjnego, jak i pooperacyjnego. W celu ustalenia wstępnych centrów skupień wybrano obserwacje tak, by zmaksymalizować odległości skupień (potem algorytm k-średnich optymalnie dopasowuje centra do rzeczywistego grupowania się danych, ale ponieważ działa on iteracyjnie, więc wybór wstępnego przybliżenia jest obligatoryjny).

Założono, że z każdego ze skupień zostaną wybrane przypadki uczące, walidacyjne i testowe w proporcjach 60% : 20% : 20%. Tego rzędu proporcje często bywają wykorzystywane w praktyce przy tworzeniu sieci neuronowych przeznaczonych do zastosowań badawczych [225]. Minimalna wielkość skupienia, które można według powyższych założeń, wynosi 5 przypadków (ponieważ wyodrębnienie 3 przypadków uczących, 1 walidacyjnego i 1 testowego odpowiada proporcjom 60% : 20% : 20%). Na tej podstawie przeprowadzono analizę liczebności skupień zależnie od założonej *a priori* ich ilości, aby ostatecznie wybrać maksymalną liczbę skupień, z których jednak żadne nie liczyłyby mniej niż 5 przypadków.

Wyniki obliczeń dla grupy I zestawiono w tabeli 4. Szczegółowa analiza wyników wykazała, że jeden z przypadków zawsze był grupowany oddzielnie. Skupienia, do których był zaliczany ten przypadek wyróżniono w tabeli. Uznano, że bezsprzecznie stanowi on obserwację odstającą, zaburzającą strukturę zbioru danych, i w związku z tym zdecydowano o jego usunięciu. Następnie ponownie przeprowadzono analizę skupień, której wyniki przedstawia tabela 5.

**Tabela 4. Grupa I - liczba przypadków w poszczególnych skupieniach przy różnej ilości skupień.**

Liczba skupień	Nr skupienia									
	1	2	3	4	5	6	7	8	9	10
10	12	27	7	18	2	19	28	14	2	11
9	21	16	12	23	27	1	5	7	28	
8	10	27	11	16	1	36	25	14		
7	29	37	11	22	22	1	18			

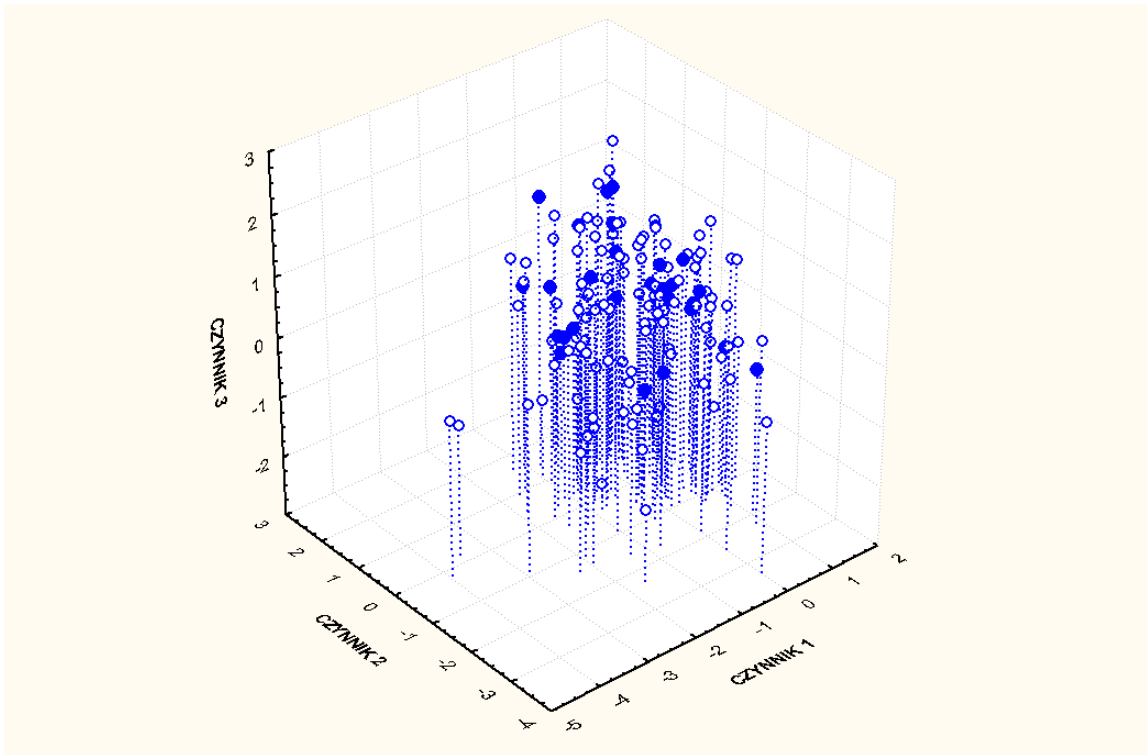
**Tabela 5. Grupa – liczba przypadków w poszczególnych skupieniach przy różnej ilości skupień (po usunięciu pojedynczej obserwacji odstającej) .**

Liczba skupień	Nr skupienia								
	1	2	3	4	5	6	7	8	9
9	18	23	15	12	1	31	8	10	21
8	14	26	9	17	7	36	18	12	
7	28	13	12	35	14	19	18		

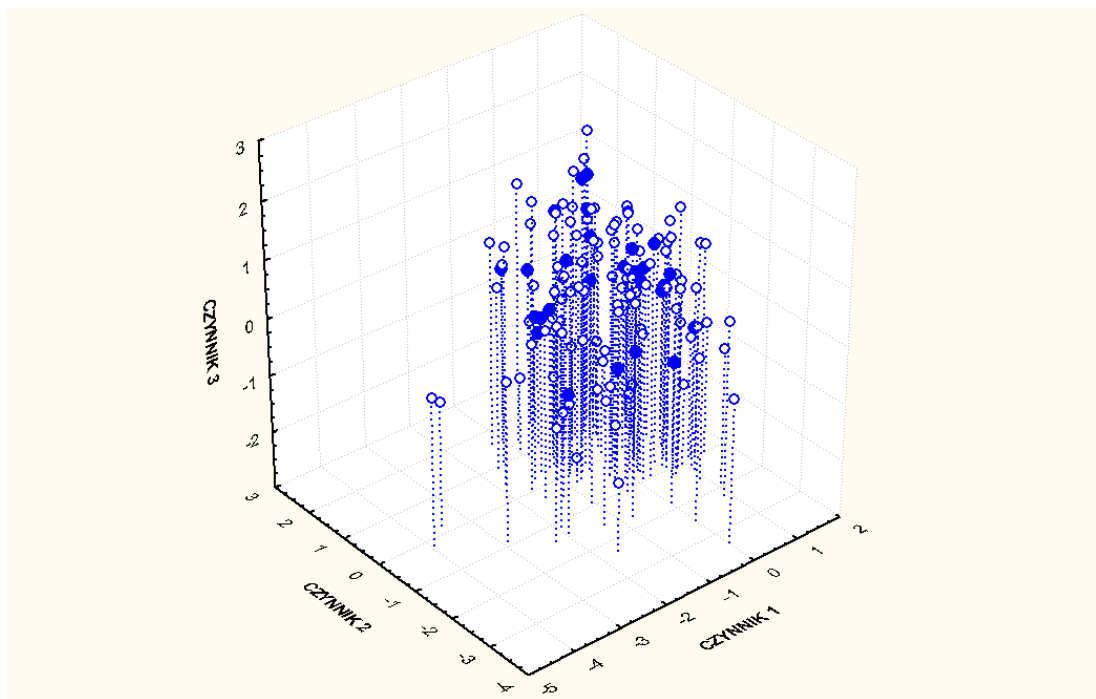
Zgodnie z przyjętymi wcześniej założeniami zdecydowano ostatecznie dokonać podziału zbioru danych na 8 skupień. Przypadki należące do każdego z wyodrębnionych skupień przyporządkowano do podzbiorów: uczącego, walidacyjnego i testowego w przybliżeniu w podanych powyżej proporcjach 60% : 20% : 20%. Przypadki najbliższe i najdalsze od centrum skupień zaliczono arbitralnie do zbioru uczącego, natomiast położone w pośrednich odległościach – do testowego i walidacyjnego. Postępowanie to miało na celu uwzględnienie w zbiorze uczącym możliwie najszerszego spektrum informacji o dostępnym zbiorze danych, a jednocześnie stosując podaną regułę nie dopuszczono do sytuacji, w której potencjalne obserwacje skrajne znalazłyby się w zbiorze walidacyjnym lub testowym, co prowadziłyby do wspomnianych wyżej niekorzystnych skutków (zubożenia zbioru uczącego, fałszywego kryterium zakończenia uczenia oraz nieadekwatnych wyników testowania). W efekcie do zbioru uczącego zaliczono 85 przypadków (61,15%), a do zbioru walidacyjnego i testowego odpowiednio po 27 przypadków (19,42%).

Po dokonaniu podziału, w celu przeprowadzenia orientacyjnej wzrokowej oceny wybranych podzbiorów, posłużono się analizą składowych głównych, którą również przeprowadzono z wykorzystaniem oprogramowania Statistica 7.0. Wykorzystano w tym celu te same zmienne, które były użyte w analizie skupień. Brakujące dane zastępowano średnimi. Wyodrębniono 3 składowe główne, co pozwoliło na odwzorowanie rozkładu danych w umożliwiającej ocenę wzrokową przestrzeni trójwymiarowej. Rozkład przypadków testowych w tej przestrzeni przedstawia rysunek 18, na którym zbiór testowy reprezentują niebieskie punkty, a zbiory uczący i walidacyjny – białe punkty. Analiza wykresu sugeruje, że w zbiorze testowym mogły się znaleźć potencjalne obserwacje skrajne. Ponieważ ocena tego rodzaju w odniesieniu do rozważanej bazy danych miała charakter wyłącznie orientacyjny, nie zdecydowano o usunięciu tych obserwacji, przeniesiono je jednak do zbioru uczącego i zamiast nich do zbioru testowego zaliczono inne przypadki należące do tych samych skupień. Rozkład przypadków testowych względem pozostałych po przeprowadzonej korekcie przedstawia rysunek 19.

**Rysunek 18. Grupa I – wyjściowy rozkład przypadków testowych w trójwymiarowej przestrzeni, której wymiarami są czynniki wyodrębnione w analizie składowych głównych (niebieskie punkty – przypadki testowe, białe punkty – pozostałe przypadki).**



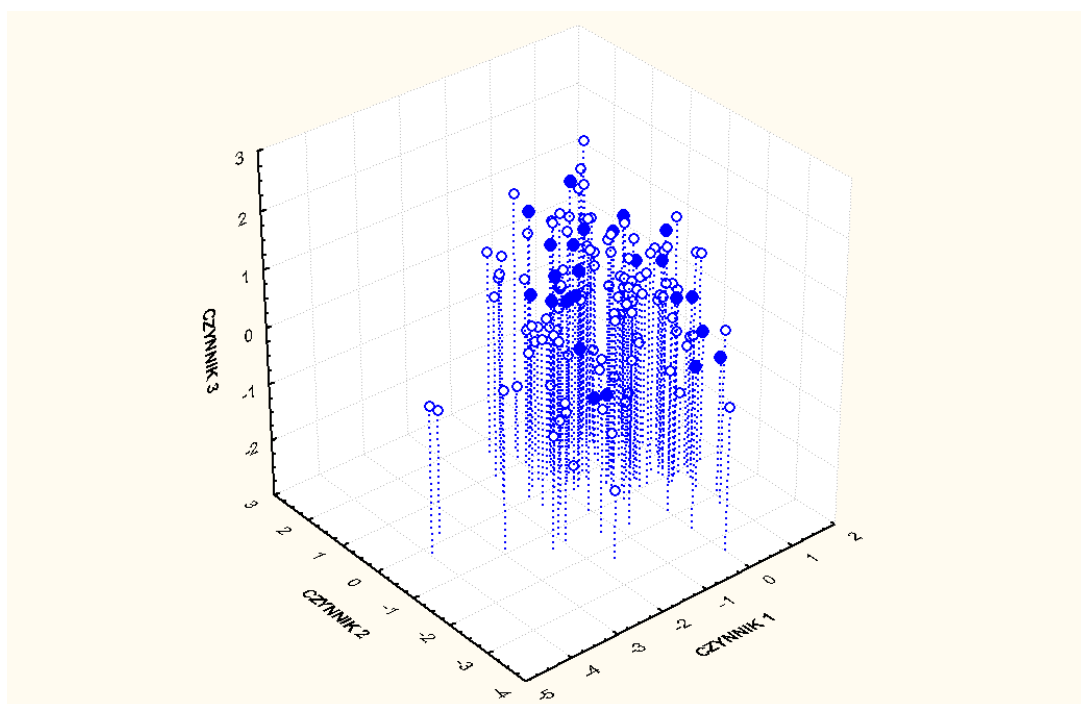
**Rysunek 19. Grupa I – rozkład przypadków testowych w trójwymiarowej przestrzeni po przeprowadzeniu korekty – opis w tekście (niebieskie punkty – przypadki testowe, białe punkty – pozostałe przypadki).**





Analogiczne postępowanie przeprowadzono dla przypadków zaliczonych do zbioru walidacyjnego, ostatecznie uzyskując ich rozkład jak na rysunku 20 Rysunek 20.

**Rysunek 20. Grupa I – ostateczny rozkład przypadków walidacyjnych w trójwymiarowej przestrzeni – opis w tekście (niebieskie punkty – przypadki walidacyjne, białe punkty – pozostałe przypadki).**

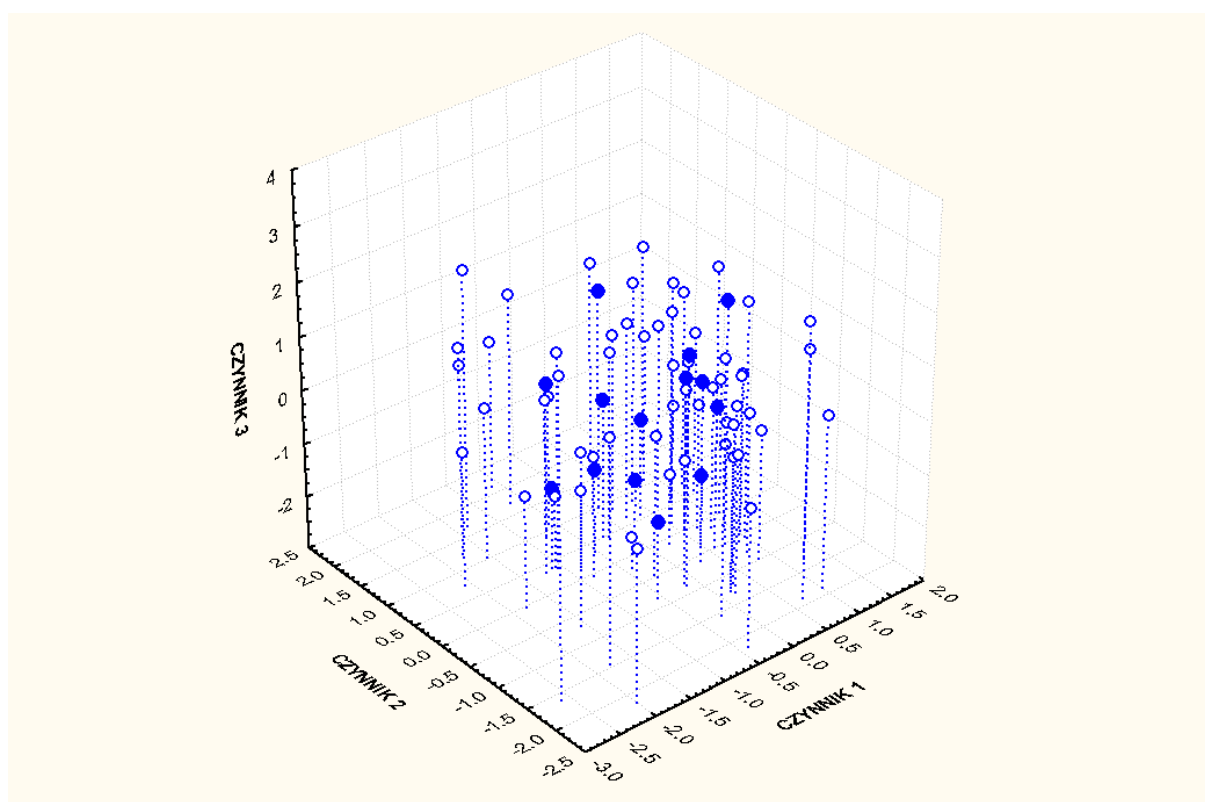


Następnie takiej samej procedurze poddano przypadki z grupy II. Liczbę przypadków w poszczególnych skupieniach zależnie od z góry założonej ilości skupień dla tej grupy pacjentów przedstawia tabela 6. Na podstawie analizy zawartości tej tabeli zdecydowano o podziale zbioru danych na 6 skupień. Przypadki należące do każdego ze skupień przydzielono odpowiednio do podzbiorów: uczącego, walidacyjnego i testowego zgodnie z opisanymi wcześniej zasadami. Łącznie do zbioru uczącego przyporządkowano 45 przypadków (62.5%), do zbioru walidacyjnego 13 przypadków (18.06%), a do zbioru testowego 14 przypadków (19.44%). Dalsze postępowanie przeprowadzono tak samo jak w grupie I. Ostateczny rozkład przypadków testowych i walidacyjnych w przestrzeni trójwymiarowej przedstawiają rysunki 21 i 22.

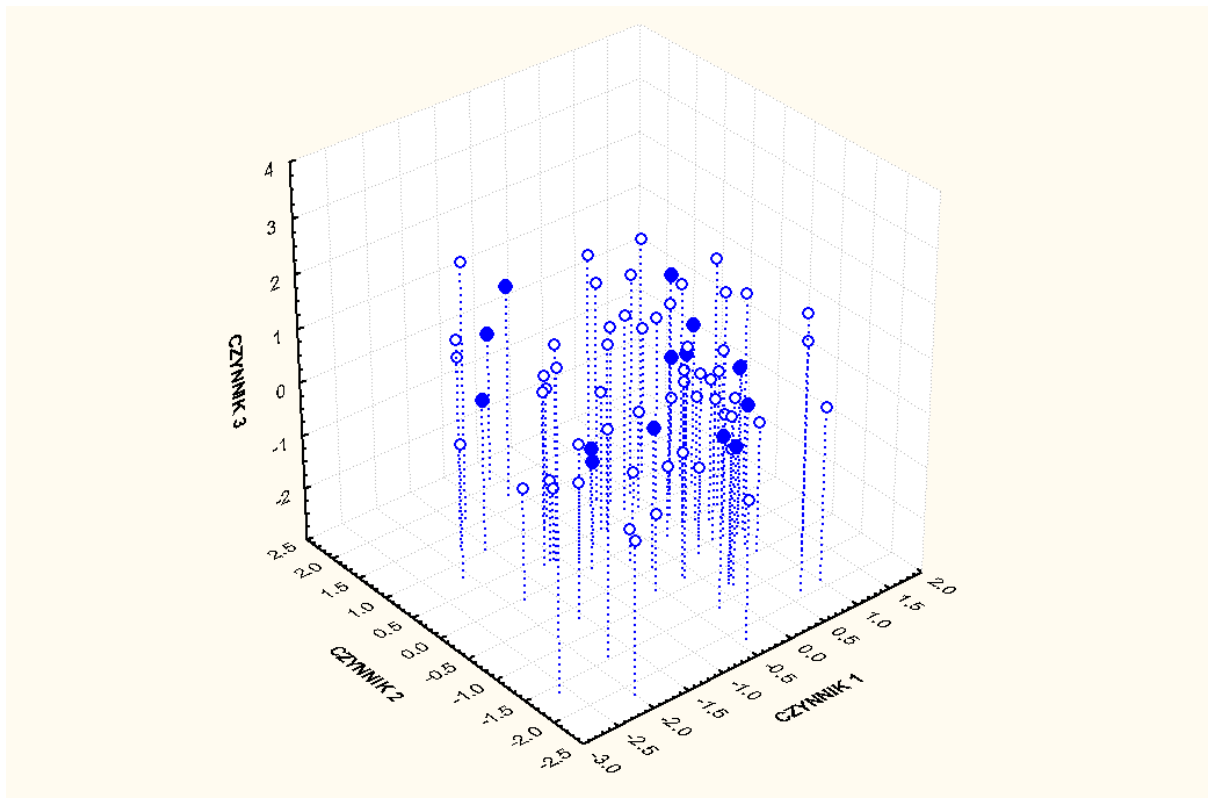
Tabela 6. Grupa II - liczba przypadków w poszczególnych skupieniach przy różnej ilości skupień.

Liczba skupień	Nr skupienia									
	1	2	3	4	5	6	7	8	9	10
10	2	2	3	6	8	11	7	13	7	13
9	4	11	2	10	7	6	18	9	5	
8	13	3	4	7	7	11	18	9		
7	6	17	17	4	10	10	8			
6	13	12	13	9	10	15				

Rysunek 21. Grupa II – ostateczny rozkład przypadków testowych w trójwymiarowej przestrzeni (niebieskie punkty – przypadki testowe, białe punkty – pozostałe przypadki).



**Rysunek 22. Grupa II – ostateczny rozkład przypadków walidacyjnych w trójwymiarowej przestrzeni (niebieskie punkty – przypadki walidacyjne, białe punkty – pozostałe przypadki).**



## 6. Wyniki

### 6.1 Uwagi wstępne

Kolejne podrozdziały w niniejszej części pracy dotyczą predykcji wyników ESS w odniesieniu do następujących zagadnień klinicznych (kodowanych jako odpowiednie zmienne wyjściowe sieci neuronowych):

- nasilenie dolegliwości po zabiegu operacyjnym (w skali VAS),
- zmiana nasilenia dolegliwości po zabiegu operacyjnym w porównaniu do stanu przed zabiegiem (w skali VAS),
- nasilenie poszczególnych objawów po zabiegu operacyjnym (w skali 4-stopniowej)
  - upośledzenie drożności nosa,
  - wydzielina z nosa,
  - spływanie wydzieliny po tylnej ścianie gardła,
  - zaburzenia węchu,
  - ból głowy,
  - ból i uczucie rozpierania w obrębie twarzy,
  - przewlekłe zmęczenie.

Wyniki w poszczególnych podrozdziałach zostały przedstawione według jednolitego porządku (por. rys. 17 w rozdziale „Materiał i metoda”). Każdej z przewidywanych zmiennych wyjściowych dotyczy oddzielny podrozdział, omawiający kolejno:

- analizy statystyczne (rozkład badanej zmiennej w obu grupach pacjentów wymienionych w rozdziale „Materiał i metoda”, zależności między badaną zmienną wyjściową a zmiennymi wejściowymi oraz sposób przekształcenia zmiennej w zmienną binarną, wprowadzoną na potrzeby korzystania z klasyfikacyjnych sieci neuronowych),
- tworzenie modeli neuronowych:
  - etap I – poszukiwanie najbardziej obiecującej architektury (wektora wejściowego, liczby neuronów w warstwie ukrytej),
  - etap II – budowanie finalnego modelu.

Dla pierwszej z omawianych zmiennych wyjściowych, po wstępnych analizach dotyczących wektora wejściowego, przeprowadzono dodatkowo także ewaluację przyjętego na wstępie podziału na podzbiory uczący, walidacyjny i testowy, dokonanego na podstawie analizy skupień. Przeprowadzono również ocenę przyjętych proporcji przypadków w podzbiorach uczącym, walidacyjnym i testowym.

**W pierwszym podrozdziale zawierającym prezentację wyników przytaczane rezultaty opatrzone obszernym komentarzem dotyczącym przebiegu prac nad tworzeniem modeli neuronowych. W odniesieniu do kolejnych eksperymentów, przeprowadzonych według analogicznych zasad, zwięźle przedstawiono tylko uzyskane wyniki, ponieważ przebieg prac był każdorazowo taki sam.**

## 6.2 Nasilenie dolegliwości po zabiegu operacyjnym

Nasilenie dolegliwości związanych z PZZP wyrażone w skali VAS opisano przy pomocy zmiennej „VAS”, przyjmującej wartości od 0 do 10. Przyjęto następujące nazewnictwo zmiennych:

- VAS\_0 – nasilenie dolegliwości w skali VAS przed zabiegiem,
- VAS\_6 – nasilenie dolegliwości w skali VAS po 3-6 miesiącach (dotyczy grupy I),
- VAS\_12 – nasilenie dolegliwości w skali VAS po 12 miesiącach (dotyczy grupy II).

### 6.2.1 Analiza statystyczna

Wartości zmiennej VAS, zanotowane przed i po zabiegu dla poszczególnych grup chorych, zestawiono w tabelach 7 i 8. W tabeli 9 przedstawiono zależności między pooperacyjnymi wartościami tej zmiennej, a zmiennymi traktowanymi w budowanym modelu jako zmienne niezależne (wejściowe).

**Tabela 7. Wartości zmiennej VAS w grupie I (140 pacjentów) – przed zabiegiem i 3-6 miesięcy po zabiegu**

Cecha	Wartości	Brak danych
dolegliwości przed zabiegiem (VAS_0)	zakres: 1-10; mediana: 8	2 (1%)
dolegliwości 3-6 miesięcy po zabiegu (VAS_6)	zakres: 0-10, mediana: 3	3 (4%)

**Tabela 8. Wartości zmiennej VAS w grupie II (72 pacjentów) – przed zabiegiem i 12 miesięcy po zabiegu**

Cecha	Wartości	Brak danych
dolegliwości przed zabiegiem (VAS_0)	zakres: 2-10; mediana: 8	0
dolegliwości 12 miesięcy po zabiegu (VAS_12)	zakres: 0-10, mediana: 3	2 (3%)

**Tabela 9. Zależności między pooperacyjnym nasileniem objawów wyrażonym w skali VAS a zmiennymi wejściowymi. Rodzaje zastosowanych testów statystycznych omówiono w rozdziale „Material i metoda”. Wartości poziomu istotności p podano tylko dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	VAS_6 - po 3-6 miesiącach (grupa I)	VAS_12 - po 12 miesiącach (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	p = 0.014964	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	p = 0.023425 R = 0.196459	-
astma	p = 0.022050	-
AIA	p = 0.037272	-
alergia	-	-
refluks	p = 0.008177	(blisko granicy istotności) p = 0.050683
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.000917	p = 0.023261
liczba przebytych ESS	p = 0.000545 R = 0.292681	p = 0.015342 R = 0.288747
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS	p = 0.040004 R = 0.177664	p = 0.000842 R = 0.390122
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	p = 0.036743 R = 0.179298	p = 0.002572 R = 0.359923

Cecha	VAS_6 - po 3-6 miesiącach (grupa I)	VAS_12 - po 12 miesiącach (grupa II)
<b>kaszel</b>	-	p = 0.000169 R = 0.443801
<b>kichanie</b>	-	-
<b>ból twarzy</b>	p = 0.048218 R = 0.169728	p = 0.037775 R = 0.252492
<b>ból czoła</b>	p = 0.007135 R = 0.230573	p = 0.001779 R = 0.380295
<b>ból głowy</b>	p = 0.015703 R = 0.208345	p = 0.005146 R = 0.338034
<b>dolegliwości uszne</b>	p = 0.018202 R = 0.202265	p = 0.045199 R = 0.245542
<b>zaburzenia węchu</b>	-	-
<b>zmęczenie</b>	p = 0.004527 R = 0.243817	p = 0.007287 R = 0.320339
<b>nieprzyjemny zapach z ust</b>	p = 0.027178 R = 0.189446	p = 0.020788 R = 0.277879
<b>ból zębów</b>	-	-

Na potrzeby sieci klasyfikacyjnych przekształcono zmienną VAS w zmienną binarną (przyjmującą wartości 0 lub 1, gdzie 0 oznacza zadowolający wynik operacji, natomiast 1 – wynik niezadowolający). W tym celu niezbędne było określenie granicznej wartości zmiennej VAS, poniżej której efekt operacji może być uznany za pozytywny (co koduje przekształcona wartość wynosząca 0). Oczywiście każdy tego rodzaju arbitralnie przyjęty „punkt odcięcia” może budzić wątpliwości, ponieważ dla wartości zmiennej bliskich progu pacjenci, u których nasilenie objawów różni się zaledwie o 1 punkt w skali VAS, zostają przyporządkowani do różnych klas [222]. Dla uniknięcia opisanego problemu niekiedy przyjmuje się dwa progi: wartości poniżej pierwszego progu klasyfikowane są jako „0”, wartości powyżej drugiego progu klasyfikowane są jako „1”, zaś wartości między jednym i drugim progiem wiązane są z brakiem decyzji (odpowiedź „nie wiadomo”). W tej pracy zrezygnowano z takiego sposobu rozgraniczania klas ze względu na wiążące się z nim kłopoty w uczeniu sieci neuronowej oraz w interpretacji dostarczanych przez nią wyników (po procesie uczenia). Jednak możliwość takich neutralnych decyzji trzeba mieć na względzie przy prowadzeniu dalszych badań na większych zbiorach danych.

Ustalając próg binaryzacji w niniejszej pracy kierowano się kryteriami klinicznymi. W badaniu dotyczącym wpływu PZZP na codzienne funkcjonowanie pacjentów Lim i wsp. stwierdzili, że obniżenie jakości życia notuje się głównie u tych chorych, u których nasilenie

dolegliwości wyrażone w skali VAS wynosi 5 lub więcej [11]. W oparciu o wyniki tych badań przyjęto próg binaryzacji VAS=5. Nową zmienną binarną (nazwaną „VAS\_bin”) zdefiniowano następująco:

- jeżeli  $VAS < 5$  to  $VAS\_bin = 0$  (wynik operacji jest zadowolający, ponieważ dolegliwości nie obniżają istotnie jakości życia pacjenta),
- jeżeli  $VAS \geq 5$  to  $VAS\_bin = 1$  (wynik operacji nie jest zadowolający).

Nazwy nowych zmiennych binarnych ustalono według schematu:

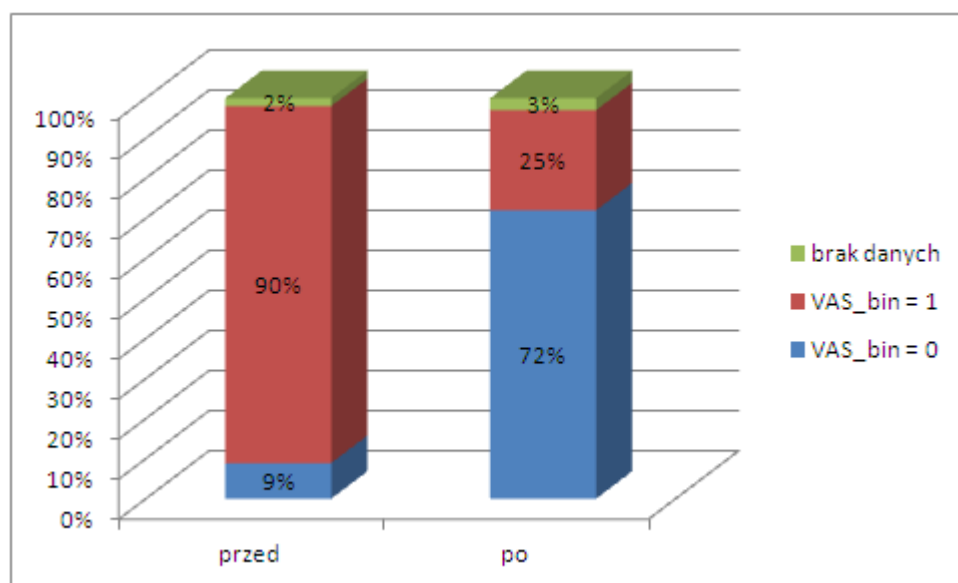
- dla VAS\_0 (przed zabiegiem): VAS\_bin\_0,
- dla VAS\_6 (3-6 miesięcy po zabiegu): VAS\_bin\_6,
- dla VAS\_12 (12 miesięcy po zabiegu): VAS\_bin\_12.

**Model predykcyjny przewidujący wartość zmiennej VAS\_bin odpowiada zatem na pytanie pacjenta:**

***„Czy po operacji moje dolegliwości związane z przewlekłym zapaleniem zatok przynosowych będą na tyle mało uciążliwe, że nie będą obniżały mojej jakości życia?”***

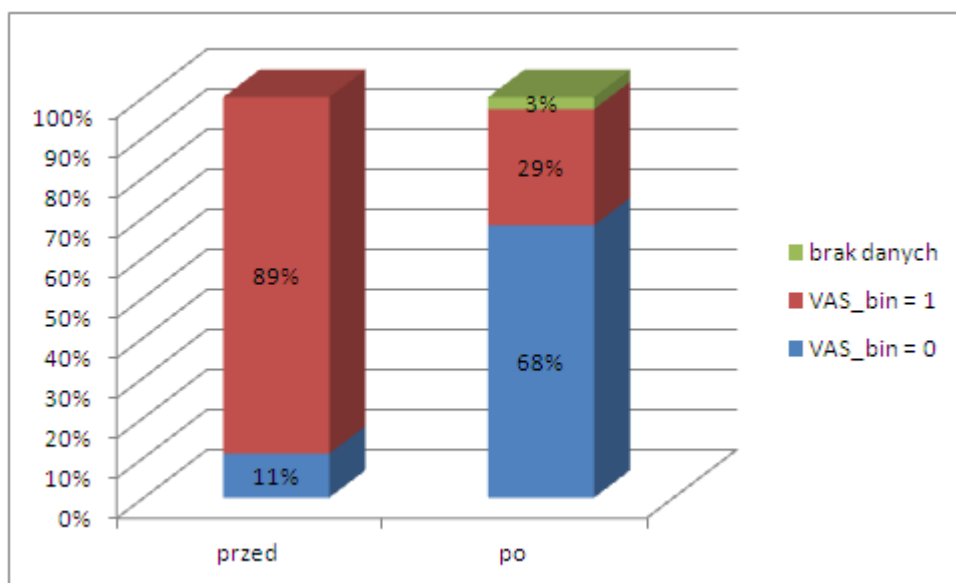
Wartości zmiennej VAS\_bin w obu grupach chorych przed zabiegiem i po zabiegu zilustrowano na rysunkach 23 i 24, a ich zależności od innych zmiennych zestawiono w tabeli 10.

**Rysunek 23. Odsetki pacjentów w grupie I, którzy zgłaszali odpowiednio zaznaczone na rysunku wartości zmiennej VAS\_bin– przed zabiegiem i 3-6 miesięcy po zabiegu**





**Rysunek 24.** Odsetki pacjentów w grupie II, którzy zgłaszali odpowiednio zaznaczone na rysunku wartości zmiennej VAS\_bin – przed zabiegiem i 12 miesięcy po zabiegu



**Tabela 10.** Zależności między pooperacyjnym nasileniem objawów wyrażonym przy pomocy zmiennej VAS\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano tylko dla zależności statystycznie istotnych.

Cecha	VAS_bin_6 (grupa I)	VAS_bin_12 (grupa II)
pleć	-	p=.04601
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	(blisko granicy istotności) p = 0.059454	-
astma	(blisko granicy istotności) p=.06979	p=.03588
AIA	p=.04975	-
alergia	-	-
refluks	p=.04040	-
przebyte polipektomie	-	-

Cecha	VAS_bin_6 (grupa I)	VAS_bin_12 (grupa II)
liczba przebytych polipektomii	-	-
przebyte ESS	p=.00016	-
liczba przebytych ESS	-	p = 0.032671
nasilenie dolegliwości przed zabiegiem:		
VAS_0	p = 0.025426	p= 0.003820
niedrożność nosa	-	-
wydzielina z nosa	-	p = 0.032008
spływanie wydzieliny	p = 0.006502	p = 0.006757
kaszel	-	p = 0.001575
kichanie	-	-
ból twarzy	-	-
ból czoła	p = 0.022270	p = 0.008665
ból głowy	p = 0.046147	p = 0.006570
dolegliwości uszne	p = 0.015886	p = 0.047731
zaburzenia węchu	-	-
zmęczenie	p = 0.017046	p = 0.011098
nieprzyjemny zapach z ust	p = 0.028737	p = 0.012406
ból zębów	-	-

## **6.2.2 Modele neuronowe przewidujące nasilenie dolegliwości po zabiegu**

### **6.2.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa VAS\_bin\_6)**

#### **➤ ETAP I:**

#### **(a) Poszukiwanie wektora wejściowego**

Poszukując najbardziej obiecującego zestawu zmiennych pozwalających przewidywać zmienną VAS\_bin\_6 porównano różne wektory wejściowe, tj. różne zestawy zmiennych wejściowych. Dla każdego z nich wygenerowano przy pomocy *Automatycznego projektanta sieci* grupę 100 modeli neuronowych, zgodnie z założeniami szczegółowo opisanymi w rozdziale „Materiał i metoda”. Korzystano z podziału danych na podzbiory uczący, walidacyjny i testowy w proporcjach 60% : 20% : 20%, przeprowadzonego na podstawie opisaną wcześniej analizy skupień. Z powodu pojedynczych braków danych ostatecznie liczby przypadków w poszczególnych podziorach wynosiły 82 : 27 : 26. Zmienne wykorzystywane przy tworzeniu wektorów wejściowych (zarówno w tym podrozdziale, jak i w kolejnych) wymieniono w tabeli 11.

**Tabela 11. Zestawienie zmiennych wykorzystywanych do tworzenia wektorów wejściowych (zestawienie dotyczy również kolejnych podrozdziałów)**

<b>Cecha (zmienna)</b>	<b>skrótowa nazwa</b>
pleć	pleć
wiek	wiek
polipy	polipy
masywne polipy	masywne polipy
skrzywienie przegrody nosa	skrzywienie przegrody
TK (skala Lund-Mackay)	TK
radykałność zabiegu	zabieg
histopatologia – dominacja eozynofili	hist-pat
czas objawów	czas objawów
astma	astma
nadwrażliwość na niesteroidowe leki przeciwzapalne	AIA
alergia	alergia
refluks	refluks
przebyte polipektomie / liczba przebytych polipektomii	polipektomie / liczba polipektomii
przebyte ESS / liczba przebytych ESS	ESS / liczba ESS
<b>nasilenie dolegliwości przed zabiegiem:</b>	
ogólnie (w skali VAS)	VAS_0
<b>dolegliwości w skali 4-stopniowej:</b>	
niedrożność nosa	niedrożność
wydzielina z nosa	wydzielina
splywanie wydzieliny	splywanie
kaszel	kaszel
kichanie	kichanie
ból twarzy	ból twarzy
ból czoła	ból czoła
ból głowy	ból głowy
dolegliwości uszne (ból, uczucie pełności)	ból ucha
zaburzenia węchu	węch
zmęczenie	zmęczenie
nieprzyjemny zapach z ust (fetor ex ore)	fetor
ból zębów	ból zębów

Jakości i błędy grup modeli dla poszczególnych wektorów wejściowych zestawiono w tabeli 12 i zilustrowano na rysunku 25. W tabeli 13 przedstawiono najlepsze modele uzyskane na tym etapie badań. Przetestowano następujące wektory wejściowe:

- Wszystkie dostępne zmienne (29 zmiennych wymienionych w tabeli 11).  
**Wynik testu:** Błędy walidacji i testowania uzyskane dla tego wektora były o kilka rzędów wielkości większe niż błędy uczenia. Oznacza to bardzo słabą zdolność otrzymanych sieci neuronowych do generalizacji i tendencję do przeuczania modelu, wynikającą ze zbyt dużych rozmiarów sieci. Redukcja wektora wejściowego dokonana przez *Automatycznego projektanta sieci* okazała się niewystarczająca, w związku z czym podjęto kolejne próby zmniejszenia jego rozmiarów innymi metodami.
- Wektory opracowane na podstawie analizy statystycznej przedstawionej w poprzednim podrozdziale:
  - Wektor „statystyczny 1” – 12 zmiennych, dla których wykazano, że zmienna wyjściowa VAS\_bin\_6 jest od nich statystycznie istotnie zależna (bądź wynik jest bliski granicy istotności statystycznej) – por. tab. 10.  
**Wynik testu:** Redukcja wektora wejściowego zgodnie z oczekiwaniami spowodowała zwiększenie błędów uczenia oraz zmniejszenie błędów walidacji i testowania, co oznacza poprawę zdolności sieci do generalizacji. Ponieważ jednak nadal przeciętne błędy walidacji i testowania były o rząd wielkości większe niż błąd uczenia, zdecydowano o usunięciu kolejnych zmiennych wejściowych.
  - Wektor „statystyczny 2” uzyskano w wyniku redukcji wektora „statystycznego 1” poprzez usunięcie zmiennych wzajemnie skorelowanych. Najsilniejszą korelację stwierdzono między zmiennymi „ból głowy” i „ból czoła” (R Spearmana = 0.54), natomiast pozostałe korelacje były znacznie słabsze. Z wektora wejściowego usunięto zmienną „ból głowy”, od której słabiej zależna była zmienna wyjściowa, uzyskując wektor złożony z 11 zmiennych.  
**Wynik testu:** Dalsza redukcja rozmiaru wektora wejściowego spowodowała dalsze zmniejszenie błędów walidacji i testowania.
  - Wektor „statystyczny 3” – 10 zmiennych – Wektor ten uzyskano w wyniku redukcji wektora „statystycznego 1” poprzez usunięcie 2 zmiennych, od których zmienna VAS\_bin\_6 nie wykazywała statystycznie istotnej zależności, a wynik był jedynie bliski granicy istotności statystycznej.  
**Wynik testu:** Uzyskano w ten sposób dalszą poprawę zdolności modelu do generalizacji.
- Wektory uzyskane dzięki algorytmom doboru cech – najlepsze z wektorów opracowanych dzięki zastosowaniu algorytmów krokowych (metody krokowej postępującej i krokowej wstecznej) oraz algorytmów genetycznych według założeń opisanych w rozdziale „Materiał i metoda”:
  - Wektor „krokowa postępująca” – 7 zmiennych (skrzywienie przegrody, ból głowy, ból ucha, zmęczenie, refluks, ESS, liczba ESS), wytypowanych dzięki zastosowaniu metody krokowej postępującej.

**Wynik testu:** Mimo niewielkich rozmiarów wektora zdolność sieci do generalizacji nie okazała się wysoka.

- Wektor „krokowa wsteczna” – 5 zmiennych (VAS\_0, AIA, refluks, ESS, liczba ESS) wybranych z zastosowaniem metody krokowej wstecznej.

**Wynik testu:** Znaczna redukcja wielkości wektora pozwoliła na obniżenie błędów walidacji i testowania, czyli poprawę zdolności do generalizacji, jednak niskie jakości uczenia uzyskiwane w tym przypadku sugerują, że sieci neuronowe wykorzystujące ten wektor są zbyt ubogie aby dostatecznie dobrze modelować badaną zależność.

- Wektor „algorytm genetyczny” – 8 zmiennych (VAS\_0, niedrożność, ból ucha, zmęczenie, fetor, ból zębów, refluks, ESS) wytypowanych dzięki zastosowaniu algorytmów genetycznych.

**Wynik testu:** Użycie wektora nieco większego niż poprzedni zgodnie z obserwowaną wcześniej tendencją pozwoliło w pewnym stopniu obniżyć błędy uczenia przy ponownym pogorszeniu zdolności sieci do generalizacji.

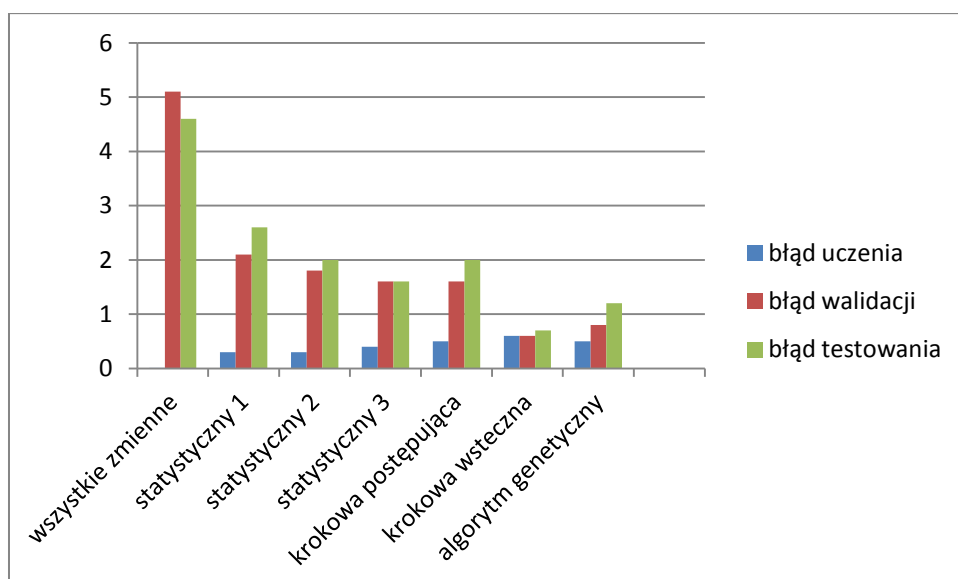
W trakcie opracowywania wektorów wejściowych przy użyciu wymienionych wyżej metod doboru cech stwierdzono ponadto, że w przypadku rozważanego zagadnienia nie dochodzi do jednoznacznego wyodrębnienia podzbioru zmiennych wejściowych, który w sposób zdecydowanie lepszy niż pozostałe pozwalałby przewidywać rozważaną zmienną wyjściową. Powtarzane próby przeprowadzenia doboru cech dawały wielokrotnie porównywalne wyniki (wartości błędu) dla wektorów złożonych z bardzo różnych zmiennych. Można założyć, że różne kombinacje zmiennych niosą w istocie podobną informację na temat rokowania u pacjenta, natomiast obserwowane różnice biorą się stąd, że przy dużych wymiarach wektora wejściowego pojawiają się trudności w dokładnym ustaleniu optymalnych wartości parametrów sieci na podstawie procesu uczenia.

**Tabela 12. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez Automatycznego projektanta sieci grup liczących po 100 modeli) uzyskanych dla różnych wektorów wejściowych – opis w tekście.**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)					
wszystkie dostępne zmienne (29)	84-100 (93)	41-67 (55)	42-73 (58)	0.00001-0.5 (0.01)	2.0-11.8 (5.1)	1.9-46 (4.6)
wektor „statystyczny 1” (12)	80-95 (87)	55-78 (63)	46-77 (58)	0.03-0.4 (0.3)	1.3-3.2 (2.1)	1.5-5.3 (2.6)

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)					
wektor „statystyczny 2” (11)	76-96 (88)	52-74 (59)	46-77 (65)	0.12-0.5 (0.3)	1.1-3.0 (1.8)	1.2-3.9 (2.0)
wektor „statystyczny 3” (10)	72-89 (79)	48-70 (59)	42-73 (62)	0.3-0.5 (0.4)	1.2-2.3 (1.6)	1.1-2.3 (1.6)
wektor „krokowa postępująca” (7)	71-85 (79)	56-78 (63)	58-77 (65)	0.4-0.6 (0.5)	1.2-2.1 (1.6)	1.4-2.8 (2.0)
wektor „krokowa wsteczna” (5)	70-78 (74)	56-78 (67)	69-85 (77)	0.6-0.7 (0.6)	0.5-0.8 (0.6)	0.6-0.9 (0.7)
wektor „algorytm genetyczny” (8)	71-84 (78)	48-70 (59)	50-77 (62)	0.4-0.6 (0.5)	0.6-1.1 (0.8)	0.7-1.8 (1.2)

Rysunek 25. Błędy uczenia, walidacji, testowania dla poszczególnych wektorów wejściowych - zmiany wartości mediana obserwowane po kolejnych modyfikacjach wektora wejściowego



**Tabela 13. Najlepsze modele neuronowe uzyskane dla poszczególnych wektorów wejściowych.**

Architektura sieci zapisywana jest poprzez podanie kolejno liczby neuronów wejściowych, ukrytych i wyjściowych (np. zapis „MLP 22-6-1” oznacza perceptron wielowarstwowy typu MLP o 22 neuronach wejściowych, 6 neuronach ukrytych i 1 neuronie wyjściowym). Następnie podane są jakości uczenia, walidacji i testowania w procentach oraz (w nawiasie) uśrednione błędy popełniane przez egzaminowaną sieć, ustalone odpowiednio na zbiorach danych używanych do uczenia, walidacji i testowania.

Wektor wejściowy	Architektura sieci	Jakości: uczenia-walidacji-testowania	Błędy: uczenia-walidacji-testowania
wszystkie dostępne zmienne	MLP 22-6-1	93%-63%-69%	0.2-3.3-3.1
wektor „statystyczny 1”	MLP 12-7-1	90%-78%-69%	0.3-1.6-1.8
wektor „statystyczny 2”	MLP 11-11-1	93%-74%-65%	0.3-1.7-2.1
wektor „statystyczny 3”	MLP 10-9-1	85%-67%-65%	0.4-1.7-1.6
wektor „krokowa postępująca”	MLP 7-11-1	82%-74%-69%	0.5-1.5-2.0
wektor „krokowa wsteczna”	MLP 5-9-1	76%-78%-81%	0.6-0.6-0.7
wektor „algorytm genetyczny”	MLP 8-9-1	77%-70%-69%	0.5-0.8-1.3

Na podstawie literatury [186, 187, 189] ustalono, że zaobserwowane trudności w uzyskaniu modeli o zadowalającej jakości uczenia, posiadających równocześnie akceptowalną zdolność do generalizacji, mogą wynikać z dwóch przyczyn: albo ze zbyt małej ilości danych uczących (brak możliwości właściwego przewidywania wyników w zbiorze walidacyjnym i testowym na podstawie informacji zawartej w zbiorze uczącym), albo z niewłaściwego podziału na podzbiory uczący, walidacyjny i testowy (zdecydowanie „trudniejsze” do zaklasyfikowania przypadki w zbiorze walidacyjnym i testowym niż w zbiorze uczącym). W celu sprawdzenia, która z tych dwóch przyczyn może stanowić źródło trudności w rozważanym tu przypadku, przeprowadzono dalsze analizy. Zbadano mianowicie zależności uzyskiwanych wyników od liczby przypadków uczących oraz od sposobu wyodrębnienia podzbiorów: uczącego, walidacyjnego i testowego.

### **(b) Modyfikacja liczby przypadków uczących.**

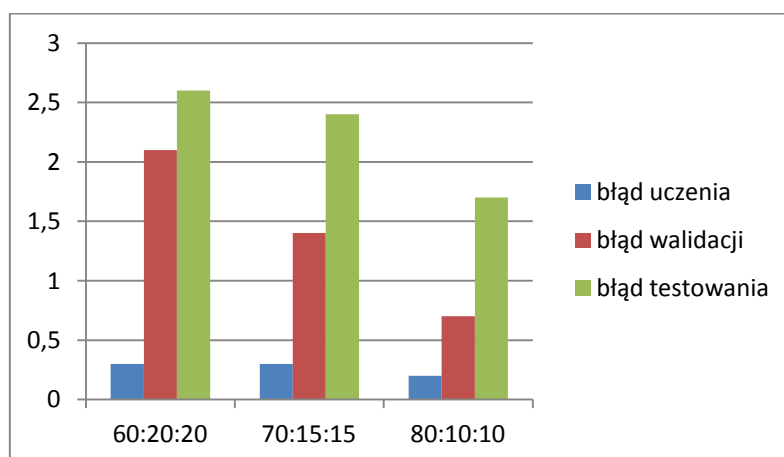
Analizę wpływu liczby przypadków uczących na efektywność modeli neuronowych przeprowadzono na przykładzie wektora „statystyczny 1”. W tym celu porównano wyniki uzyskane dla tego wektora przy zastosowaniu podziału zasobu posiadanych danych na podzbiory uczący, walidacyjny i testowy w proporcjach 60% : 20% : 20% (82:27:26 przypadków) z wynikami uzyskanymi dla podziałów w proporcjach 70% : 15% : 15% (96:20:19 przypadków) oraz 80% : 10% : 10% (106:15:14 przypadków). Dla każdego z wymienionych trzech podziałów wygenerowano przy pomocy *Automatycznego projektanta sieci* grupę 100 modeli zgodnie z zasadami opisanymi w rozdziale „Materiał i metoda”. Rezultaty porównań otrzymanych grup sieci zestawiono w tabeli 14 oraz zilustrowano na rysunku 26. Zmiany proporcji w podzbiórach uzyskano przenosząc w ramach poszczególnych

skupień odpowiednią liczbę przypadków ze zbiorów walidacyjnego i testowego do zbioru uczącego.

**Tabela 14. Wpływ zmian liczby przypadków uczących na efektywność modeli neuronowych (na przykładzie wektora wejściowego „statystyczny 1”).** Dla oszczędności miejsca tutaj (i w dalszej części pracy) przedstawiono łącznie dane prezentowane uprzednio w formie dwóch oddzielnych tabel (tabele 12 i 13). W ostatniej kolumnie zapisano architekturę najlepszych uzyskanych sieci wg schematu jak w tabeli 13, a następnie podano jakości uczenia, walidacji i testowania w procentach oraz (w nawiasie) uśrednione błędy popełniane przez egzaminowaną sieć, ustalane odpowiednio na zbiorach danych używanych do uczenia, walidacji i testowania.

Proporcje podzbiorów uczącego, walidacyjnego i testowego	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
<b>60 : 20 : 20</b>	80-95 (87)	55-78 (63)	46-77 (58)	0.03-0.4 (0.3)	1.3-3.2 (2.1)	1.5-5.3 (2.6)	MLP 12-7-1 90%-78%-69% (0.3-1.6-1.8)
<b>70 : 15 : 15</b>	81-97 (92)	50-85 (65)	37-79 (53)	0.06-0.4 (0.3)	0.8-2.7 (1.4)	1.5-4.3 (2.4)	MLP 12-11-1 86%-80%-74% (0.2-1.7-2.3)
<b>80 : 10 : 10</b>	86-100 (96)	60-100 (80)	50-79 (71)	0.08-0.4 (0.2)	0.3-1.5 (0.7)	0.7-3.1 (1.7)	MLP 12-10-1 97%-93%-78% (0.3-0.3-1.8)

**Rysunek 26. Błędy uczenia, walidacji, testowania dla różnych proporcji podzbiorów uczącego, walidacyjnego i testowego - zmiany wartości mediany**





Jak widać, zwiększanie rozmiarów zbioru uczącego kosztem walidacyjnego i testowego prowadzi do postępującej poprawy jakości klasyfikacji i do zmniejszenia błędów sieci. Zdecydowano w związku z tym w dalszej części pracy posługiwać się zbiorami uczącym, walidacyjnym i testowym w proporcjach 70% : 15% : 15%. Dla proporcji 80% : 10% : 10% przypuszczalnie możliwe byłoby uzyskanie lepszych wyników, jednak liczba przypadków w zbiorach walidacyjnym i testowym (zwłaszcza w mniej licznej II grupie pacjentów) byłaby wtedy nieakceptowalnie mała. W związku z tym ocena rzeczywistej zdolności modeli do generalizacji mogłaby okazać się mało precyzyjna.

### **(c) Ocena przyjętego podziału na zbiory uczący, walidacyjny i testowy**

W opisanych powyżej analizach zwraca uwagę fakt, że nawet dla dużej liczby przypadków uczących błąd testowania nadal znacząco przewyższa błędy uczenia i walidacji. Jak wspomniano powyżej, tego rodzaju efekt może być wynikiem nieoptymalnego wyboru zbioru testowego. Mimo starannego przeprowadzenia procedury podziału na podzbiory na podstawie analizy skupień, w zbiorze testowym miały szanse znaleźć się przypadki szczególnie „trudne” do zaklasyfikowania (przynajmniej dla rozważanej aktualnie zmiennej wyjściowej).

W celu weryfikacji tej hipotezy porównano modele uzyskane dla wektora „statystyczny 1” przy podziale na podzbiory uczący, walidacyjny i testowy (70% : 20% : 20%) przeprowadzonym według analizy skupień z analogicznymi modelami, dla których podzbiory uczący, walidacyjny i testowy w tych samych proporcjach wyodrębniono losowo (oddzielnie dla każdej sieci). Rezultaty przedstawiono w tabeli 15.

**Tabela 15. Porównanie wyników uzyskiwanych przez sieci neuronowe przy podziale na podzbiory uczący, walidacyjny i testowy wyodrębnione na podstawie analizy skupień oraz losowo (oddzielnie dla każdej sieci).**

Sposób podziału na podzbiory uczący, walidacyjny i testowy	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
losowo	82-100 (94)	35-90 (65)	40-90 (65)	0.04-0.4 (0.2)	0.5-4.2 (1.8)	0.4-3.9 (1.7)	MLP 12-11-1 94%-90%-75% (0.3-0.5-0.7)
wg analizy skupień	81-97 (92)	50-85 (65)	37-79 (53)	0.06-0.4 (0.3)	0.8-2.7 (1.4)	1.5-4.3 (2.4)	MLP 12-11-1 86%-80%-74% (0.2-1.7-2.3)

Przeprowadzone porównanie pozwala zauważyć, że w przypadku losowego podziału na podzbiory uczący, walidacyjny i testowy, przeciętne jakości uczenia i walidacji są porównywalne jak dla podziału przeprowadzonego na podstawie analizy skupień, podczas gdy wyniki testowania są w tym drugim przypadku zauważalnie gorsze. Oznacza to, że dla zmiennej VAS\_bin\_6 rzeczywista zdolność modeli do generalizacji może być wyższa niż wskazuje na to jakość testowania i wartość błędu testowania, ponieważ „trudność” zbioru testowego dla tej zmiennej wydaje się zawyżona.

Analiza wyników zamieszczonych w tabeli 15 pozornie sugeruje, że zastosowanie losowego podziału na podzbiór uczący, walidacyjny i testowy może prowadzić do uzyskania skuteczniejszych modeli. Należy jednak pamiętać, że uzyskane w ten sposób wyniki nie są wiarygodne, ponieważ jako pozornie najlepsza traktowana jest taka sieć, dla której wylosowano „najłatwiejszy” zbiór testowy. Ocenionej w ten sposób jakości testowania nie można uważać za reprezentatywną dla ewentualnych nowych przypadków, które miałyby być klasyfikowane przez sieć w trakcie jej dalszego praktycznego użytkowania. Z tego powodu właściwszym podejściem jest jednak wykorzystanie zbioru testowego wyodrębnionego na podstawie logicznych kryteriów (np. analizy skupień). W przypadku uzyskania wyników sugerujących, że zbiór testowy może zawierać przypadki odbiegające trudnością od przeciętnych, konieczna jest ocena zbioru w odniesieniu do konkretnych zmiennych wyjściowych, a następnie odpowiednia do uzyskanych wyników interpretacja otrzymanych rezultatów uczenia sieci. W szczególności w przypadku zmiennej VAS\_bin\_6 można uznać, że nieco wyższy błąd testowania w stosunku do błędu uczenia i walidacji nie musi świadczyć o słabej zdolności do generalizacji, a jedynie wynika z właściwości zbioru testowego.

#### **(d) Ocena liczby neuronów ukrytych**

Kierując się wynikami przytoczonych powyżej analiz zdecydowano w dalszych poszukiwaniach wykorzystać wektory „statystyczny 3” i „algorytm genetyczny” i posługiwać się dokonaniem według analizy skupień podziałem na podzbiory uczący, walidacyjny i testowy w proporcjach 70% : 15% : 15% (96 : 20 : 19 przypadków).

#### **Wektor wejściowy „statystyczny 3”**

Wyniki uzyskane dla wektora „statystyczny 3” zestawiono w tabeli 16 i zilustrowano na rysunku 27. Najlepszy model uzyskany dotychczas dla tego wektora liczył 9 neuronów ukrytych, jednak wykazywał tendencję do przeuczenia się, co sugerowało nadmierną wielkość sieci. Ponadto optymalne wartości parametru  $\rho$  (którego znaczenie omówiono w rozdziałach „Wstęp” oraz „Materiał i metoda”) dla wektora wejściowego liczącego 10 zmiennych wejściowych uzyskuje się dla około 4 neuronów ukrytych. Z tego powodu przetestowano w praktyce i porównano sieci o liczbie neuronów ukrytych od 2 do 8.

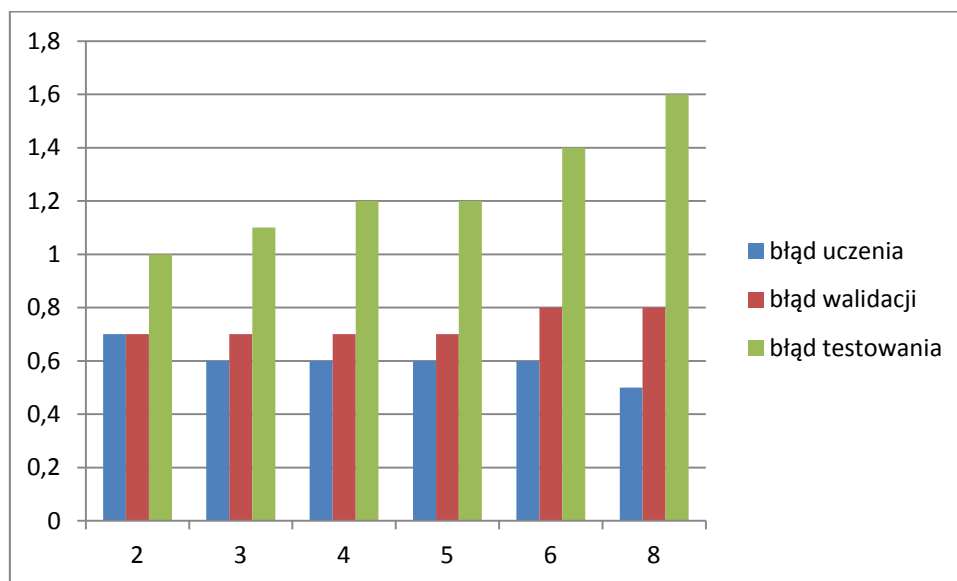
Analiza wyników przedstawionych w tabeli 16 obrazuje, jak ze wzrostem liczby neuronów ukrytych poprawia się jakość uczenia sieci (rośnie zdolność sieci do dopasowania się do zbioru uczącego) przy jednoczesnym wzroście błędów walidacji i testowania (utrata zdolności

do generalizacji). Najlepszym z uzyskanych modeli okazała się sieć neuronowa posiadająca 4 neurony w warstwie ukrytej. Jakości uczenia, walidacji i testowania dla tej sieci wynosiły odpowiednio: 79%-85%-79%, błędy 0.6-0.6-0.9, wartość parametru  $\rho$  1.8 (tj. w zakresie wartości optymalnych). Analizę wrażliwości przedstawiono w tabeli 17. Jej wyniki sugerują, że zmienna „ból głowy” może być zbędna dla działania modelu (problem ten brano pod uwagę w II etapie eksperymentów dotyczących omawianej sieci neuronowej).

**Tabela 16. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez Automatycznego projektanta sieci grup liczących po 100 modeli) przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 3”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	58-79 (71)	45-85 (70)	42-89 (70)	0.5-0.8 (0.7)	0.4-1.4 (0.7)	0.5-1.4 (1.0)	MLP 9-2-1 76%-75%-73% (0.7-0.7-1.0)
3	53-83 (75)	50-90 (70)	53-84 (68)	0.5-0.8 (0.6)	0.4-1.2 (0.7)	0.7-2.0 (1.1)	MLP 10-3-1 79%-70%-79% (0.6-0.8-0.9)
4	39-83 (77)	40-90 (70)	32-84 (68)	0.5-0.8 (0.6)	0.4-1.5 (0.7)	0.6-1.8 (1.2)	<b>MLP 9-4-1 79%-85%-79%</b> <b>(0.6-0.6-0.9)</b>
5	69-86 (77)	50-85 (70)	58-79 (68)	0.5-0.7 (0.6)	0.4-1.4 (0.7)	0.7-2.0 (1.2)	MLP 9-5-1 80%-85%-74% (0.5-0.4-1.1)
6	71-89 (79)	50-85 (70)	58-79 (68)	0.5-0.8 (0.6)	0.4-1.6 (0.8)	0.8-2.6 (1.4)	MLP 9-6-1 82%-80%-74% (0.6-0.4-1.1)
8	72-92 (81)	50-85 (70)	53-84 (68)	0.4-0.6 (0.5)	0.4-1.5 (0.8)	0.6-2.8 (1.6)	MLP 9-8-1 83%-85%-74% (0.5-0.4-1.8)

**Rysunek 27. Błędy uczenia, walidacji, testowania dla wektora „statystyczny 3” – zmiany wartości mediany przy zmieniającej się liczbie neuronów ukrytych**



**Tabela 17. Analiza wrażliwości dla najlepszego z modeli wykorzystujących wektor wejściowy „statystyczny 3” (MLP 9-4-1). Pojęcie ilorazu i rangi objaśniono w rozdziale „Materiał i metoda”.**

Zmienna	splywanie	ból czoła	ból głowy	ból ucha	zmęczenie	fetor	AIA	refluks	ESS
Iloraz	1.0317	1.0351	0.9976	1.0211	1.0216	1.0004	1.2375	1.0600	1.0849
Ranga	5	4	9	7	6	8	1	3	2

### **Wektor wejściowy „algorytm genetyczny”**

Najlepszy model dla wektora wejściowego „algorytm genetyczny”, uzyskany na etapie poszukiwań wektora wejściowego, posiadał 9 neuronów ukrytych. Błędy uczenia, walidacji i testowania dla tego modelu wynosiły odpowiednio: 0.5, 0.8 i 1.3, co sugeruje, że zmniejszenie sieci mogłoby prowadzić do poprawy zdolności do generalizacji. Optymalne wartości parametru  $\rho$  dla sieci o 8 neuronach wejściowych uzyskuje się dla 4 lub 5 neuronów ukrytych. Z tego względu zdecydowano przetestować modele posiadające od 2 do 10 neuronów w warstwie ukrytej.

Podobnie jak w przypadku omawianego wcześniej wektora „statystyczny 3”, również dla wektora „algorytm genetyczny” wraz ze wzrostem liczby neuronów ukrytych obserwowano stopniowe zmniejszanie błędów uczenia i wzrost błędów walidacji i testowania (tab. 18, rys. 28).

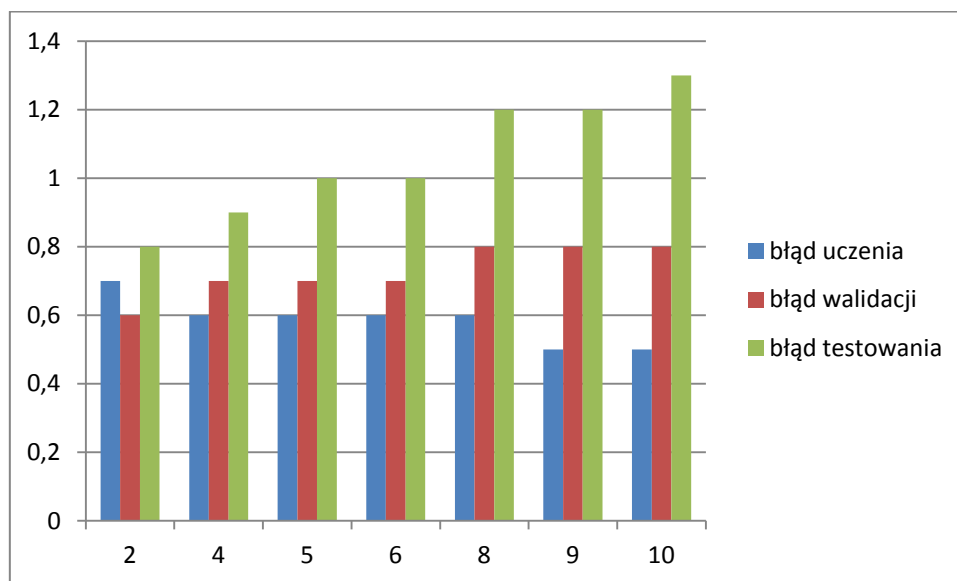
W toku eksperymentów uzyskano liczne modele o jakościach testowania 74-79%. Najwyższą jakość testowania zanotowano dla sieci o 9 neuronach ukrytych. Mimo dużych rozmiarów ( $\rho = 1.2$ ) wartości błędów walidacji i testowania dla tej sieci były stosunkowo niskie (odpowiednio 0.6 i 0.8 przy błędzie uczenia 0.5), co sugeruje akceptowalną zdolność sieci do generalizacji. Analiza wrażliwości dla tej sieci nie wykazała obecności zbędnych zmiennych wejściowych (tab. 19).

Nieznacznie gorsze wyniki przy optymalnych wartościach parametru  $\rho$  uzyskano dla sieci posiadających 5 lub 6 neuronów ukrytych. Analiza wrażliwości dla tych sieci wykazała jednak obecność pojedynczych potencjalnie zbędnych zmiennych (tab. 20 i 21).

**Tabela 18. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez Automatycznego projektanta sieci grup liczących po 100 modeli) przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „algorytm genetyczny”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	38-77 (67)	20-85 (70)	26-84 (58)	0.6-0.7 (0.7)	0.5-0.9 (0.6)	0.6-1.4 (0.8)	MLP 5-2-1 75%-75%-74% (0.6-0.6-0.8)
4	42-82 (73)	40-85 (70)	32-79 (58)	0.5-0.7 (0.6)	0.5-0.9 (0.7)	0.6-1.5 (0.9)	MLP 6-4-1 79%-80%-74% (0.6-0.6-0.9)
5	65-84 (75)	55-90 (70)	42-84 (58)	0.5-0.7 (0.6)	0.5-1.1 (0.7)	0.6-1.8 (1.0)	<b>MLP 7-5-1 77%-85%-74% (0.5-0.7-1.0)</b>
6	42-84 (76)	40-90 (70)	37-84 (61)	0.5-0.7 (0.6)	0.5-1.2 (0.7)	0.6-1.9 (1.0)	<b>MLP 7-6-1 76%-80%-74% (0.6-0.6-0.7)</b>
8	63-89 (78)	60-90 (75)	37-74 (63)	0.4-0.7 (0.6)	0.5-1.1 (0.8)	0.6-1.7 (1.2)	MLP 6-8-1 78%-90%-74% (0.6-0.6-1.0)
9	69-89 (79)	60-90 (75)	42-79 (63)	0.4-0.8 (0.5)	0.5-1.3 (0.8)	0.7-1.9 (1.2)	<b>MLP 7-9-1 77%-85%-79% (0.5-0.6-0.8)</b>
10	71-89 (80)	60-90 (70)	37-79 (63)	0.4-0.6 (0.5)	0.5-1.1 (0.8)	0.7-2.1 (1.3)	MLP 7-10-1 78%-80%-74% (0.5-0.7-0.9)

**Rysunek 28. Błędy uczenia, walidacji, testowania dla wektora „algorytm genetyczny” – zmiany wartości mediany przy zmieniającej się liczbie neuronów ukrytych**



**Tabela 19. Analiza wrażliwości dla najlepszego modelu wykorzystującego wektor wejściowy „algorytm genetyczny” i posiadającego 9 neuronów ukrytych**

Zmienna	niedrożność	ból ucha	zmęczenie	fetor	ból zębów	refluks	liczba ESS
Iloraz	1.0336	1.2209	1.0495	1.0050	1.0562	1.0152	1.0781
Ranga	5	1	4	7	3	6	2

**Tabela 20. Analiza wrażliwości dla najlepszego modelu wykorzystującego wektor wejściowy „algorytm genetyczny” i posiadającego 5 neuronów ukrytych**

Zmienna	niedrożność	ból ucha	zmęczenie	fetor	ból zębów	refluks	liczba ESS
Iloraz	0.9908	1.3981	1.0407	1.0000	1.1156	1.0105	1.0651
Ranga	7	1	4	6	2	5	3

**Tabela 21. Analiza wrażliwości dla najlepszego modelu wykorzystującego wektor wejściowy „algorytm genetyczny” i posiadającego 6 neuronów ukrytych**

Zmienna	niedrożność	ból ucha	zmęczenie	fetor	ból zębów	refluks	liczba ESS
Iloraz	0.9889	1.1683	1.0675	0.9883	1.0253	1.0527	1.1071
Ranga	6	1	3	7	5	4	2

## ➤ ETAP II

W drugim etapie eksperymentów, tj. przy poszukiwaniu optymalnego modelu, za punkt wyjścia przyjęto architekturę najlepszych sieci neuronowych uzyskanych w etapie I. Ze względu na mniej uporządkowany charakter poszukiwań optymalnego modelu w tym etapie oraz ograniczenia związane z objętością pracy, poniżej przedstawiono skrótowo jedynie najistotniejsze wyniki przeprowadzonych testów.

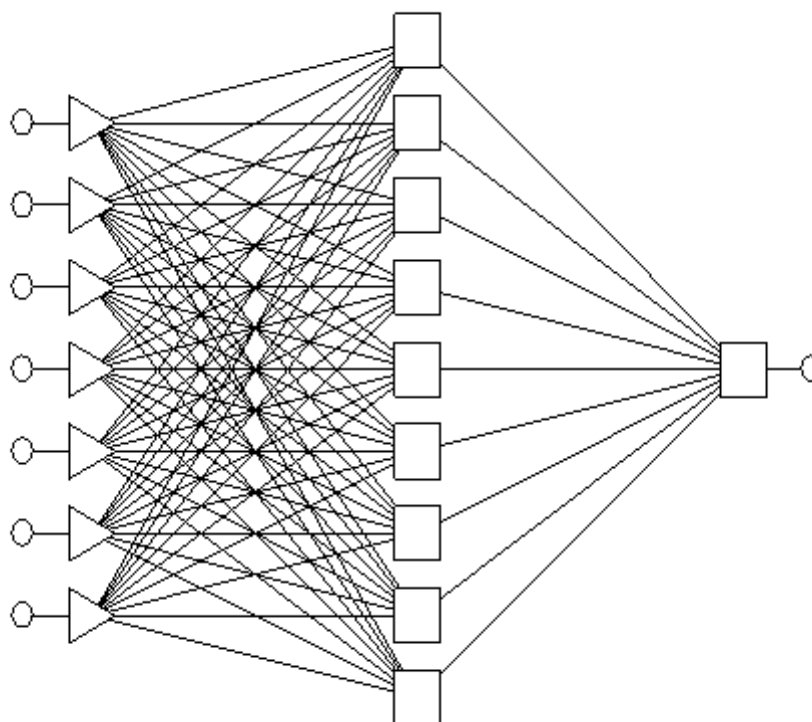
- Wektor „statystyczny 3”, 4 neurony ukryte.  
Najlepszy model uzyskany w etapie I (MLP 9-4-1) charakteryzował się jakościami uczenia, walidacji i testowania odpowiednio 79%-85%-79% i błędami odpowiednio 0.6-0.6-0.9.  
Przy pomocy *Automatycznego projektanta sieci* przetestowano 1000 modeli o takiej samej architekturze według zasad opisanych w rozdziale „Materiał i metoda”. Uzyskany w efekcie najlepszy model to MLP 9-4-1 o jakościach 76%-90%-79% i błędach 0.6-0.4-0.8. Wartość parametru  $\rho$  dla tego modelu wynosiła 2.1 (tj. w przedziale optymalnym), a w przeciwieństwie do wyjściowego modelu analiza wrażliwości nie wykazała obecności zbędnych zmiennych wejściowych.
- Wektor „algorytm genetyczny”, 5-6 neuronów ukrytych.  
Najlepszymi dotychczas uzyskanymi modelami były sieci MLP 7-5-1 (jakości 77%-85%-74%, błędy 0.5-0.7-1.0) i MLP 7-6-1 (jakości 76%-80%-74%, błędy 0.6-0.6-0.7). Przetestowanie przy pomocy *Automatycznego projektanta* 1000 sieci o analogicznych architekturach nie doprowadziło do uzyskania znacząco lepszego modelu. Najlepsza otrzymana sieć to MLP 7-6-1 (jakości 81%-75%-76%, błędy 0.5-0.5-1.1)
- Wektor „algorytm genetyczny”, 9 neuronów ukrytych.  
Przetestowanie 1000 sieci doprowadziło do uzyskania najlepszego modelu o dokładnie takich samych jakościach i błędach, jak w przypadku modelu wyjściowego (jakości 77%-85%-79%, błędy 0.5-0.6-0.8). Ponieważ wartość parametru  $\rho$  dla tego modelu wynosiła 1.2, co sugerowało nadmierną wielkość sieci, a analiza wrażliwości wykazywała obecność jednej zbędnej zmiennej wejściowej (fektor), przetestowano 1000 modeli zbudowanych z wyłączeniem tej zmiennej, jednak żaden z otrzymanych modeli nie okazał się lepszy od uzyskanych wcześniej.

Tabela 22 zawiera podsumowujące zestawienie najefektywniejszych spośród uzyskanych sieci neuronowych. Przedstawiono w niej również, w jakim stopniu spełniają one przyjęte kryteria wyboru najlepszych modeli (opisane w rozdziale „Materiał i metoda”, w podrozdziale „Ocena efektywności sieci neuronowych”). Rysunki 29 i 30 przedstawiają schematycznie widok obu najlepszych sieci neuronowych.

**Tabela 22. Najlepsze uzyskane modele dokonujące predykcji nasilenia dolegliwości po zabiegu w I grupie pacjentów.**

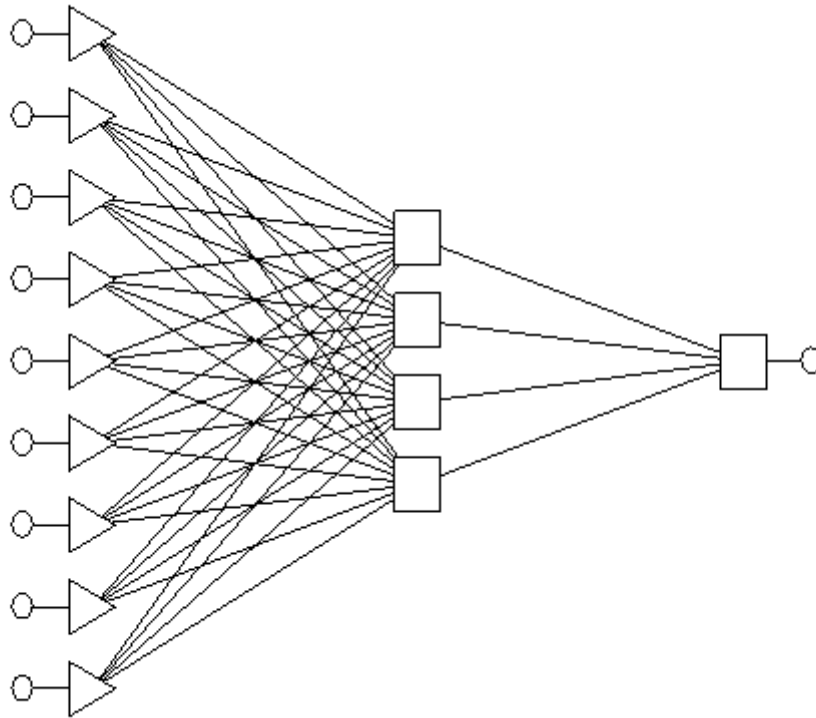
Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 7-9-1	niedrożność, ból ucha, zmęczenie, fetor, ból zębów, refluks, liczba ESS	77%-85%-79%	0.5-0.6-0.8	nie	nie
MLP 9-4-1	spływanie, ból czoła, ból głowy, ból ucha, zmęczenie, fetor, AIA, refluks, ESS	76%-90%-79%	0.6-0.4-0.8	nie	tak

**Rysunek 29. Widok sieci neuronowej o architekturze MLP 7-9-1.**





**Rysunek 30. Widok sieci neuronowej o architekturze MLP 9-4-1.**



### 6.2.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa VAS\_bin\_12)

#### ➤ ETAP I

##### (a) Poszukiwanie wektora wejściowego

Poszukiwania najbardziej obiecującego wektora wejściowego dla grupy II przeprowadzono w sposób analogiczny jak w grupie I. Korzystano z podziału zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 70% : 15% : 15%. Po uwzględnieniu braków danych liczba przypadków w poszczególnych podziorach wynosiła 48 : 11 : 11. Jakości i błędy dla grup po 100 sieci neuronowych wykorzystujących różne wektory wejściowe zestawiono w tabeli 23. Testowano następujące wektory wejściowe:

- Wszystkie dostępne zmienne.  
**Wynik testu:** Podobnie jak dla grupy I, stwierdzono znaczącą przewagę błędów walidacji i testowania nad błędami uczenia, co oznacza niską zdolność do generalizacji wynikającą ze zbyt dużych rozmiarów sieci. Modele dopasowują się nadmiernie do danych uczących.
- Wektory opracowane na podstawie analizy statystycznej.

- Wektor „statystyczny 1” – 11 zmiennych, od których zmienna wyjściowa VAS\_bin\_12 była statystycznie istotnie zależna.  
**Wynik testu:** Stwierdzono mniejszą niż poprzednio, ale zauważalną tendencję do przeuczania sieci.
- Wektor „statystyczny 2” – 9 zmiennych, od których statystycznie istotnie zależna była zarówno zmienna VAS\_bin\_12, jak i zmienna VAS\_12.  
**Wynik testu:** Usunięcie 2 zmiennych z wektora wejściowego nie spowodowało znaczącej poprawy działania sieci.
- Wektor „statystyczny 3” – zestaw 7 zmiennych (płeć, VAS\_0, wydzielina, ból czoła, ból głowy, fetor, liczba ESS) otrzymany w wyniku usunięcia z wektora „statystycznego 1” zmiennych wzajemnie skorelowanych.  
**Wynik testu:** Redukcja wektora wejściowego spowodowała wprawdzie zmniejszenie błędów walidacji, ale doprowadziła równocześnie do pogorszenia jakości walidacji i testowania.
- Wektory uzyskane dzięki algorytmom doboru cech.
  - Wektor „krokowa postępująca” – 7 zmiennych (kaszel, ból czoła, ból głowy, zmęczenie, fetor, polipektomie, ESS).  
**Wynik testu:** Mimo wykorzystania takiej samej liczby zmiennych jak w wektorze „statystyczny 3”, zastosowanie tego wektora pozwoliło uzyskać wyższe jakości walidacji i testowania.
  - Wektor „krokowa wsteczna” – 6 zmiennych (VAS\_0, wydzielina, sptywanie, kaszel, kichanie, ból czoła).  
**Wynik testu:** W porównaniu z poprzednim wektorem zaobserwowano niewielkie pogorszenie walidacji i poprawę testowania.
  - Wektor „algorytm genetyczny” – 7 zmiennych (TK, VAS\_0, kaszel, ból czoła, ból głowy, astma, refluks).  
**Wynik testu:** Utworzone sieci wykazywały podobne błędy jak dla poprzedniego wektora, jednak niższe jakości walidacji i testowania.

W wyniku omówionych powyżej eksperymentów nie wyodrębniono wektora wejściowego, który można by było jednoznacznie uznać za bardziej obiecujący od pozostałych. Dla żadnego z testowanych wektorów nie uzyskano modeli spełniających jednocześnie wszystkie przyjęte kryteria świadczące o zadowalającej efektywności budowanych sieci neuronowych. Zdecydowano, że w dalszych eksperymentach zostaną wykorzystane wektory: „statystyczny 1” (jako potencjalny punkt wyjścia do budowania mniejszych modeli, mogących uzyskać lepszą zdolność generalizacji) oraz „krokowa postępująca” (ze względu na uzyskanie wysokich jakości walidacji i testowania).

**Tabela 23. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez *Automatycznego projektanta sieci grup liczących po 100 modeli*) uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wszystkie dostępne zmienne (29)	88-100 (100)	45-91 (72)	55-91 (82)	<0.0000 01 – 1.4 (0.004)	<0.0000 01 – 6.8 (1.8)	0.5-21.9 (2.7)	MLP 18-4-1 100%-91%-81% (0.3-1.2-2.0)
wektor „statystyczny 1” (11)	83-96 (88)	64-82 (73)	64-91 (82)	0.0005- 0.4 (0.1)	1.6-7.0 (2.4)	1.1-5.4 (2.3)	MLP 11-6-1 88%-82%-82% (0.3-1.6-1.3)
wektor „statystyczny 2” (9)	77-96 (88)	55-82 (73)	73-91 (82)	0.0002- 0.4 (0.2)	0.8-5.4 (2.1)	1.2-8.9 (2.5)	MLP 9-4-1 81%-82%-82% (0.4-1.6-1.2)
wektor „statystyczny 3” (7)	79-94 (88)	55-91 (64)	55-73 (64)	0.3-0.5 (0.4)	0.4-1.2 (0.6)	1.3-2.6 (2.0)	MLP 7-6-1 83%-73%-73% (0.4-0.6-1.7)
wektor „krokowa postępująca” (7)	79-96 (90)	64-91 (73)	64-82 (73)	0.0009- 0.5 (0.3)	0.7-4.2 (1.6)	1.1-6.6 (2.3)	MLP 7-4-1 88%-91%-82% (0.3-1.3-1.5)
wektor „krokowa wsteczna” (6)	83-100 (90)	64-73 (64)	73-82 (82)	0.2-0.5 (0.3)	1.2-2.2 (1.8)	1.0-2.9 (2.0)	MLP 6-6-1 92%-73%-82% (0.3-1.2-2.2)
wektor „algorytm genetyczny” (7)	75-98 (86)	27-81 (55)	27-91 (73)	0.02-0.6 (0.3)	0.9-4.6 (1.6)	1.6-7.0 (2.6)	MLP 7-5-1 85%-73%-82% (0.3-1.6-1.6)

## **(b) ocena liczby neuronów ukrytych**

### **Wektor wejściowy „statystyczny 1”**

Najlepszym dotychczas uzyskanym modelem wykorzystującym ten wektor wejściowy była sieć MLP 11-6-1 o jakościach uczenia, walidacji i testowania odpowiednio 88%-82%-82% i błędach 0.3-1.6-1.3. Zarówno wartości błędów, jak i wartość parametru  $\rho$  ( $\rho = 0.6$ ) wskazują, że liczba połączeń w sieci jest zbyt duża. Z tego względu przetestowano sieci o mniejszej liczbie neuronów ukrytych z możliwością redukcji wektora wejściowego przez *Automatycznego projektanta sieci*. Wyniki eksperymentów zestawiono w tabeli 24. Wykazano wyraźne pogorszenie zdolności do generalizacji ze wzrostem liczby neuronów ukrytych. Najlepszym z uzyskanych modeli okazała się sieć o 2 neuronach ukrytych (MLP 11-2-1, jakości uczenia, walidacji i testowania: 85%-82%-82%, błędy 0.5-1.6-1.2,  $\rho = 1.8$ ). Analiza wrażliwości dla tej sieci, przedstawiona w tabeli 25, sugeruje jednak obecność w wektorze wejściowym potencjalnie zbędnych zmiennych (w związku z tym dalsze eksperymenty dotyczące tego modelu przeprowadzono w etapie II i opisano poniżej).

**Tabela 24. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez *Automatycznego projektanta sieci* grup liczących po 100 modeli) przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 1”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	79-94 (86)	64-82 (73)	64-82 (82)	0.2-1.4 (0.4)	0.6-3.6 (1.8)	0.9-3.8 (1.5)	<b>MLP 11-2-1</b> <b>85%-82%-82%</b> <b>(0.5-1.6-1.2)</b>
4	81-96 (86)	63-82 (73)	63-90 (92)	0.001- 0.5 (0.2)	1.2-7.8 (2.3)	0.9-5.5 (2.1)	MLP 9-4-1 85%-82%-82% (0.5-1.6-1.3)
6	79-94 (88)	64-82 (73)	63-91 (82)	0.0004- 0.4 (0.1)	1.1-6.6 (2.4)	1.2-6.2 (2.3)	MLP 11-6-1 94%-82%-82% (0.02-2.5-3.2)

**Tabela 25. Analiza wrażliwości dla najlepszego modelu wykorzystującego wektor wejściowy „statystyczny 1” i posiadającego 2 neurony ukryte.**

Zmienna	pleć	VAS_0	wydzie- lina	splywa- nie	kaszel	ból czoła	ból głowy	ból ucha	zmę- czenie	fetor	liczba ESS
Iloraz	1.0023	1.0263	1.0841	1.0466	0.9110	0.9865	0.9981	0.9976	0.9961	1.0591	0.9799
Ranga	5	4	1	3	11	9	6	7	8	2	10

### Wektor wejściowy „krokowa postępująca”

Jak zilustrowano w tabeli 26, najlepszy model uzyskano dla sieci posiadającej 7 neuronów wejściowych i 4 neurony w warstwie ukrytej (jakości uczenia, walidacji i testowania: 90%-82%-82%, błędy 0.4-1.2-1.4,  $\rho = 1.3$ ). Również dla tego modelu analiza wrażliwości (tab. 27) sugeruje obecność zbędnych zmiennych wejściowych (dalszą redukcję ich liczby przeprowadzono w etapie II).

**Tabela 26. Porównanie jakości i błędów sieci neuronowych (wygenerowanych przez *Automatycznego projektanta sieci grup* liczących po 100 modeli) przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „krokowa postępująca”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	71-94 (80)	55-82 (73)	45-82 (73)	0.3-0.7 (0.5)	0.7-2.1 (1.4)	0.8-2.6 (1.3)	MLP 6-2-1 85%-82%-82% (0.5-1.4-1.2)
3	71-94 (85)	64-91 (73)	36-82 (73)	0.2-0.6 (0.4)	1.1-2.6 (1.5)	0.8-3.5 (1.9)	MLP 7-3-1 85%-82%-82% (0.6-1.5-1.6)
4	79-96 (88)	64-91 (73)	64-82 (82)	0.01-0.6 (0.4)	0.7-3.1 (1.5)	0.8-4.9 (2.0)	<b>MLP 7-4-1</b> <b>90%-82%-82%</b> <b>(0.4-1.2-1.4)</b>
6	77-96 (90)	64-91 (73)	64-82 (82)	0.00006 -0.6 (0.3)	0.6-5.2 (1.5)	0.9-7.9 (2.3)	MLP 7-6-1 90%-82%-82% (0.4-1.4-1.3)

**Tabela 27. Analiza wrażliwości dla najlepszego modelu wykorzystującego wektor wejściowy „krokowa postępująca” i posiadającego 4 neurony ukryte**

Zmienna	kaszel	ból czoła	ból głowy	zmęczenie	fetor	polipektomie	ESS
Iloraz	0.9332	1.0352	1.0158	1.0926	1.1037	1.0592	0.9649
Ranga	7	4	5	2	1	3	6

## ➤ ETAP II

W II etapie badań przeprowadzono eksperymenty wykorzystujące następujące architektury sieci:

- Wektor „statystyczny 1”, 2 neurony ukryte.  
 Jak opisano powyżej, najlepszy uzyskany dotychczas model o tej architekturze (MLP 11-2-1) wykazywał jakości uczenia, walidacji i testowania: 85%-82%-82% i błędy 0.5-1.6-1.2 oraz wartość parametru  $\rho = 1.8$ .  
 Przy pomocy *Automatycznego projektanta* przetestowano 1000 sieci o tej samej architekturze, uzyskując model lepszy od wyjściowego, tj. MLP 11-2-1 o jakościach uczenia, walidacji i testowania odpowiednio 90%-82%-82% i błędach 0.4-0.9-0.9. Następnie przeprowadzono analogiczny eksperyment po usunięciu z wektora „statystyczny 1” jednej zmiennej wejściowej (kaszel), ponieważ analiza wrażliwości przeprowadzona dla wyjściowego modelu sugerowała, że zmienna ta jest zbędna. Wynik eksperymentu dowiódł jednak, że usunięcie tej zmiennej z wektora wejściowego spowodowało pogorszenie działania sieci (jakości testowania osiągały maksymalną wartość 72%). W związku z tym nie kontynuowano redukcji wektora wejściowego na podstawie analizy wrażliwości wejść sieci.
- Wektor „krokowa postępująca”, 4 neurony ukryte.  
 Dla opisanego powyżej najlepszego modelu o tej architekturze (MLP 7-4-1) jakości uczenia, testowania i walidacji wynosiły 90%-82%-82%, a błędy 0.4-1.2-1.4. Przetestowanie przy pomocy *Automatycznego projektanta* 1000 podobnych sieci doprowadziło do uzyskania modelu MLP 6-4-1 o jakościach uczenia, walidacji i testowania 90%-82%-82% oraz błędach 0.3-1.0-1.9. Sieć wykorzystywała 6 zmiennych wejściowych (kaszel, ból czoła, ból głowy, zmęczenie, fetor, polipektomie), a analiza wrażliwości wejść nie wykazała obecności zbędnych zmiennych. Wartość parametru  $\rho$  wynosiła 1.5.  
 Ponieważ w wyjściowym modelu o rozważanej architekturze analiza wrażliwości sugerowała, że najmniejsze znaczenie ma zmienna „kaszel”, przetestowano 1000 sieci wykorzystujących wektor wejściowy „krokowa postępująca” zmniejszony o tę zmienną. Efektem było obniżenie maksymalnej jakości testowania do 54%.

Informacje o najlepszych uzyskanych modelach zestawiono w tabeli 28.

**Tabela 28. Najlepsze uzyskane modele dokonujące predykcji nasilenia dolegliwości po zabiegu w II grupie pacjentów.**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 11-2-1	płeć, liczba ESS, VAS_0, wydzielina, spływanie, kaszel, ból czoła, ból głowy, ból ucha, zmęczenie, fetor	90%-82%-82%	0.4-0.9-0.9	tak	tak
MLP 6-4-1	kaszel, ból czoła, ból głowy, zmęczenie, fetor, polipektomie	90%-82%-82%	0.3-1.0-1.9	nie	tak

### **6.2.3 Podsumowanie**

W grupie I zadowolający wynik operacji zatok zgłosiło 101 (72%) pacjentów, wynik niezadowolający 35 (25%), natomiast 4 pacjentów (3%) nie udzieliło odpowiedzi. Po odrzuceniu danych niekompletnych odsetek pacjentów zgłaszających wynik zadowolający wynosił 74%. Analiza statystyczna danych historycznych dotyczących wyników leczenia upoważnia zatem do informowania pacjentów, że każdy z nich ma 74% szans uzyskania pozytywnego wyniku leczenia operacyjnego. Należy jednak zauważyć, że na podstawie samej tylko analizy danych historycznych nie można orzekać, u których pacjentów prawdopodobieństwo powodzenia operacji jest większe, a u których mniejsze. Wyniki te nie mogą w związku z tym wspomagać decyzji dotyczącej wyboru optymalnego sposobu leczenia u poszczególnych chorych.

Najlepsze modele neuronowe przewidujące nasilenie dolegliwości po zabiegu, uzyskane dla grupy I, osiągały jakość testowania do 79% (należy przy tym pamiętać, że ocena jakości modelu może być w tym wypadku zaniżona ze względu na charakterystykę zbioru testowego). Wynik ten można interpretować następująco: dzięki zastosowaniu sieci neuronowych można z co najmniej 79% prawdopodobieństwem odróżnić pacjentów, u których wynik operacji będzie zadowolający, od chorych, którzy nie wyniosą korzyści z operacji. Pozwala to między innymi wskazać chorych mających niewielkie szanse na

satysfakcjonujący efekt leczenia operacyjnego. U tych pacjentów wskazane może być podjęcie decyzji o odstąpieniu od zabiegu na rzecz intensywnego leczenia zachowawczego.

Zastosowanie modelu neuronowego pozwala w pewnym stopniu poprawić predykcję wyników leczenia pod względem ilościowym (79% zamiast 74% prawdopodobieństwa), przede wszystkim jednak wnosi istotną różnicę jakościową (wytypowanie pacjentów o wyraźnie gorszym rokowaniu).

W przypadku grupy II zadowalający wynik operacji zanotowano – po uwzględnieniu braków danych - u 70% pacjentów. Najlepsze z modeli predykcyjnych uzyskanych dla tej grupy wykazywały jakość testowania 82%. Oznacza to nie tylko jakościową, ale również ilościową poprawę predykcji efektów operacji obserwowanych 12 miesięcy po zabiegu. Powyżej opisano trudności z opracowaniem sieci neuronowej spełniającej wszystkie przyjęte kryteria definiujące wiarygodny model. Wyniki eksperymentów wskazują, że dysponując obszerniejszą bazą danych można się spodziewać dalszej poprawy jakości wiarygodności predykcji.

### **6.3 Zmiana nasilenia dolegliwości po zabiegu operacyjnym w porównaniu do stanu przed zabiegiem**

**W tym podrozdziale, podobnie jak w kolejnych podrozdziałach, ze względu na ograniczoną objętość pracy skrótowo przedstawiono tylko najistotniejsze wyniki uzyskane na poszczególnych etapach eksperymentów. Jest to uzasadnione faktem, że obliczenia prowadzono w sposób analogiczny jak w przypadku omawianej wcześniej zmiennej VAS\_bin. Szczegółowe komentarze dotyczące przebiegu badań oraz sposobu prezentacji wyników można więc znaleźć w poprzednim podrozdziale.**

Zmianę nasilenia dolegliwości po zabiegu operacyjnym w porównaniu do stanu przed zabiegiem opisano przy pomocy zmiennej „dVAS”. Ujemne wartości tej zmiennej oznaczają zmniejszenie dolegliwości (poprawę), a dodatnie – ich nasilenie (pogorszenie). Przyjęto następujące nazwy zmiennych:

- dVAS\_6 – zmiana nasilenia dolegliwości w skali VAS po 3-6 miesiącach w stosunku do stanu sprzed zabiegu (dotyczy grupy I); wartość zmiennej wyliczana jest ze wzoru:

$$dVAS_6 = VAS_6 - VAS_0$$

- dVAS\_12 – zmiana nasilenia dolegliwości w skali VAS po 12 miesiącach w stosunku do stanu sprzed zabiegu (dotyczy grupy II); wartość wyliczana jest ze wzoru:

$$dVAS_{12} = VAS_{12} - VAS_0$$



### 6.3.1 Analiza statystyczna

Wartości zmiennej dVAS dla grupy I przedstawia tabela 29, a dla grupy II tabela 30. Zależności między zmienną dVAS a zmiennymi wejściowymi budowanych modeli zestawiono w tabeli 31.

Tabela 29. Wartości zmiennej dVAS w grupie I

Cecha	Wartości	Brak danych
zmiana dolegliwości 3-6 miesięcy po zabiegu (dVAS_6)	zakres: -10 do 4, mediana: -4	6 (4%)

Tabela 30. Wartości zmiennej dVAS w grupie II

Cecha	Wartości	Brak danych
zmiana dolegliwości 12 miesięcy po zabiegu (dVAS_12)	zakres: -10 do 3, mediana: -4.5	2 (3%)

Tabela 31. Zależności między pooperacyjnymi zmianami nasilenia objawów wyrażonymi w skali VAS a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).

Cecha	dVAS_6 (grupa I)	dVAS_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	p = 0.021447	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	R = -0.193650 p = 0.031864	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	-	-
alergia	-	-
refluks	-	-
przebyte polipektomie	-	-

Cecha	dVAS_6 (grupa I)	dVAS_12 (grupa II)
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.015132	-
liczba przebytych ESS	R = 0.213167 p = 0.013757	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	R = -0.504609 p = 0.000000	R = -0.366890 p = 0.001784
niedrożność nosa	R = -0.320337 p = 0.000191	R = -0.364487 p = 0.002077
wydzielina z nosa	R = -0.206634 p = 0.017017	R = -0.394544 p = 0.000795
splywanie wydzieliny	-	-
kaszel	R = -0.263131 p = 0.002301	-
kichanie	-	R = -0.302134 p = 0.012277
ból twarzy	-	-
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	-
zaburzenia węchu	R = -0.188187 p = 0.030066	(blisko granicy istotności) R = -0.224893 p = 0.065209
zmęczenie	-	-
nieprzyjemny zapach z ust	-	-
ból zębów	-	-

Na potrzeby klasyfikacyjnych sieci neuronowych niezbędne było przekształcenie zmiennej dVAS w zmienną binarną. Istotny problem stanowił w tym przypadku wybór właściwego progu odcięcia, tj. decyzja, jak dużą zmianę dolegliwości wyrażoną w skali VAS należy traktować jako klinicznie istotną poprawę. Dostępne piśmiennictwo nie dostarcza na ten temat jednoznacznych wskazówek. Definiując próg istotności klinicznej dla zmiany dolegliwości związanych z PZZP wyrażonych w skali VAS, Soler i wsp. oparli się na wynikach badań

dotyczących ostrego bólu (gdzie za klinicznie istotną uznano zmianę w skali VAS o nie mniej niż 0.9-1.3) oraz wynikami badań psychologicznych (wykazujących, że próg dyskryminacji wynosi około  $\frac{1}{2}$  odchylenia standardowego od wartości wyjściowej) [87]. Kierując się pierwszym z kryteriów należy założyć, że za „subiektywną poprawę po operacji” można uważać pozytywną zmianę o co najmniej 2 punkty w skali VAS. W przypadku drugiego z kryteriów, ze względu na brak normalności rozkładu zmiennej VAS\_0 w badanych grupach, posłużono się rozstępem kwartylowym (odchyleniem ćwiartkowym) zamiast odchylenia standardowego. Rozstęp kwartylowy dla zmiennej VAS\_0 w obu grupach wynosił 3. Przyjęto zatem próg dyskryminacji równy 1.5, potwierdzając wniosek, że jako minimalną klinicznie istotną należy traktować zmianę o 2 punkty w skali VAS.

Zgodnie z powyższymi założeniami następująco zdefiniowano zmienną binarną dVAS\_bin:

- jeżeli  $dVAS > -2$  to  $dVAS\_bin = 0$  (brak poprawy),
- jeżeli  $dVAS \leq -2$  to  $dVAS\_bin = 1$  (poprawa).

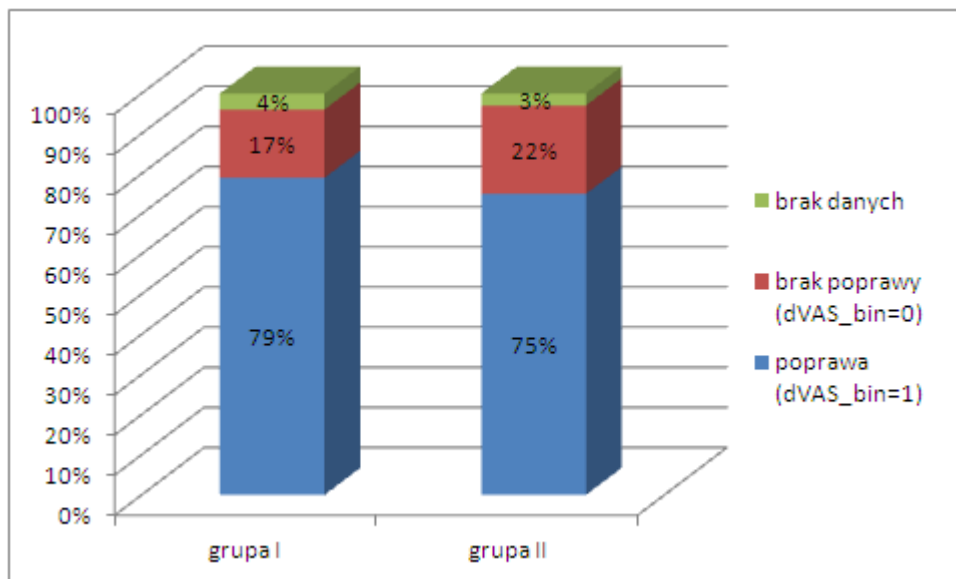
Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (ocenianej 3-6 miesięcy po zabiegu): dVAS\_bin\_6,
- dla grupy II (ocenianej 12 miesięcy po zabiegu): dVAS\_bin\_12.

**Model przewidujący wartość zmiennej dVAS\_bin odpowiada na pytanie pacjenta:  
„Czy po operacji moje dolegliwości związane z przewlekłym zapaleniem zatok  
przynosowych będą zauważalnie mniejsze niż obecnie?”**

Rozkład zmiennej dVAS\_bin w obu grupach chorych przedstawiono na rysunku 31. Zwraca uwagę znaczna przewaga liczebna tych pacjentów, którzy zgłosili poprawę po operacji. Tego rodzaju nierównowaga jest na ogół bardzo niekorzystna w przypadku modelowania z zastosowaniem sieci neuronowych, ponieważ powoduje, że jedna z wartości (w tym przypadku  $dVAS\_bin=1$ ) jest „faworyzowana” przez model. Na ogół przy stosowaniu algorytmów należących do kategorii *machine learning* (a do tej kategorii należą sieci neuronowe) dąży się do eliminacji tego rodzaju dysproporcji. Można to osiągnąć poprzez wykluczenie z analiz odpowiedniej liczby przypadków z części reprezentowanej klasy, jednak jest to uzasadnione tylko wtedy, gdy nierównowaga w grupie badanej nie odzwierciedla proporcji występujących rzeczywiście w populacji. Tymczasem - w odniesieniu do dolegliwości związanych z PZZP - przegląd literatury i doświadczenie dowodzi, że faktycznie większość chorych uzyskuje poprawę po ESS. Dlatego wyeliminowanie z analiz części pacjentów, u których zanotowano poprawę, stanowiłoby w istocie zafałszowanie danych. Inną skuteczną metodą poprawy funkcjonowania modelu jest taka zmiana przyjętego progu binaryzacji, aby uzyskać w przybliżeniu równe proporcje w obu grupach [226]. Jednak, jak przedyskutowano powyżej, arbitralne przyjęcie progu innego niż 2 punkty w skali VAS nie znajduje uzasadnienia klinicznego. Przyjmując, że kryteria medyczne mają w tym wypadku zdecydowany priorytet, postanowiono zaakceptować przyjęty próg mimo świadomości negatywnych konsekwencji tego wyboru dla jakości modelowania.

**Rysunek 31. Odsetki pacjentów w grupie I i II zgłaszających subiektywną poprawę po operacji (odpowiednio po 3-6 miesiącach i po 12 miesiącach od zabiegu) przy założeniu, że klinicznie istotna jest wartość zmiany o co najmniej 2 punkty w skali VAS**



W tabeli 32 przedstawiono wyniki analiz statystycznych dotyczących zależności między zmienną dVAS\_bin a zmiennymi wejściowymi.

**Tabela 32. Zależności między pooperacyjną zmianą objawów wyrażoną za pomocą zmiennej dVAS\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	dVAS_bin_6 (grupa I)	dVAS_bin_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	p=.03379	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	-	-
alergia	-	-

Cecha	dVAS_bin_6 (grupa I)	dVAS_bin_12 (grupa II)
refluks	-	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	-	p=.01954
liczba przebytych ESS	-	(blisko granicy istotności) p = 0.078023
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	p = 0.001118	-
niedrożność nosa	p = 0.023293	p = 0.006758
wydzielina z nosa	-	p = 0.004461
splywanie wydzieliny	-	-
kaszel	-	-
kichanie	-	p = 0.010068
ból twarzy	-	-
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	-
zaburzenia węchu	-	(blisko granicy istotności) p = 0.068503
zmęczenie	-	-
nieprzyjemny zapach z ust	-	-
ból zębów	-	-

### **6.3.2 Modele neuronowe przewidujące zmianę nasilenia dolegliwości po zabiegu operacyjnym**

#### **6.3.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa dVAS\_bin\_6)**

##### **➤ ETAP I:**

##### **(a) Poszukiwanie wektora wejściowego**

W tabeli 33 zestawiono jakości i błędy sieci neuronowych otrzymanych dla różnych wektorów wejściowych. Wykorzystano podział zbioru danych na podzbiory w proporcjach

70% : 15% : 15%. Po uwzględnieniu braków danych liczby przypadków w poszczególnych podzbiorach wynosiły odpowiednio 95 : 20 : 18.

Testowano następujące wektory wejściowe:

- Wszystkie dostępne zmienne (29 zmiennych).  
**Wynik testu:** Użycie wektora zawierającego wszystkie 29 zmiennych wejściowych pozwoliło uzyskać sieci o bardzo wysokich jakościach uczenia, walidacji i testowania, jednak analiza wartości błędów uczenia, walidacji i testowania wskazywała na ich niską zdolność generalizacji. Wynik ten wskazuje, że użycie takiego wektora mogłoby prowadzić do uzyskania bardzo efektywnego modelu, jednak niezbędny do tego byłby znacznie liczniejszy zbiór przypadków uczących.
- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” - 3 zmienne, od których zmienna dVAS\_bin\_6 była istotnie statystycznie zależna (por. tab. 32).  
**Wynik testu:** Mimo bardzo niewielkiej liczby zmiennych wejściowych, sieci neuronowe wykorzystujące ten wektor wykazywały wysokie jakości i niskie błędy uczenia, walidacji i testowania.
  - Wektor „statystyczny 2” - 8 zmiennych, od których statystycznie istotnie zależna była zmienna dVAS (tab. 31).  
**Wynik testu:** Zwiększenie wektora wejściowego pogorszyło zdolność modelu do generalizacji.
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 15 zmiennych (płeć, wiek, polipy, TK, histopat, czas objawów, VAS\_0, sphywanie, kichanie, ból czoła, ból głowy, wdech, alergia, refluks, liczba polipektomii).  
**Wynik testu:** Wykorzystanie tego wektora pozwoliło uzyskać sieci o bardzo wysokich jakościach uczenia, walidacji i testowania oraz niskich błędach.
  - Wektor „krokowa wsteczna” – 3 zmienne (skrzywienie przegrody, TK, zabieg).  
**Wynik testu:** Sieci uzyskane dla tego wektora były zdecydowanie mniej efektywne niż opisane powyżej.
  - Wektor „algorytm genetyczny” - 11 zmiennych (wiek, polipy, skrzywienie przegrody, TK, niedrożność, kaszel, ból ucha, ból zębów, astma, AIA, refluks).  
**Wynik testu:** Uzyskane sieci wykazywały niskie jakości testowania.

**Tabela 33. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych.**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wszystkie dostępne zmienne (29)	87-100 (99)	35-85 (65)	50-89 (72)	0.0005-0.4 (0.02)	0.2-3.2 (1.6)	1.9-7.3 (3.6)	MLP 29-11-1 100%-85%-89% (0.02-1.7-3.7)
wektor „statystyczny 1” (3)	67-82 (73)	45-75 (60)	61-89 (72)	0.5-0.7 (0.6)	0.3-0.6 (0.4)	0.4-0.9 (0.6)	<b>MLP 3-9-1</b> <b>82%-75%-89%</b> <b>(0.6-0.4-0.8)</b>
wektor „statystyczny 2” (8)	77-94 (86)	55-80 (75)	61-78 (72)	0.4-0.6 (0.4)	0.4-1.2 (0.9)	0.9-1.7 (1.6)	MLP 8-11-1 88%-80%-78% (0.4-1.0-1.3)
wektor „krokowa postępująca” (15)	88-100 (99)	50-100 (80)	67-94 (83)	0.02-0.4 (0.1)	0.2-1.7 (0.7)	0.3-2.8 (1.2)	<b>MLP 15-10-1</b> <b>100%-95%-89%</b> <b>(0.2-0.3-0.5)</b> <b>MLP 15-11-1</b> <b>100%-90%-94%</b> <b>(0.1-0.6-0.6)</b>
wektor „krokowa wsteczna” (3)	47-72 (59)	30-70 (60)	33-72 (56)	0.6-0.9 (0.6)	0.4-0.5 (0.4)	0.6-1.0 (0.7)	MLP 3-9-1 65%-70%-61% (0.6-0.4-0.7)
wektor „algorytm genetyczny” (11)	79-95 (89)	55-85 (70)	61-72 (67)	0.2-0.4 (0.2)	0.5-1.8 (1.1)	2.9-5.1 (3.9)	MLP 11-12-1 96%-85%-67% (0.2-0.9-3.2)

**(b) Ocena liczby neuronów ukrytych**

Ocenę liczby neuronów ukrytych przeprowadzono dla dwóch wektorów wejściowych, których wykorzystanie pozwoliło uzyskać najlepsze modele w omówionych powyżej testach, tj. wektorów „statystyczny 1” i „krokowa postępująca”.

## **Wektor wejściowy „statystyczny 1”**

W tabeli 34 przedstawiono wyniki testów sieci neuronowych wykorzystujących wektor wejściowy „statystyczny 1”, lecz posiadających różną liczbę neuronów w warstwie ukrytej. Podobnie jak w przypadku omawianej wcześniej zmiennej VAS\_bin, zwiększanie rozmiarów sieci powodowało zmniejszanie błędu uczenia przy wzroście błędów walidacji i testowania, jednak dla analizowanego obecnie zagadnienia zmiany te były stosunkowo niewielkie.

Najlepsze modele uzyskiwały jakość testowania do 89% niezależnie od liczby neuronów ukrytych (w rozważanym przedziale wartości tego parametru). Analizując uzyskiwane wyniki stwierdzono, że pojedyncze przypadki w zbiorach walidacyjnym i testowym okazały się szczególnie trudne do prawidłowego zaklasyfikowania na podstawie danych dostarczanych przez rozważany wektor wejściowy. Dla kilku pacjentów większość modeli niezmiennie dokonywała błędnych predykcji. Prawidłowość tę potwierdzono dodatkowo zamieniając zbiór walidacyjny z testowym. W efekcie uzyskano jakości walidacji około 89%, a jakości testowania 75%, tj. dokładnie odwrotnie niż w modelach prezentowanych w tabeli 34. Następnie dokonano losowego podziału na podzbiory uczący, walidacyjny i testowy i uzyskano jakości walidacji i testowania około 80% oraz nieprawidłowe wyniki dokładnie dla tych samych przypadków, które były błędnie przyporządkowane przez omawiane wcześniej modele. Można przypuszczać, że były to przypadki w jakimś sensie nie pasujące do ogólnej reguły, którą wykrywała i ustalała każda ucząca się sieć neuronowa.

W związku z uzyskaniem równie efektywnych modeli nawet dla sieci posiadających tylko 1 neuron ukryty, przeprowadzono także dodatkowe testy z wykorzystaniem sieci liniowych (z definicji w ogóle pozbawionych warstwy ukrytej), jednak stwierdzono, że skuteczność tych modeli była znacznie niższa (jakość testowania 61-72%). Można zatem wnioskować, że zadany problem jest rozwiązywany przez sieci neuronowe w bardzo podobny sposób niezależnie od liczby neuronów ukrytych, jednak zastosowanie modelu nieliniowego jest niezbędne.

Podsumowując powyższe rozważania można założyć, że wykorzystanie 3 zmiennych składających się na wektor wejściowy „statystyczny 1” w rzeczywistości pozwala dokonać prawidłowych predykcji u około 80% pacjentów. Najprawdopodobniej ilość informacji zawarta w tym niewielkim wektorze jest jednak zbyt mała, by można było oczekiwać znacznego zwiększenia jakości testowania na etapie dalszych badań, nawet przy znacznie większej liczbie przypadków uczących.



**Tabela 34. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 1”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
<b>1</b>	71-80 (73)	50-75 (60)	89-94 (89)	0.7-0.8 (0.7)	0.3-0.5 (0.3)	0.4-0.6 (0.5)	MLP 2-1-1 80%-75%-89% 0.7-0.3-0.6
<b>2</b>	70-82 (77)	53-74 (74)	71-94 (82)	0.6-0.7 (0.7)	0.3-0.5 (0.4)	0.4-0.8 (0.6)	MLP 2-2-1 80%-73%-88% (0.7-0.3-0.7)
<b>3</b>	67-80 (72)	45-85 (60)	67-89 (72)	0.6-0.7 (0.6)	0.3-0.5 (0.4)	0.5-0.8 (0.6)	MLP 2-3-1 80%-75%-89% (0.6-0.5-0.5)
<b>4</b>	67-80 (72)	45-75 (60)	67-89 (72)	0.5-0.7 (0.6)	0.3-0.6 (0.4)	0.5-0.9 (0.7)	MLP 2-4-1 80%-75%-89% (0.6-0.4-0.5)
<b>10</b>	67-80 (72)	45-75 (60)	61-83 (72)	0.5-0.7 (0.6)	0.3-0.6 (0.5)	0.5-0.8 (0.6)	MLP 3-10- 80%-75%-83% (0.6-0.4-0.8)

### **Wektor wejściowy „krokowa postępująca”**

Sieci neuronowe wykorzystujące wektor wejściowy „krokowa postępująca” wykazywały bardzo wysokie jakości uczenia, walidacji i testowania, wzrastające wraz ze wzrostem liczby neuronów ukrytych (tab. 35). Większe modele wykazywały jednak wyższe błędy testowania. Wartości parametru  $\rho$  sugerujące optymalną zdolność do uczenia i generalizacji otrzymano dla sieci posiadającej 3 neurony w warstwie ukrytej. Należy jednak zauważyć, że najwyższe jakości uczenia, walidacji i testowania (100%-100%-94%) wykazywał model posiadający aż 12 neuronów w warstwie ukrytej. Model ten posiada tak wiele połączeń, że jego zdolność do generalizacji może się wydawać dyskusyjna dla zbioru uczącego liczącego zaledwie 95 przypadków ( $\rho=0.5$ ), jednak niskie błędy walidacji i testowania (odpowiednio 0.2 i 0.3 przy błędzie uczenia równym 0.2) sugerują prawidłowe działanie nauczonej sieci. Analiza wrażliwości dla tej sieci nie wykazała też obecności zbędnych zmiennych.

**Tabela 35. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „krokowa postępująca”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
1	50-81 (73)	25-95 (45)	56-83 (61)	0.6-0.8 (0.7)	0.4-1.2 (0.7)	0.7-1.3 (0.9)	MLP 11-1-1 81%-75%-78% (0.6-0.4-0.7)
2	64-92 (79)	40-85 (55)	50-89 (67)	0.4-0.8 (0.6)	0.4-1.6 (0.8)	0.5-1.6 (1.0)	MLP 14-2-1 92%-85%-83% (0.6-0.5-0.7)
3	66-96 (83)	40-90 (65)	56-94 (72)	0.3-0.7 (0.5)	0.4-1.6 (0.9)	0.5-1.7 (0.9)	<b>MLP 13-3-1</b> <b>89%-90%-94%</b> <b>(0.5-0.7-0.6)</b>
4	68-99 (87)	40-95 (65)	56-89 (75)	0.2-0.6 (0.5)	0.3-1.7 (0.9)	0.4-1.9 (1.0)	MLP 15-4-1 96%-95%-89% (0.3-0.3-0.5)
6	80-100 (94)	45-90 (75)	56-94 (78)	0.1-0.6 (0.3)	0.2-1.5 (0.7)	0.4-2.3 (0.9)	MLP 15-6-1 98%-90%-94% (0.3-0.6-0.5)
8	86-100 (98)	60-95 (75)	61-100 (78)	0.06-0.4 (0.2)	0.3-1.5 (0.8)	0.3-2.1 (1.0)	MLP 15-8-1 99%-90%-94% (0.2-0.6-0.5)
10	89-100 (99)	55-100 (80)	67-100 (83)	0.04-0.4 (0.1)	0.1-1.6 (0.7)	0.3-2.3 (1.1)	MLP 15-10-1 100%-95%-94% (0.08-0.6-0.8)
12	91-100 (100)	60-100 (85)	67-94 (83)	0.001-0.3 (0.12)	0.2-1.6 (0.7)	0.2-2.9 (1.2)	<b>MLP 15-12-1 nr 106</b> <b>100%-100%-94%</b> <b>(0.2-0.2-0.3)</b>

## ➤ ETAP II

Za punkt wyjścia dla dalszych prac przyjęto sieci neuronowe wykorzystujące wektor wejściowy „krokowa postępująca” i posiadające odpowiednio 3 lub 12 neuronów w warstwie ukrytej. Informacje o najlepszych uzyskanych ostatecznie modelach zestawiono w tabeli 36.

- Wektor „krokowa postępująca”, 3 neurony ukryte.  
Najlepszym modelem uzyskanym w etapie I, traktowanym tutaj jako „punkt wyjścia”, była sieć MLP 13-3-1 o jakościach 89%-90%-94% i błędach 0.5-0.7-0.6.  
Przetestowano przy pomocy *Automatycznego projektanta sieci* 1000 modeli o tej samej architekturze, otrzymując model o nieco niższej jakości testowania (MLP 14-3-1, jakości: 94%-85%-89%, błędy: 0.5-0.3-0.9). Podobnie jak w modelu wyjściowym, również dla tej sieci analiza wrażliwości wykazała obecność zbędnych zmiennych. Ponieważ w modelu wyjściowym analiza wrażliwości sugerowała, że zbędna jest zmienna „ból czoła”, przetestowano 1000 sieci wykorzystujących wektor wejściowy bez tej zmiennej, jednak ponownie uzyskano nieco gorszy model (MLP 13-1-1, jakości 89%-85%-89%, błędy 0.6-0.3-0.5). Z kolei z tego modelu usunięto zmienną wskazaną przez analizę wrażliwości jako zbędna (refluks), uzyskując model o podobnej efektywności. Najlepszą siecią był perceptron MLP 12-3-1 o jakościach 94%-90%-89% i błędach 0.3-0.4-0.5.
- Wektor „krokowa postępująca”, 12 neuronów ukrytych.  
Przy pomocy *Automatycznego projektanta* przetestowano 1000 sieci o tej samej architekturze. Większość z nich wykazywała wyraźną tendencję do przeuczania się (błędy uczenia o jeden lub dwa rzędy wielkości niższe niż błędy walidacji i testowania). Najlepszy uzyskany model MLP 15-12-1 wykazał równie wysokie jakości uczenia, walidacji i testowania jak model wyjściowy (100%-100%-94%) przy błędach 0.1-0.3-0.6. Przeprowadzone testy dowodzą, że rozważany wektor wejściowy pozwala dokonywać prawidłowych predykcji u niemal każdego pacjenta. Należy oczekiwać, że dysponując większym zbiorem uczącym można będzie uzyskać jeszcze bardziej wiarygodną i dokładną prognozę.

**Tabela 36. Najlepsze modele dokonujące predykcji zmian nasilenia dolegliwości po zabiegu operacyjnym w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 12-3-1	pleć, polipy, TK, hist-pat, czas objawów, VAS_0, spływanie, kichanie, ból głowy, węch, alergia, liczba polipektomii	94%-90%-89%	0.3-0.4-0.5	tak	tak
MLP 15-12-1	pleć, wiek, polipy, TK, hist-pat, czas objawów, VAS_0, spływanie, kichanie, ból czoła, ból głowy, węch, alergia, refluks, liczba polipektomii	100%-100%-94%	0.2-0.2-0.3	nie	nie

### 6.3.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa dVAS\_bin\_12)

**Ponieważ obliczenia dotyczące II grupy pacjentów były prowadzone w sposób analogiczny do opisanego powyżej, aby nie zwiększać dodatkowo rozmiarów pracy poniżej przedstawiono wyłącznie ostateczne wyniki eksperymentów, tj. opis najlepszych modeli neuronowych uzyskanych dla tej grupy chorych. Według podobnej zasady prezentowane są wyniki w kolejnych podrozdziałach.**

Obliczenia dla II grupy chorych prowadzono korzystając z podziału zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 70% : 15% : 15% (liczby przypadków 48 : 11 : 11). Spośród testowanych wektorów wejściowych najbardziej obiecujący okazał się wektor zawierający 6 zmiennych, od których statystycznie istotnie zależna była zmienna dVAS\_bin\_12 lub wynik testu był bliski granicy istotności statystycznej

(por. tab. 32). Najlepsze modele posiadały 3 neurony w warstwie ukrytej. Po przeprowadzeniu drugiego etapu eksperymentów za najefektywniejszą uznano sieć neuronową, której charakterystyka przedstawiona jest w tabeli 37.

**Tabela 37. Najlepszy model dokonujący predykcji zmian nasilenia dolegliwości po zabiegu operacyjnym w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 6-3-1	niedrożność, wydzielina, kichanie, wdech, ESS, liczba ESS	88% -82% -82%	0.4-0.6-0.7	nie	tak

### **6.3.3 Podsumowanie**

Dla I grupy pacjentów zbudowano modele neuronowe dokonujące prawidłowych predykcji nasilenia dolegliwości po zabiegu operacyjnym u 94% pacjentów. Oznacza to, że dzięki zastosowaniu sieci neuronowych można z prawdopodobieństwem graniczącym z pewnością przewidywać, którzy z chorych mogą oczekiwać poprawy w okresie 3-6 miesięcy operacji, a u których celowość przeprowadzenia zabiegu jest wątpliwa. Wyniki eksperymentów sugerują, że dla liczniejszych zbiorów danych uczących będzie możliwe zbudowanie jeszcze skuteczniejszych narzędzi predykcyjnych. Dla grupy II uzyskano nieco mniej efektywne modele (jakość testowania do 82%), co najprawdopodobniej wynika z mniejszej liczebności tej grupy chorych, nie musi natomiast świadczyć o większych trudnościach w przewidywaniu wyniku w okresie 12 miesięcy po operacji.

## **6.4 Upośledzenie drożności nosa**

W niniejszym podrozdziale omówiono modele neuronowe dokonujące predykcji upośledzenia drożności nosa po zabiegu operacyjnym. Jest to pierwsza z dolegliwości, które pacjenci oceniali w skali 4-stopniowej (od 0 oznaczającego brak objawu do 3 oznaczającego objaw bardzo nasilony).

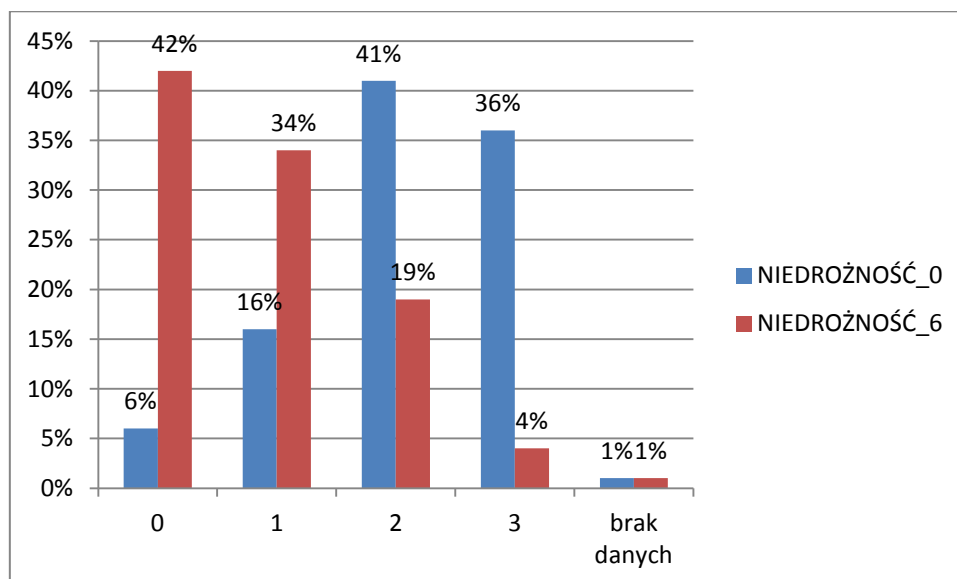
Przyjęto następujące nazwy zmiennych:

- NIEDROŻNOŚĆ\_0 – nasilenie niedrożności nosa w skali 4-stopniowej przed zabiegiem,  
NIEDROŻNOŚĆ\_6 – nasilenie niedrożności nosa w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- NIEDROŻNOŚĆ\_12 – nasilenie niedrożności nosa w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

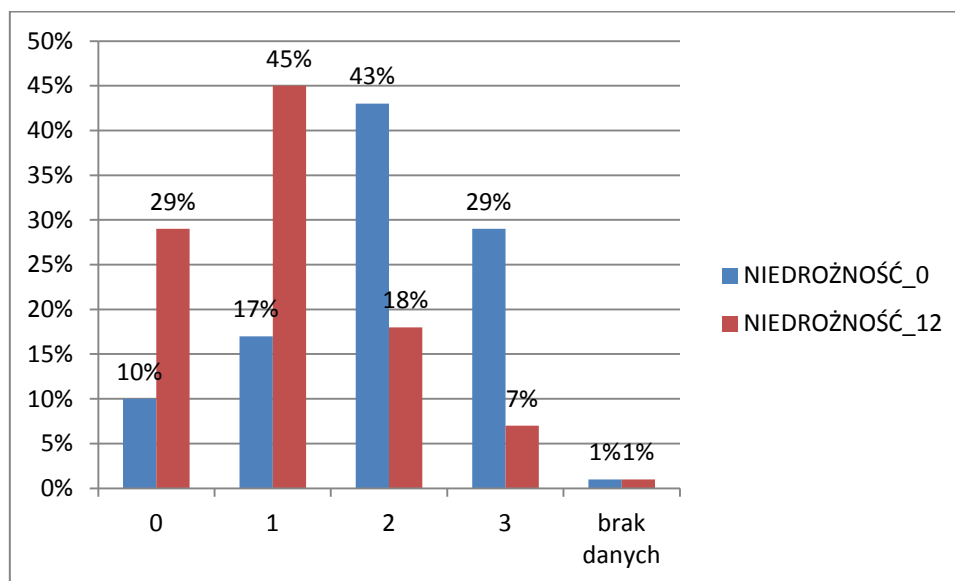
#### **6.4.1 Analiza statystyczna**

Wartości zmiennej NIEDROŻNOŚĆ podane przez pacjentów z grupy I przed zabiegiem i 3-6 miesięcy po zabiegu przedstawiono na rysunku 32. Wartości zgłoszone przez pacjentów z grupy II przed zabiegiem i 12 miesięcy po zabiegu przedstawiono na rysunku 33. Zależności między zmienną NIEDROŻNOŚĆ a zmiennymi wejściowymi budowanych modeli zestawiono w tabeli 38.

**Rysunek 32. Wartości zmiennej NIEDROŻNOŚĆ w grupie I – przed zabiegiem (NIEDROŻNOŚĆ\_0) i 3-6 miesięcy po zabiegu (NIEDROŻNOŚĆ\_6)**



**Rysunek 33. Wartości zmiennej NIEDROŻNOŚĆ w grupie II – przed zabiegiem (NIEDROŻNOŚĆ\_0) i 12 miesięcy po zabiegu (NIEDROŻNOŚĆ\_12)**



**Tabela 38. Zależności między nasileniem zgłaszanej po operacji niedrożności nosa (w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	NIEDROŻNOŚĆ_6 (grupa I)	NIEDROŻNOŚĆ_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	(blisko granicy istotności) R = 0.233281 p = 0.050245
astma	p = 0.041907	-
AIA	p = 0.012343	-
alergia	-	-
refluks	-	-

Cecha	NIEDROŻNOŚĆ_6 (grupa I)	NIEDROŻNOŚĆ_12 (grupa II)
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.010216	-
liczba przebytych ESS	R = 0.234270 p = 0.005862	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	(blisko granicy istotności) R = 0.167311 p = 0.052429	-
wydzielina z nosa	-	-
splýwanie wydzieliny	-	-
kaszel	-	-
kichanie	R = 0.185986 p = 0.030792	-
ból twarzy	-	-
ból czoła	-	-
ból głowy	R = 0.221972 p = 0.009669	-
dolegliwości uszne	-	-
zaburzenia węchu	-	-
zmęczenie	R = 0.258282 p = 0.002586	-
nieprzyjemny zapach z ust	-	-
ból zębów	-	-

W dalszej części opisywanych tu badań a potrzeby budowy klasyfikacyjnych sieci neuronowych przekształcono zmienną NIEDROŻNOŚĆ w zmienną binarną NIEDR\_bin. Kierując się rozkładem wartości zmiennej NIEDROŻNOŚĆ w obu grupach chorych przyjęto, że zadaniem modeli będzie rozróżnienie chorych, u których po zabiegu dochodzi do całkowitego ustąpienia niedrożności nosa, od pacjentów, u których objaw ten się utrzymuje w stopniu małym, średnim bądź znacznym. Przyjęcie takiego progu binaryzacji powoduje stosunkowo korzystną dystrybucję nowej zmiennej binarnej w obu grupach mimo



obserwowanych różnic między grupami (po uwzględnieniu braków danych proporcje 42%:58% w grupie I i 30%:70% w grupie II).

Alternatywna binaryzacja polegałaby na wyróżnieniu grupy pacjentów, u których po operacji niedrożność nosa nie występuje wcale bądź występuje jedynie w małym stopniu. Zadaniem modelu byłoby w tym przypadku odróżnienie ich od chorych, u których po operacji utrzymuje się średnia lub znaczna niedrożność nosa. Przyjęcie takiego progu binaryzacji prowadziło jednak do nieco mniej równomiernego podziału w proporcjach 77%:23% w grupie I i 75%:25% w grupie II (po uwzględnieniu braków danych). Ponadto założono, że informacja o możliwości całkowitego udrożnienia nosa dzięki operacji jest bardziej jednoznaczna i istotniejsza z punktu widzenia pacjenta.

W związku z tym zmienną NIEDR\_bin zdefiniowano następująco:

- jeżeli NIEDROŻNOŚĆ=0 to NIEDR\_bin = 0 (niedrożność nosa po zabiegu nie występuje),
- jeżeli NIEDROŻNOŚĆ>0 to NIEDR\_bin = 1 (po zabiegu utrzymuje się mała, średnia lub znaczna niedrożność nosa).

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (ocenianej 3-6 miesięcy po zabiegu): NIEDR\_bin\_6,
- dla grupy II (ocenianej 12 miesięcy po zabiegu): NIEDR\_bin\_12.

**Model przewidujący wartość zmiennej NIEDR\_bin odpowiada na pytanie pacjenta:  
„Czy po operacji mój nos będzie drożny?”**

Wyniki analiz statystycznych dotyczących zależności między zmienną NIEDR\_bin a zmiennymi wejściowymi budowanych modeli zestawiono w tabeli 39. Zwraca uwagę fakt, że w II grupie pacjentów wartość zmiennej NIEDR\_bin po 12 miesiącach od zabiegu nie okazała się statystycznie istotnie zależna od żadnej z wykorzystywanych przez sieci neuronowe zmiennych.

**Tabela 39. Zależności pomiędzy zgłaszaną po operacji niedrożnością nosa wyrażoną za pomocą zmiennej NIEDR\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	NIEDR_bin_6 (grupa I)	NIEDR_bin_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-

Cecha	NIEDR_bin_6 (grupa I)	NIEDR_bin_12 (grupa II)
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	p=.04151	-
AIA	-	-
alergia	-	-
refluks	(blisko granicy istotności) p=.05073	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	-	-
liczba przebytych ESS	-	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	p = 0.042904	-
wydzielina z nosa	(blisko granicy istotności) p = 0.058962	-
splýwanie wydzieliny	-	-
kaszel	p = 0.029401	-
kichanie	-	-
ból twarzy	-	-
ból czoła	-	-
ból głowy	(blisko granicy istotności) p = 0.058910	-
dolegliwości uszne	-	-
zaburzenia węchu	-	-
zmęczenie	0.017016	-
nieprzyjemny zapach z ust	-	-
ból zębów	-	-

## **6.4.2 Modele neuronowe przewidujące upośledzenie drożności nosa po zabiegu operacyjnym**

### **6.4.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa NIEDR\_bin\_6)**

#### **➤ ETAP I:**

#### **(a) Poszukiwanie wektora wejściowego**

Informacje o sieciach neuronowych wykorzystujących różne wektory wejściowe zestawiono w tabeli 40. Dla omawianych eksperymentów proporcje liczby przypadków w podzbiorach uczącym, walidacyjnym i testowym wynosiły 70% : 20% : 20%, tj. po uwzględnieniu braków danych 97 : 20 : 20 przypadków. Wyniki testów dotyczących poszczególnych wektorów wejściowych były następujące:

- Wszystkie dostępne zmienne (29 zmiennych).  
**Wynik testu:** Podobnie jak w przypadku zmiennych wyjściowych omawianych w poprzednich podrozdziałach, stwierdzono bardzo słabą zdolność modeli do generalizacji.
- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 4 zmienne, od których statystycznie istotnie zależna była zmienna NIEDR\_bin\_6 (astma, niedrożność, kaszel, zmęczenie),
  - Wektor „statystyczny 2” – 7 zmiennych, od których statystycznie istotnie zależna była zmienna NIEDR\_bin\_6 lub wynik był bliski granicy istotności statystycznej (astma, niedrożność, kaszel, zmęczenie, refluks, wydzielina, ból głowy),
  - Wektor „statystyczny 3” – 7 zmiennych, od których statystycznie istotnie zależna była zmienna NIEDROŻNOŚĆ\_6 (astma, AIA, ESS, kichanie, ból głowy, zmęczenie, niedrożność),
  - Wektor „statystyczny 4” - wszystkie 10 zmiennych zawartych w opisanych powyżej trzech wektorach.  
**Wynik testów:** Spośród opisanych powyżej wektorów jako najbardziej obiecujący wybrano wektor „statystyczny 2”, złożony z 7 zmiennych. Wektor złożony z 4 zmiennych nie zapewniał dostatecznej jakości uczenia, natomiast wektor zawierający 10 zmiennych powodował tendencję do przeuczania modeli.
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 16 zmiennych (masywne polipy, hist-pat, czas objawów, niedrożność, wydzielina, sływanie, kaszel, kichanie, ból twarzy, ból głowy, ból ucha, węch, zmęczenie, fetor, astma, refluks)  
**Wynik testu:** Ten wektor wejściowy okazał się zbyt duży i spowodował tendencję do przeuczania modeli.
  - Wektor „krokowa wsteczna” – 7 zmiennych (polipy, skrzywienie przegrody, zabieg, hist-pat, czas objawów, AIA, refluks)  
**Wynik testu:** Efektywność uzyskanych modeli okazała się niższa niż sieci

neuronowych wykorzystujących tej samej wielkości wektory opracowane na podstawie analizy statystycznej.

- Wektor „algorytm genetyczny” – 17 zmiennych (płeć, wiek, polipy, masywne polipy, TK, zabieg, VAS\_0, niedrożność, wydzielina, spływanie, kichanie, ból głowy, ból ucha, węch, zmęczenie, AIA, refluks)

**Wynik testu:** Mimo niskich błędów uzyskiwanych dla tego wektora na etapie doboru cech, stwierdzono, że jest on zbyt duży dla właściwego rozwiązania rozważanego problemu i powoduje utratę zdolności modeli do generalizacji.

**Tabela 40. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wszystkie dostępne zmienne (29)	94-100 (100)	35-70 (45)	35-65 (55)	0.00001 -0.3 (0.007)	2.6-52.6 (5.5)	1.2-11.2 (4.1)	MLP 29-12-1 100%-65%-65% (0.00006-8.0-8.3)
wektor „statystyczny 1” (4)	55-71 (66)	40-75 (60)	35-85 (50)	0.6-0.7 (0.7)	0.7-0.8 (0.7)	0.6-0.9 (0.7)	MLP 4-10-1 71%-65%-60% (0.7-0.8-0.7)
wektor „statystyczny 2” (7)	68-90 (79)	35-75 (55)	60-90 (75)	0.4-0.6 (0.6)	0.7-1.6 (1.0)	0.5-1.0 (0.7)	<b>MLP 7-11-1</b> <b>86%-75%-80%</b> <b>(0.5-0.9-0.8)</b>
wektor „statystyczny 3” (7)	61-84 (75)	25-80 (60)	40-85 (70)	0.5-0.6 (0.6)	0.6-0.9 (0.8)	0.6-1.2 (0.8)	MLP 7-9-1 69%-75%-70% 0.6-0.7-0.7
wektor „statystyczny 4” (10)	75-97 (87)	40-75 (50)	40-80 (65)	0.3-0.6 (0.4)	0.7-2.2 (1.2)	0.6-2.1 (1.2)	MLP 10-11-1 93%-75%-70% (0.3-1.2-1.3)

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „krokowa postępująca” (16)	95-100 (100)	45-85 (65)	40-75 (60)	0.006-0.3 (0.09)	1.2-4.1 (2.3)	1.0-4.7 (2.3)	MLP 16-9-1 95%-75%-70% (0.3-1.5-1.2)
wektor „krokowa wsteczna” (7)	68-80 (76)	40-65 (50)	50-75 (60)	0.55-0.6 (0.6)	1.1-2.4 (1.8)	0.7-1.4 (1.0)	MLP 7-12-1 75%-60%-60% (0.6-1.8-1.1)
wektor „algorytm genetyczny” (17)	92-100 (100)	30-65 (45)	50-90 (65)	0.008-0.3 (0.08)	1.8-6.9 (4.0)	1.0-3.3 (1.8)	MLP 17-12-1 100%-65%-90% (0.05-3.6-1.5)

### **(b) Ocena liczby neuronów ukrytych**

W kolejnym etapie prac wykorzystano wektor wejściowy „statystyczny 2”. Wyniki eksperymentów zestawiono w tabeli 41. Zwracają uwagę wysokie jakości testowania uzyskiwane nawet przez bardzo małe modele o niskich jakościach uczenia. Obserwacja ta może świadczyć o przypadkowym doborze łatwych przypadków w zbiorze testowym. Modele, dla których jakość uczenia i walidacji była znacznie niższa niż jakość testowania, uznano za mało wiarygodne. Ponadto, mimo że teoretycznie parametr  $\rho$  osiąga optymalną wartość dla sieci o około 5 neuronach ukrytych, to w praktyce akceptowalne proporcje jakości uczenia, walidacji i testowania zanotowano dopiero dla modeli posiadających 8 lub więcej neuronów ukrytych. Najlepszą ze zbudowanych sieci był perceptron mający aż 11 neuronów w warstwie ukrytej. Ze względu na niskie błędy i wysokie jakości uczenia, walidacji oraz testowania, zdecydowano wykorzystać tę właśnie rozbudowaną sieć jako model wyjściowy w drugim etapie eksperymentów. (Warto zauważyć, że taką samą architekturę miała najlepsza sieć zbudowana na etapie doboru wektora wejściowego – por. tab. 40.)

**Tabela 41. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 2”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
<b>1</b>	56-68 (61)	45-75 (60)	55-95 (80)	0.66-0.72 (0.7)	0.7-0.8 (0.7)	0.6-0.7 (0.6)	MLP 5-1-1 63%-75%-80% (0.7-0.7-0.7)
<b>2</b>	43-74 (64)	40-75 (60)	40-90 (70)	0.6-0.7 (0.7)	0.6-1.0 (0.7)	0.6-0.8 (0.7)	MLP 7-2-1 64%-75%-65% (0.7-0.6-0.7)
<b>4</b>	59-77 (71)	45-75 (55)	45-95 (70)	0.6-0.7 (0.7)	0.6-1.1 (0.8)	0.5-1.0 (0.7)	MLP 6-4-1 70%-70%-85% (0.7-0.7-0.6)
<b>6</b>	63-83 (74)	40-70 (55)	45-90 (75)	0.5-0.7 (0.6)	0.6-1.7 (0.8)	0.5-1.0 (0.7)	MLP 7-6-1 76%-70%-75% (0.6-0.9-1.0)
<b>8</b>	65-88 (75)	45-75 (55)	55-90 (75)	0.4-0.7 (0.6)	0.7-1.7 (0.9)	0.5-1.0 (0.7)	MLP 7-8-1 80%-75%-85% (0.5-0.8-0.9)
<b>10</b>	68-84 (77)	45-70 (55)	60-90 (75)	0.5-0.6 (0.6)	0.7-1.6 (0.9)	0.5-1.0 (0.7)	MLP 7-10-1 78%-70%-80% (0.5-0.9-0.8)
<b>11</b>	69-84 (79)	40-80 (55)	60-85 (75)	0.5-0.7 (0.6)	0.7-1.5 (1.0)	0.5-1.2 (0.7)	<b>MLP 7-11-1</b> <b>81%-80%-85%</b> <b>(0.6-0.7-0.7)</b>
<b>12</b>	66-87 (79)	40-75 (55)	65-90 (75)	0.5-0.7 (0.6)	0.7-1.6 (1.0)	0.5-1.0 (0.7)	MLP 7-12-1 85%-75%-80% (0.5-1.4-1.0)
<b>14</b>	73-86 (80)	40-75 (55)	65-90 (75)	0.4-0.6 (0.6)	0.7-1.9 (1.1)	0.5-1.5 (0.7)	MLP 7-14-1 83%-70%-75% (0.6-1.0-0.9)

## ➤ ETAP II

Przedmiotem analiz w II etapie eksperymentów były sieci neuronowe o następującej architekturze:

- Wektor wejściowy „statystyczny 2”, 11 neuronów ukrytych.  
Analiza wrażliwości najlepszych uzyskanych dotychczas modeli nie wykazała obecności zbędnych zmiennych. W związku z tym nie dokonywano ręcznych modyfikacji wektora wejściowego. Również przetestowanie z pomocą *Automatycznego projektanta* 1000 sieci o podobnej architekturze nie doprowadziło do uzyskania modelu lepszego niż dotychczasowe, ani też nie doprowadziło do korzystnej redukcji wielkości którejkolwiek z warstw sieci. Próby ręcznego uczenia sieci przy pomocy *Edytora modelu* dostępnego w pakiecie *Statistica* nie doprowadziły do zmniejszenia błędów popełnianych przez sieć. Również ręczne usuwanie neuronów z warstwy ukrytej, mające na celu zmniejszenie rozmiarów sieci, powodowało obniżenie jakości uczenia, walidacji i testowania. Biorąc pod uwagę wyniki przeprowadzonych eksperymentów, za najlepsze uznano modele uzyskane w etapie I, mimo ich znacznych rozmiarów. Informacje o najefektywniejszych sieciach zestawiono w tabeli 42.

**Tabela 42. Najlepsze uzyskane modele dokonujące predykcji nasilenia niedrożności nosa po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 7-11-1	astma, niedrożność, kaszel, zmęczenie, refluks, wydzielina, ból głowy	86%-75%-80%	0.5-0.9-0.8	nie	nie
MLP 7-11-1	j.w.	81%-80%-85%	0.6-0.7-0.7	nie	nie

### 6.4.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa NIEDR\_bin\_12)

Prowadząc obliczenia dla grupy II, wykorzystywano podzbiory uczący, walidacyjny i testowy liczące odpowiednio 49, 12 i 10 przypadków. Na etapie poszukiwania wektora wejściowego napotkano istotny problem, ponieważ wyniki analiz statystycznych nie pozwoliły na wskazanie zmiennych wejściowych, od których statystycznie istotnie zależne byłyby zmienne wyjściowe NIEDROŻNOŚĆ\_12 i NIEDR\_bin\_12. Jedynie dla zmiennej „czas objawów” uzyskano wynik bliski istotności statystycznej (tab. 38 i 39). Dlatego budując wektory wejściowe określane jako „statystyczne” wykorzystano zestawy zmiennych jak dla grupy I,

dołączając dodatkowo zmienną „czas objawów”. Dla tych wektorów nie uzyskano jednak zadowalających wyników modelowania. Jakości testowania większości modeli osiągały maksymalnie 50%, co oznacza, że efektywność dokonywanych predykcji nie była lepsza niż w przypadku rzutu monetą. Znacznie lepsze modele zbudowano wykorzystując wektor wygenerowany dzięki algorytmom doboru cech (metoda krokowa postępująca). Najefektywniejsza okazała się sieć neuronowa o jakości testowania 70%, posiadająca 3 neurony w warstwie ukrytej. Dla tej sieci uzyskano zarówno optymalne wartości parametru  $\rho$ , jak i dobre wyniki analizy wrażliwości. Wadą tego modelu, której nie udało się skorygować w toku dalszych eksperymentów, były stosunkowo wysokie błędy walidacji i testowania, mogące świadczyć o niedostatecznej zdolności sieci do generalizacji. Informacje na temat najlepszego modelu zestawiono w tabeli 43.

**Tabela 43. Najlepsze uzyskane modele dokonujące predykcji nasilenia niedrożności nosa po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 6-3-1	skrzywienie przegrody, hist-pat, czas objawów, niedrożność, ból twarzy, węch	88%-75%-70%	0.4-2.1-2.6	nie	tak

### **6.4.3 Podsumowanie**

Modele neuronowe dokonujące predykcji nasilenia niedrożności nosa po zabiegu w I grupie pacjentów osiągnęły jakości testowania rzędu 80-85% (jakkolwiek szczegółowa analiza działania sieci sugeruje, że wartości te mogą być nieco zawyżone). Oznacza to jednak w pewnym uproszczeniu, że na podstawie dokonanej przez model predykcji będzie można poinformować pacjenta, czy jego nos po zabiegu będzie drożny, czy też nie, a prognozy te będą prawidłowe u znacznej większości chorych. Dla porównania warto ponownie zaznaczyć, że na podstawie samej tylko analizy danych historycznych można przewidywać uzyskanie prawidłowej drożności nosa u 42% chorych, jednak nie da się orzec, którzy z pacjentów mogą oczekiwać dobrego wyniku, a którzy nie. Dla mniej licznej II grupy pacjentów ponownie uzyskano mniej sprawne modele, których jakość testowania osiągała maksymalnie 70%. Analiza danych historycznych w tej grupie pozwoliłaby przewidywać całkowite udrożnienie nosa u 30% chorych.



## 6.5 Wydzielina z nosa

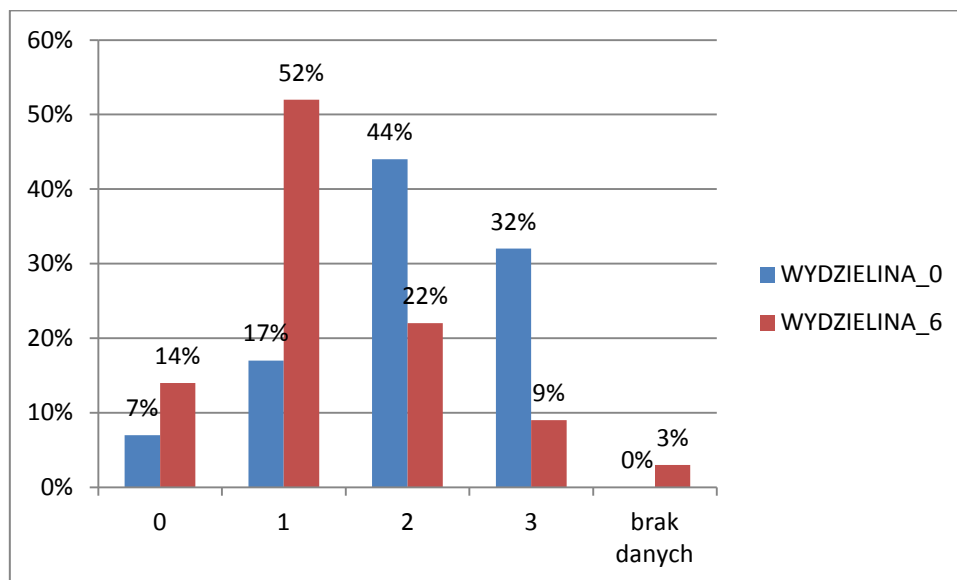
Drugą z dolegliwości typowych dla PZZP, której nasilenie pacjenci oceniali w skali 4-stopniowej, była obecność wydzieliny z nosa, opisana przez następujące zmienne:

- WYDZIELINA\_0 – uciążliwość związana z obecnością wydzieliny z nosa w skali 4-stopniowej przed zabiegiem,
- WYDZIELINA\_6 – uciążliwość związana z obecnością wydzieliny z nosa w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- WYDZIELINA\_12 – uciążliwość związana z obecnością wydzieliny z nosa w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

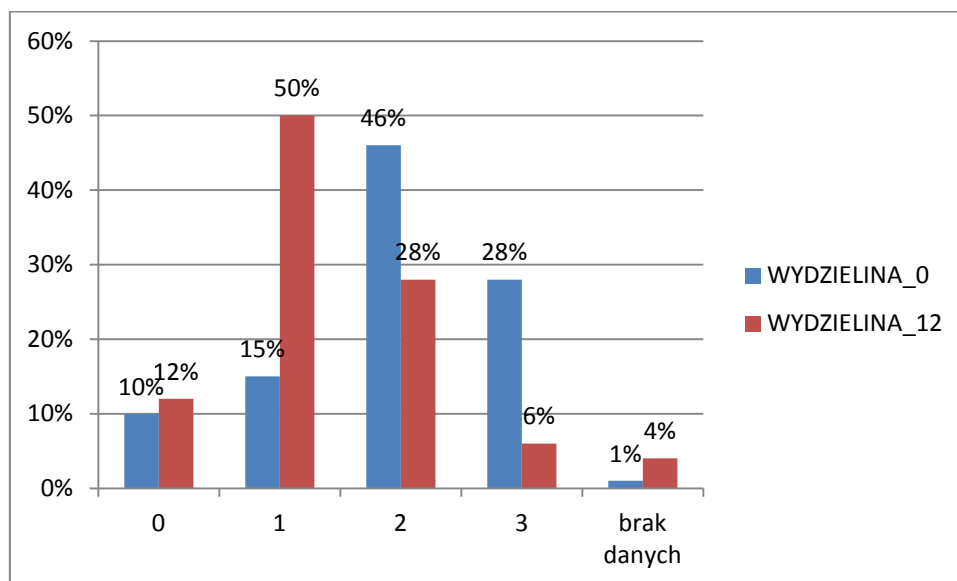
### 6.5.1 Analiza statystyczna

Wartości zmiennej WYDZIELINA zgłaszane przez pacjentów przed i po zabiegu przedstawiono na rysunkach 34 i 35, natomiast w tabeli 44 zestawiono zależności między zmienną WYDZIELINA a zmiennymi wejściowymi modeli neuronowych.

**Rysunek 34. Wartości zmiennej WYDZIELINA w grupie I – przed zabiegiem (WYDZIELINA\_0) i 3-6 miesięcy po zabiegu (WYDZIELINA\_6)**



**Rysunek 35. Wartości zmiennej WYDZIELINA w grupie II – przed zabiegiem (WYDZIELINA\_0) i 12 miesięcy po zabiegu (WYDZIELINA\_12)**



**Tabela 44. Zależności między ilością wydzieliny z nosa po operacji (wyrażoną w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	WYDZIELINA_6 (grupa I)	WYDZIELINA_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	(blisko granicy istotności) p = 0.056057
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	p = 0.006165	-
alergia	-	-
refluks	-	-

Cecha	WYDZIELINA_6 (grupa I)	WYDZIELINA_12 (grupa II)
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.045330	p = 0.037363
liczba przebytych ESS	R = 0.205323 p = 0.016896	R = 0.281118 p = 0.019290
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	R = 0.294445 p = 0.014054
niedrożność nosa	-	-
wydzielina z nosa	(blisko granicy istotności) R = 0.147725 p = 0.087294	-
splywanie wydzieliny	-	-
kaszel	-	R = 0.296641 p = 0.014787
kichanie	-	-
ból twarzy	-	-
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	-
zaburzenia węchu	-	-
zmęczenie	-	-
nieprzyjemny zapach z ust	-	R = 0.322610 p = 0.007293
ból zębów	-	-

Zmienną WYDZIELINA przekształcono w zmienną binarną WYDZ\_bin. Dla zmiennej WYDZIELINA zaobserwowano odmienny rozkład wartości podawanych przez pacjentów po zabiegu niż w przypadku niedrożności nosa (zmiennej NIEDROŻNOŚĆ). Całkowite ustąpienie dolegliwości związanych z obecnością wydzieliny z nosa zgłosiło jedynie kilkanaście procent pacjentów. Z tego względu binaryzacja analogiczna jak w przypadku niedrożności nosa prowadziłyby do znaczącej nierównowagi w zbiorze danych. Po uwzględnieniu braków danych, w grupie I proporcje liczby pacjentów zgłaszających

kompletne ustąpienie dolegliwości do pozostałych chorych wynosiły 14%:86%, a w grupie II 13%:87%.

Zdecydowano w związku z tym dokonać binaryzacji następująco:

- jeżeli  $WYDZIELINA \leq 1$  to  $WYDZ\_bin = 0$  (tzn. po zabiegu brak wydzieliny z nosa lub jej obecność jest mało uciążliwa),
- jeżeli  $WYDZIELINA \geq 2$  to  $WYDZ\_bin = 1$  (tzn. po zabiegu obecność wydzieliny z nosa jest średnio lub bardzo uciążliwa).

Przy tak przyjętym progu binaryzacji (oraz po uwzględnieniu braków danych) proporcje pacjentów, dla których zmienna  $WYDZ\_bin$  przyjmowała wartość 0, do chorych, u których zanotowano wartość 1, wynosiły odpowiednio: w grupie I 68%:32%, a w grupie II 65%:35%.

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (oceniaanej 3-6 miesięcy po zabiegu):  $WYDZ\_bin\_6$ ,
- dla grupy II (oceniaanej 12 miesięcy po zabiegu):  $WYDZ\_bin\_12$ .

**Model przewidujący wartość zmiennej  $WYDZ\_bin$  odpowiada na pytanie pacjenta: „Czy po zabiegu wydzielina z nosa będzie nieobecna lub jej obecność będzie mało uciążliwa?”**

W tabeli 45 przedstawiono wyniki analiz statystycznych dotyczących zależności między zmienną  $WYDZ\_bin$  a zmiennymi wejściowymi.

**Tabela 45. Zależności pomiędzy ilością wydzieliny z nosa po operacji wyrażoną za pomocą zmiennej  $WYDZ\_bin$  a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	$WYDZ\_bin\_6$ (grupa I)	$WYDZ\_bin\_12$ (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	p = 0.02903
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	p = 0.02632	-
alergia	-	-

Cecha	WYDZ_bin_6 (grupa I)	WYDZ_bin_12 (grupa II)
refluks	-	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	(blisko granicy istotności) p = 0.06485
przebyte ESS	p = 0.04027	-
liczba przebytych ESS	p = 0.015817	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	p = 0.009624
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	-	-
kaszel	-	p = 0.019587
kichanie	-	-
ból twarzy	-	-
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	-
zaburzenia węchu	(blisko granicy istotności) p = 0.069624	-
zmęczenie	-	-
nieprzyjemny zapach z ust	-	p = 0.041304
ból zębów	-	-

## **6.5.2 Modele neuronowe przewidujące ilość wydzieliny z nosa po zabiegu operacyjnym**

### **6.5.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa WYDZ\_bin\_6)**

#### **➤ ETAP I:**

#### **(a) Poszukiwanie wektora wejściowego**

Wyniki uzyskane przez sieci neuronowe dla różnych wektorów wejściowych przedstawiono w tabeli 46. Korzystano z podzbiorów: uczącego, walidacyjnego i testowego liczących odpowiednio 96, 19 i 20 przypadków. Dla kolejnych testowanych wektorów wejściowych uzyskano następujące wyniki:

- Wszystkie dostępne zmienne (29 zmiennych).  
**Wynik testu:** Zaobserwowano efekty typowe dla tak dużego wektora wejściowego, tj. wysoką jakość uczenia oraz bardzo niską jakość działania sieci podczas testowania – w praktyce nie lepszą niż w przypadku losowania wyników.
- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 3 zmienne, od których istotnie statystycznie zależna była zmienna WYDZ\_bin\_6 (AIA, liczba ESS, węch).
  - Wektor „statystyczny 2” – 4 zmienne, od których istotnie statystycznie zależne były zmienne WYDZ\_bin\_6 lub WYDZIELINA\_6 (AIA, liczba ESS, węch, wydzielina).
  - Wektor „statystyczny 3” – 8 zmiennych, od których istotnie statystycznie zależne były zmienne WYDZ\_bin\_6, WYDZ\_bin\_12, WYDZIELINA\_6 lub WYDZIELINA\_12 (AIA, liczba ESS, węch, wydzielina, skrzywienie przegrody, VAS\_0, kaszel, fetor).
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 7 zmiennych (hist-pat, ból ucha, węch, fetor, AIA, refluks, ESS)
  - Wektor „krokowa wsteczna” – 16 zmiennych (masywne polipy, TK, VAS\_0, niedrożność, spływanie, ból twarzy, ból czoła, ból ucha, węch, fetor, ból zębów, AIA, alergia, refluks, polipektomie, ESS)
  - Wektor „algorytm genetyczny” – 15 zmiennych (wiek, polipy, masywne polipy, TK, zabieg, czas objawów, VAS\_0, ból czoła, ból głowy, ból ucha, węch, fetor, AIA, alergia, ESS)

**Wyniki testów:** W przypadku wektorów wejściowych złożonych z 3-4 zmiennych modele były zbyt małe, aby uzyskać zadowalającą jakość uczenia. Zwiększanie wektora wejściowego powodowało poprawę jakości uczenia, ale wydaje się, że dodatkowe zmienne nie wносиły informacji istotnej do prawidłowego rozwiązania zadanego problemu, ponieważ błędy walidacji i testowania rosły przy tym zamiast ulegać zmniejszeniu. Przepuszczalnie wykorzystanie dodatkowych zmiennych powodowało jedynie zwiększenie rozmiarów sieci, które mogły dzięki temu lepiej dopasować się do nieistotnych szczegółów charakteryzujących zbiór uczący, ale nie wpływało to pozytywnie na modelowanie rzeczywistej zależności. Na skutek tego nie uzyskano tą metodą oczekiwanej poprawy jakości działania sieci ustalonej podczas walidacji i testowania. Otrzymane na tym etapie wyniki budzą podejrzenie, że dostępny zestaw danych uczących w ogóle nie zawiera informacji potrzebnych do tego, żeby prawidłowo przewidywać rozważaną zmienną wyjściową WYDZ\_bin\_6.

**Tabela 46. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wszystkie dostępne zmienne (29)	91-100 (99)	31-74 (0.6)	25-60 (40)	0.00008 -0.3 (0.01)	1.1-7.5 (2.7)	1.8-8.5 (3.9)	MLP 27-11-1 100%-73%-45% 0.08-2.1-2.8
wektor „statystyczny 1” (3)	51-68 (61)	37-79 (63)	40-75 (60)	0.7-0.8 (0.7)	0.5-0.6 (0.6)	0.7-0.9 (0.7)	MLP 3-7-1 68%-79%-75% (0.7-0.5-0.8)
wektor „statystyczny 2” (4)	52-68 (58)	37-79 (73)	40-75 (55)	0.7-0.9 (0.7)	0.5-0.6 (0.5)	0.6-1.0 (0.8)	MLP 4-5-1 64%-79%-75% (0.7-0.5-0.7)
wektor „statystyczny 3” (8)	68-88 (76)	31-68 (47)	35-70 (50)	0.5-0.7 (0.6)	0.7-1.6 (1.2)	0.7-1.9 (1.3)	MLP 8-10-1 79%-63%-65% (0.5-1.4-1.5)
wektor „krokowa postępująca” (7)	71-83 (78)	47-79 (68)	35-65 (50)	0.5-0.6 (0.6)	0.6-1.3 (0.8)	0.8-1.4 (1.1)	MLP 7-11-1 80%-73%-65% (0.5-0.9-1.1)
wektor „krokowa wsteczna” (16)	83-100 (93)	42-79 (58)	25-70 (45)	0.07-0.4 (0.2)	1.4-4.3 (2.3)	1.4-4.1 (2.2)	MLP 16-11-1 95%-74%-65% (0.1-2.1-2.2)
wektor „algorytm genetyczny” (15)	85-100 (94)	37-74 (58)	25-65 (45)	0.007-0.4 (0.2)	1.0-3.6 (2.2)	1.7-5.9 (2.7)	MLP 15-11-1 93%-74%-60% (0.2-1.0-2.3)

## **(b) Ocena liczby neuronów ukrytych**

Aby bardziej szczegółowo ocenić możliwość przewidywania ilości wydzieliny z nosa po zabiegu operacyjnym na podstawie dostępnych danych, zdecydowano przeanalizować działanie sieci neuronowych wykorzystujących wektor wejściowy „statystyczny 2” o większej niż wcześniej testowana liczbie neuronów ukrytych oraz wektory „krokowa postępująca” i „krokowa wsteczna” o zmniejszonej liczbie neuronów ukrytych. Uzyskane wyniki zestawiono w tabelach 47, 48 i 49.

Dla wektora „statystyczny 2” (tab. 47) zwiększanie liczby neuronów ukrytych nie poprawiło działania sieci. Najlepsze wyniki uzyskiwały sieci neuronowe wykorzystujące 3 zmienne wejściowe: węch, AIA i liczba ESS. Niezależnie od wielkości warstwy ukrytej uzyskiwano dla nich jakość testowania dochodzącą do 75%. Ponieważ jednak jakość uczenia nie przekraczała przy tym 68%, należy założyć, że wartość uzyskana dla zbioru testowego mogła być zawyżona i wynikać z przypadkowego nagromadzenia w nim wielu łatwych przypadków. Oznacza to, że 3 zmienne, od których statystycznie istotnie zależna była zmienna przewidywana, mają największe znaczenie dla modelowanej zależności, ale niosą co najwyżej tyle informacji, aby poprawnie przewidywać wyniki u około 70% pacjentów. Ponadto wątpliwości co do rzeczywistej efektywności uzyskanych modeli budzi fakt, że uzyskane sieci wykazywały tendencję do przewidywania wyłącznie wartości 0 w zbiorze testowym, a dla najlepszej sieci (o 8 neuronach ukrytych) w analizie wrażliwości dla wszystkich zmiennych uzyskano ilorazy poniżej 1.

**Tabela 47. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 2”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość wali-dacji [%]	Jakość testowa-nia [%]	Błąd uczenia	Błąd wali-dacji	Błąd testo-wania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
6	51-68 (58)	58-84 (73)	40-75 (55)	0.7-0.8 (0.7)	0.5-0.6 (0.6)	0.6-0.9 (0.8)	MLP 3-6-1 68%-78%-75% (0.7-0.5-0.8)
8	53-68 (58)	47-79 (73)	40-75 (55)	0.7-0.8 (0.7)	0.5-0.6 (0.5)	0.6-0.9 (0.8)	<b>MLP 3-8-1</b> <b>68%-79%-</b> <b>75%</b> <b>(0.7-0.5-0.8)</b>
10	53-63 (58)	31-74 (68)	35-65 (55)	0.7-0.8 (0.7)	0.5-0.6 (0.5)	0.6-0.9 (0.8)	MLP 3-10-1 63%-63%-60% (0.8-0.5-0.9)



Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
12	54-67 (57)	31-79 (68)	40-75 (55)	0.7-0.8 (0.7)	0.5-0.6 (0.5)	0.6-0.9 (0.8)	MLP 3-12-1 67%-79%-75% (0.7-0.5-0.8)
14	54-68 (58)	42-79 (74)	40-75 (55)	0.7-0.8 (0.7)	0.5-0.6 (0.5)	0.6-0.9 (0.8)	MLP 3-14-1 68%-79%-75% (0.7-0.5-0.8)

Dla wektora „krokowa postępująca” (tab. 48) uzyskano sieci o lepszych jakościach uczenia, jednak jakości działania sieci ustalone podczas testowania nie przekroczyły 70%. Modele, dla których uzyskano najlepsze wyniki, wykazywały ponadto stosunkowo wysokie błędy walidacji i testowania.

**Tabela 48. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „krokowa postępująca”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	55-72 (63)	47-74 (63)	20-50 (35)	0.6-0.7 (0.7)	0.5-0.7 (0.6)	0.7-1.0 (0.8)	MLP 7-2-1 71%-73%-50% (0.7-0.5-1.0)
6	63-82 (73)	47-79 (63)	30-60 (45)	0.6-0.7 (0.6)	0.5-1.2 (0.7)	0.7-1.4 (1.0)	MLP 7-6-1 76%-68%-60% (0.6-0.9-1.2)
10	69-87 (79)	52-84 (68)	30-70 (50)	0.5-0.6 (0.6)	0.5-1.2 (0.8)	0.8-1.5 (1.1)	<b>MLP 7-10-1</b> <b>79%-74%-70%</b> <b>(0.5-1.0-1.1)</b>

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
12	69-86 (78)	47-79 (68)	35-70 (50)	0.5-0.6 (0.6)	0.6-1.3 (0.8)	0.8-1.5 (1.1)	MLP 7-12-1 80%-74%-65% (0.6-0.9-1.2)

Dla bardzo dużego wektora „krokowa wsteczna” obserwowano tendencję do przeuczania modelu nawet przy bardzo małych rozmiarach warstwy ukrytej (tab. 49). Wysokie błędy walidacji i testowania w porównaniu z błędami uczenia uzyskano już dla sieci liczącej 2 neurony ukryte, chociaż dla tej sieci wartość parametru  $\rho$  wyniosła aż 3.1, co sugeruje zbyt małe rozmiary modelu i znajduje odzwierciedlenie w jego niskiej jakości uczenia. Dla żadnego z modeli nie uzyskano jakości testowania powyżej 65%.

**Tabela 49. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „krokowa wsteczna”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	41-82 (65)	42-79 (58)	20-65 (40)	0.6-0.8 (0.6)	0.5-1.3 (0.8)	0.7-2.1 (1.1)	MLP 13-2-1 65%-79%-60% (0.6-1.0-1.0)
6	73-97 (86)	31-79 (63)	25-65 (45)	0.2-0.6 (0.4)	0.6-2.2 (1.4)	0.8-3.4 (1.8)	MLP 13-6-1 80%-79%-60% (0.5-1.1-1.8)
10	58-88 (72)	47-79 (63)	30-65 (48)	0.5-0.7 (0.6)	0.5-1.0 (0.7)	0.8-1.8 (1.1)	MLP 8-10-1 76%-79%-60% (0.6-0.7-0.9)

## ➤ ETAP II

Przy pomocy *Automatycznego projektanta* wygenerowano zbiory po 1000 sieci neuronowych o następujących architekturach:

- Wektor wejściowy „statystyczny 2”, 8 neuronów ukrytych.  
Najlepsza z uzyskanych sieci (MLP 3-8-1) wykorzystywała zmienne węch, AIA i liczba ESS. Posiadała jakości uczenia, walidacji i testowania 68%-79%-75% i błędy 0.7-0.5-0.8. Podobnie jak poprzednie modele o tej architekturze, dla zbioru testowego sieć przewidywała wyłącznie wartość zmiennej wyjściowej równą 0, a analiza wrażliwości wykazała, że dwie z trzech zmiennych wejściowych potencjalnie pogarszają działanie sieci. W związku z tym model trudno bezwzględnie uznać za wiarygodny. Informacje o tej sieci – jako najlepszej z uzyskanych w toku eksperymentów - zestawiono w tabeli 50.
- Wektor „krokowa postępująca”, 10-11 neuronów ukrytych.  
Najlepszy z uzyskanych modeli wykazał jakość testowania 65%, a próby poprawy jego działania poprzez modyfikacje architektury i ponowne uczenie sieci nie spowodowały poprawy tego wyniku.

**Tabela 50. Najlepszy z uzyskanych modeli dokonujących predykcji ilości wydzieliny z nosa po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 3-8-1	węch, AIA, liczba ESS	68%-79%-75%	0.7-0.5-0.8	tak	nie

### 6.5.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa WYDZ\_bin\_12)

Dla grupy II liczby przypadków w podzbiorach uczącym, walidacyjnym i testowym wynosiły odpowiednio 48 : 11 : 10. Zdecydowanie najlepsze wyniki uzyskano wykorzystując wektory wejściowe opracowane przy pomocy algorytmów doboru cech (zwłaszcza metody krokowej wstecznej). W omawianym przypadku algorytmy te okazały się znacznie skuteczniejsze niż próby wyboru zmiennych wejściowych na podstawie analiz statystycznych. Najlepsze modele liczyły od 3 do 5 neuronów w warstwie ukrytej. Ostatecznie w II etapie eksperymentów uzyskano najefektywniejszy model, którego opis przedstawiono w tabeli 51.

**Tabela 51. Najlepszy z uzyskanych modeli dokonujących predykcji ilości wydzieliny z nosa po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 5-5-1	pleć, fetor, astma, refluks, ESS	83%-82%-70%	0.6-0.6-0.9	nie	nie

### **6.5.3 Podsumowanie**

Wyniki uzyskane przez sieci neuronowe dokonujące predykcji ilości wydzieliny z nosa po zabiegu operacyjnym są zdecydowanie gorsze, niż opisane w poprzednich podrozdziałach wyniki zdolności przewidywania innych wskaźników określających stopień poprawy stanu zdrowia pacjentów. Jakość testowania większości modeli uzyskanych dla grupy I nie przekraczała 65%, a sieci uzyskujące 75% poprawnych odpowiedzi w zbiorze testowym wykazywały cechy sugerujące, że przy wyciąganiu ostatecznych wniosków dotyczących ich rzeczywistej efektywności, należy zachować szczególną ostrożność. Lepsze wyniki uzyskano dla grupy II, dla której zbudowano stosunkowo wiarygodne modele wykazujące jakość testowania 70%. Jednak przebieg eksperymentów także i w tym przypadku wskazywał, że dla rozważanej w tym podrozdziale zmiennej wyjściowej uzyskanie znacząco lepszych predykcji na podstawie wykorzystanych w niniejszej pracy danych klinicznych może nie być osiągalne.

Podsumowując należy stwierdzić, że uciążliwość związana z obecnością wydzieliny z nosa po zabiegu operacyjnym u chorych z PZZP okazała się zdecydowanie trudniejsza do przewidzenia (przynajmniej na podstawie dostępnych danych) niż ogólne nasilenie dolegliwości czy niedrożność nosa. Może to wynikać ze znacznego subiektywizmu oceny tej dolegliwości, bądź też z istnienia innych - poza rozważanymi w tej pracy - uwarunkowań jej nasilenia.

### **6.6 Splywanie wydzieliny po tylnej ścianie gardła**

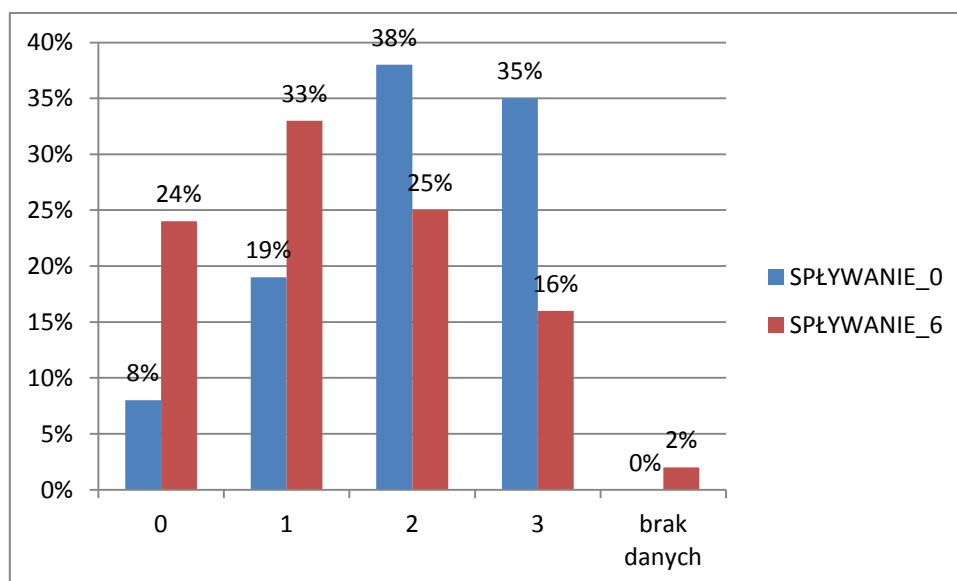
Uciążliwość kolejnego istotnego objawu PZZP, jakim jest splywanie wydzieliny po tylnej ścianie gardła, również była oceniana przez pacjentów w skali 4-stopniowej. Dla jej opisu przyjęto następujące nazwy zmiennych:

- SPŁYWANIE\_0 – uciążliwość sływania wydzieliny po tylnej ścianie gardła w skali 4-stopniowej przed zabiegiem,
- SPŁYWANIE\_6 – uciążliwość sływania wydzieliny po tylnej ścianie gardła w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- SPŁYWANIE\_12 – uciążliwość sływania wydzieliny po tylnej ścianie gardła w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

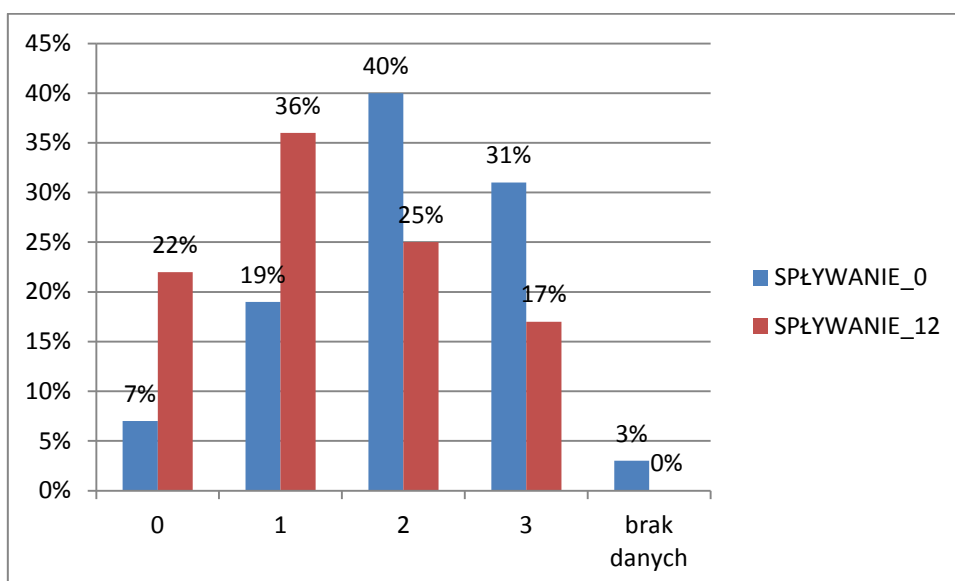
### 6.6.1 Analiza statystyczna

Analogicznie jak w poprzednich podrozdziałach, wartości zmiennej SPŁYWANIE zgłaszane przez pacjentów przed zabiegiem i po nim przedstawiono na rysunkach 36 i 37, a wyniki analiz statystycznych dotyczących zależności między zmienną SPŁYWANIE a zmiennymi wejściowymi budowanych modeli zebrano w tabeli 52. Zwraca uwagę fakt, że sływanie wydzieliny po tylnej ścianie gardła okazało się statystycznie istotnie zależne od znacznie większej liczby czynników niż omawiana wcześniej ilość wydzieliny z nosa.

**Rysunek 36. Wartości zmiennej SPŁYWANIE w grupie I – przed zabiegiem (SPŁYWANIE\_0) i 3-6 miesięcy po zabiegu (SPŁYWANIE\_6)**



**Rysunek 37. Wartości zmiennej SPŁYWANIE w grupie II – przed zabiegiem (SPŁYWANIE\_0) i 12 miesięcy po zabiegu (SPŁYWANIE\_12)**



**Tabela 52. Zależności między uciążliwością sływania wydzieliny po tylnej ścianie gardła po operacji (wyrażoną w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	SPŁYWANIE_6 (grupa I)	SPŁYWANIE_12 (grupa II)
pleć	-	p = 0.030427
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	R = 0.236389 p = 0.006156	R = 0.288577 p = 0.013958
astma	(blisko granicy istotności) p = 0.063758	(blisko granicy istotności) p = 0.074365
AIA	p = 0.005711	-
alergia	-	-
refluks	(blisko granicy istotności) p = 0.066959	-

Cecha	SPLYWANIE_6 (grupa I)	SPLYWANIE_12 (grupa II)
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.016635	-
liczba przebytych ESS	R = 0.224695 p = 0.008541	(blisko granicy istotności) R = 0.217174 p = 0.066882
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	R = 0.283112 p = 0.000918	R = 0.378425 p = 0.001047
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	R = 0.394315 p = 0.000002	R = 0.500571 p = 0.000010
kaszel	R = 0.218991 p = 0.010716	R = 0.458828 p = 0.000073
kichanie	-	-
ból twarzy	-	R = 0.278299 p = 0.019659
ból czoła	-	-
ból głowy	-	R = 0.460696 p = 0.000068
dolegliwości uszne	R = 0.339414 p = 0.000053	R = 0.402786 p = 0.000601
zaburzenia węchu	-	-
zmęczenie	R = 0.343352 p = 0.000052	R = 0.445645 p = 0.000098
nieprzyjemny zapach z ust	R = 0.249578 p = 0.003387	R = 0.277281 p = 0.019230
ból zębów	R = 0.189994 p = 0.026728	-

Podobnie jak w przypadku ilości wydzieliny z nosa, odsetek pacjentów podających całkowite ustąpienie splywania po tylnej ścianie gardła był stosunkowo niewielki (24% w grupie I i

22% w grupie II). Z tego powodu zmienną SPŁYWANIE przekształcono w zmienną binarną SPŁYW\_bin w taki sam sposób, jak w przypadku omawianej poprzednio zmiennej WYDZIELINA, tj:

- jeżeli  $SPŁYWANIE \leq 1$  to  $SPŁYW\_bin = 0$  (tzn. po zabiegu splotowanie wydzieliny po tylnej ścianie gardła jest niezauważalne lub mało uciążliwe),
- jeżeli  $SPŁYWANIE \geq 2$  to  $SPŁYW\_bin = 1$  (tzn. po zabiegu splotowanie wydzieliny po tylnej ścianie gardła jest średnio lub bardzo uciążliwe).

Przy tak przyjętej definicji (i po uwzględnieniu braków danych) proporcje pacjentów, dla których wartość zmiennej SPŁYW\_bin wynosiła 0 do pozostałych chorych wynosiły 58%:42% w obu badanych grupach.

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (oceniaanej 3-6 miesięcy po zabiegu): SPŁYW\_bin\_6,
- dla grupy II (oceniaanej 12 miesięcy po zabiegu): SPŁYW\_bin\_12.

**Model przewidujący wartość zmiennej SPŁYW\_bin odpowiada na pytanie pacjenta:  
„Czy po zabiegu splotowanie wydzieliny po tylnej ścianie gardła ustąpi lub będzie mało uciążliwe?”**

**Tabela 53. Zależności pomiędzy uciążliwością splotowania wydzieliny po tylnej ścianie gardła po operacji wyrażoną za pomocą zmiennej SPŁYW\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	SPŁYW_bin_6 (grupa I)	SPŁYW_bin_12 (grupa II)
pleć	-	p = 0.03710
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	p = 0.025850	p = 0.046875
astma	-	-
AIA	p=0.01096	-
alergia	-	-
refluks	-	-
przebyte polipektomie	-	-



Cecha	SPLYW_bin_6 (grupa I)	SPLYW_bin_12 (grupa II)
liczba przebytych polipektomii	-	-
przebyte ESS	p=0.02424	-
liczba przebytych ESS	-	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	p = 0.010721	p = 0.000769
niedrożność nosa	-	-
wydzielina z nosa	-	-
spływanie wydzieliny	p = 0.000678	p = 0.000022
kaszel	(blisko granicy istotności) p = 0.067096	p = 0.000084
kichanie	-	-
ból twarzy	-	p = 0.003489
ból czoła	-	(blisko granicy istotności) p = 0.065589
ból głowy	-	p = 0.000666
dolegliwości uszne	p = 0.010198	p = 0.000972
zaburzenia węchu	-	-
zmęczenie	p = 0.000881	p = 0.000257
nieprzyjemny zapach z ust	(blisko granicy istotności) p = 0.058190	-
ból zębów	p = 0.039643	(blisko granicy istotności) p = 0.077811

## **6.6.2 Modele neuronowe przewidujące ilość wydzieliny z nosa po zabiegu operacyjnym**

### **6.6.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa SPŁYW\_bin\_6)**

#### **➤ ETAP I:**

#### **(a) Poszukiwanie wektora wejściowego**

W trakcie eksperymentów dotyczących poszukiwania najbardziej obiecującego wektora wejściowego korzystano z podziału zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 96:20:20 przypadków. Informacje o wynikach testów krótko opisano poniżej i zestawiono w tabeli 54.

- Wszystkie dostępne zmienne (29 zmiennych).  
**Wynik testu:** Tak jak w przypadku poprzednio omawianych zmiennych obserwowano przecuczenie modeli.
- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 8 zmiennych, od których istotnie statystycznie zależna była zmienna SPŁYW\_bin\_6 (por. tab. 53).  
**Wynik testu:** Mimo stosunkowo niewielkich rozmiarów wektora wejściowego modele cechowały się słabą zdolnością do generalizacji. W związku z tym zdecydowano zmniejszyć liczbę zmiennych wejściowych eliminując trzy z nich, dla których wartość poziomu istotności p była najwyższa (por. tab. 53). W ten sposób otrzymano kolejny wektor wejściowy opisany poniżej.
  - Wektor „statystyczny 2” – 5 zmiennych (astma, VAS\_1, spływanie, ból ucha, zmęczenie).  
**Wynik testu:** Mimo redukcji wielkości wektora wejściowego utrzymywały się wysokie błędy walidacji i testowania.
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 8 zmiennych (czas objawów, VAS\_1, spływanie, ból ucha, zmęczenie, ból zębów, AIA, alergia).  
**Wynik testu:** Po raz kolejny stwierdzono wysokie błędy walidacji i testowania dla wszystkich otrzymanych sieci neuronowych.
  - Wektor „krokowa wsteczna” – 6 zmiennych (czas objawów, spływanie, kaszel, ból ucha, zmęczenie, ból zębów).  
**Wynik testu:** Mimo redukcji błędów walidacji, błędy testowania pozostały stosunkowo wysokie.
  - Wektor „algorytm genetyczny”- 6 zmiennych (czas objawów, spływanie, zmęczenie, ból zębów, AIA, refluks).  
**Wynik testu:** Dla tego wektora uzyskano najkorzystniejsze wyniki dla zbioru walidacyjnego, jednak w dalszym ciągu nie uległa dostatecznej poprawie jakość testowania ani błęd testowania.

**Tabela 54. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
<b>wszystkie dostępne zmienne (29)</b>	95-100 (100)	40-70 (55)	40-70 (55)	0.00001 -0.5 (0.006)	1.4-8.9 (3.8)	3.6-12 (6.0)	MLP 23-11-1 97%-60%-60% (0.09-2.6-5.2)

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 1” (8)	78-91 (85)	50-75 (60)	40-60 (50)	0.4-0.6 (0.5)	1.4-3.0 (2.0)	1.7-4.0 (2.3)	MLP 8-10-1 86%-65%-60% (0.4-2.0-2.5)
wektor „statystyczny 2” (5)	67-81 (73)	45-60 (50)	50-65 (55)	0.6-0.7 (0.6)	1.5-2.1 (1.8)	1.5-2.3 (2.0)	MLP 5-9-1 73%-60%-60% (0.6-2.1-2.0)
wektor „krokowa postępująca” (8)	77-92 (86)	45-70 (60)	35-65 (50)	0.4-0.6 (0.5)	1.0-2.1 (1.4)	1.0-3.0 (2.0)	MLP 8-8-1 84%-70%-60% (0.6-1.2-1.6)
wektor „krokowa wsteczna” (6)	71-84 (79)	50-85 (65)	35-60 (45)	0.5-0.6 (0.5)	0.5-1.2 (0.8)	1.2-1.8 (1.4)	MLP 6-9-1 80%-70%-60% (0.5-0.9-1.4)
wektor „algorytm genetyczny” (6)	77-90 (85)	55-80 (70)	45-60 (55)	0.5-0.7 (0.6)	0.5-1.0 (0.6)	1.0-1.9 (1.5)	<b>MLP 6-10-1</b> <b>83%-75%-60%</b> <b>(0.5-0.8-1.3)</b>

Systematyczne uzyskiwanie niezadowolających wyników testowania przez wszystkie budowane modele neuronowe sugeruje, że dla zmiennej SPŁYW\_bin\_6 w zbiorze testowym znalazły się wyjątkowo trudne przypadki. Powzięto zatem podejrzenie, że zbiór ten może być niereprezentatywny dla całej populacji. Aby sprawdzić tę hipotezę, porównano wyniki uzyskane dla stosowanego stale podziału na podzbiory uczący, walidacyjny i testowy z wynikami dla podziałów losowych. Przeprowadzono eksperyment, w którym przy pomocy *Automatycznego projektanta sieci* zbudowano 100 modeli wykorzystujących wektor wejściowy „algorytm genetyczny”, ale dla każdego z nich oddzielnie dokonywano rozlosowania przypadków do podzbiorów uczącego, walidacyjnego i testowego w proporcjach 96:20:20. Wyniki zestawiono w tabeli 55. Oczywiście w wyniku losowania uzyskiwano zarówno bardzo łatwe, jak i bardzo trudne zbiory testowe, stąd jakości testowania wahały się w szerokich granicach od 35% do 90%. Stwierdzono jednak, że mediana jakości testowania modeli wykorzystujących losowe podziały na podzbiory jest o około 10% wyższa

niż w eksperymentach z wykorzystaniem podziału dokonanego według analizy skupień. Można stąd wnioskować, że rzeczywista efektywność opisywanych w tym podrozdziale modeli jest wyższa, niż wskazują ich wyniki uzyskiwane dla wykorzystywanego w tej pracy zbioru testowego.

**Tabela 55. Porównanie wyników uzyskiwanych przez sieci neuronowe wykorzystujące wektor wejściowy „algorytm genetyczny” przy podziale na podzbiory uczący, walidacyjny i testowy wyodrębnione na podstawie analizy skupień oraz losowo (oddzielnie dla każdej sieci).**

Sposób podziału na podzbiory uczący, walidacyjny i testowy	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wg analizy skupień	77-90 (85)	55-80 (70)	45-60 (55)	0.5-0.7 (0.6)	0.5-1.0 (0.6)	1.0-1.9 (1.5)	<b>MLP 6-10-1</b> <b>83%-75%-60%</b> <b>(0.5-0.8-1.3)</b>
losowo (oddzielnie dla każdej sieci)	71-91 (81)	45-85 (65)	35-90 (65)	0.4-0.8 (0.6)	0.5-2.7 (1.0)	0.5-2.2 (1.0)	MLP 6-10-1 73%-75%-75% (0.7-0.6-0.7)

### **(b) Ocena liczby neuronów ukrytych**

Ocenę działania sieci w zależności od wielkości warstwy ukrytej przeprowadzono dla najbardziej obiecującego zestawu zmiennych wejściowych, tj. wektora „algorytm genetyczny”. Najlepszy model uzyskany dotychczas dla tego wektora posiadał 10 neuronów w warstwie ukrytej. Optymalną wartość parametru  $\rho$  uzyskuje się natomiast dla 6 neuronów w tej warstwie. Z tego powodu zdecydowano przetestować modele o 4, 6 i 8 neuronach w rozważanej warstwie. Ponieważ jednak eksperymenty wykazały, że lepsze wyniki uzyskuje się dla mniejszych sieci, przetestowano dodatkowo sieci posiadające 1 lub 2 neurony w warstwie ukrytej. Ostatecznie najlepszym uzyskanym modelem był perceptron o strukturze MLP 6-2-1. Dla rozważanego zadania nie udało się jednak uzyskać jakości testowania powyżej 65% w odniesieniu do zbioru testowego wyodrębnionego na podstawie analizy skupień. Wyniki testów przedstawiono szczegółowo w tabeli 56.

**Tabela 56. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „algorytm genetyczny”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
1	40-80 (79)	35-85 (70)	35-65 (60)	0.7-0.8 (0.7)	0.5-0.8 (0.6)	0.7-1.2 (1.0)	MLP 6-1-1 nr 7 77%-70%-65% (0.7-0.7-1.0)
2	71-83 (78)	45-80 (65)	45-65 (60)	0.6-0.8 (0.7)	0.5-0.9 (0.6)	1.0-1.5 (1.1)	<b>MLP 6-2-1</b> <b>nr 86</b> <b>80%-70%-65%</b> <b>(0.7-0.6-1.0)</b>
4	74-88 (81)	55-80 (70)	45-65 (55)	0.6-0.7 (0.6)	0.5-1.0 (0.6)	1.0-1.6 (1.2)	MLP 6-4-1 nr 37 79%-70%-65% (0.7-0.7-1.2)
6	75-88 (84)	60-85 (70)	45-60 (55)	0.5-0.7 (0.6)	0.5-1.1 (0.6)	0.9-1.8 (1.3)	MLP 6-6-1 nr 79 85%-75%-60% (0.6-0.6-1.3)
8	77-89 (84)	60-85 (70)	45-60 (55)	0.5-0.7 (0.6)	0.5-1.1 (0.6)	1.0-1.9 (1.3)	MLP 6-8-1 nr 96 84%-80%-60% (0.6-0.5-1.2)

## ➤ ETAP II

W drugim etapie eksperymentów testowano sieci neuronowe o następujących architekturach:

- Wektor wejściowy „algorytm genetyczny”, 2 neurony ukryte.  
Najlepszą z 1000 sieci wygenerowanych przez *Automatycznego projektanta* był perceptron MLP 5-2-1 o jakościach uczenia, walidacji i testowania odpowiednio 78%-75%-65% i błędach 0.7-0.4-1.3. Nie uzyskano zatem poprawy predykcji dla zbioru testowego.

Dla porównania przeprowadzono analogiczny eksperyment, w którym zastosowano jednokrotny (przykładowy) podział losowy na podzbiory uczący, walidacyjny i testowy w tych samych proporcjach, jakich użyto wcześniej. Dla tak wyodrębnionego zbioru testowego uzyskano 70% poprawnych predykcji, co ponownie sugeruje, że wartości uzyskiwane dla standardowo stosowanego zbioru testowego mogą być niedoszacowane.

Następnie przeprowadzono eksperymenty polegające na usuwaniu z wektora wejściowego tych zmiennych, które na podstawie analizy wrażliwości można było uznać za pogarszające działanie modelu. W efekcie doszło jednak do dalszego pogorszenia jakości testowania.

- Wektor wejściowy „algorytm genetyczny”, 6 neuronów ukrytych.

Mimo mało obiecujących wyników uzyskanych dla tej architektury w pierwszym etapie badań, zdecydowano o ponownym jej przetestowaniu ze względu na uzyskiwaną dla niej optymalną wartość parametru  $\rho$ . Wygenerowanie 1000 modeli o tej architekturze przez *Automatycznego projektanta* pozwoliło uzyskać najlepszy model o najwyższej jakości testowania, wynoszącej 70%. Jego szczegółową charakterystykę przedstawiono w tabeli 57.

Ponieważ analiza wrażliwości sugerowała, że zmienna AIA może pogarszać działanie modelu, przeprowadzono również testy sieci neuronowych wykorzystującej wektor wejściowy „algorytm genetyczny” bez tej zmiennej, jednak ta hipoteza robocza nie znalazła potwierdzenia w przeprowadzonych badaniach, ponieważ maksymalna jakość testowania dla modeli z usuniętą zmienną AIA wynosiła 60%.

**Tabela 57. Najlepszy z uzyskanych modeli dokonujących predykcji uciążliwości sływania wydzieliny po tylnej ścianie gardła po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 5-6-1	czas objawów, sływanie, zmęczenie, AIA, refluks	75%-75%-70%	0.7-0.5-1.2	tak	tak

### 6.6.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa SPŁYW\_bin\_12)

Dla grupy II korzystano z podziału zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 49:12:11 przypadków. Należy zauważyć, że dla tej grupy chorych stwierdzono wysoce istotne statystyczne zależności między licznymi zmiennymi wejściowymi a zmienną wyjściową (por. tab. 53). Najlepsze modele neuronowe zbudowano korzystając z wektora

wejściowego złożonego z tych 6 zmiennych, od których zmienna wyjściowa była statystycznie istotnie zależna i poziom istotności statystycznej  $p$  wynosił poniżej 0.01. Warto zauważyć, że wybrane zmienne opisywały wyłącznie subiektywne nasilenie i rodzaj objawów zgłaszanych przez pacjentów przed operacją (VAS\_0, sływanie, kaszel, ból twarzy, ból głowy, ból ucha, zmęczenie). Wektor wejściowy nie zawierał natomiast danych o charakterze obiektywnym, takich jak wiek i płeć pacjenta, liczba przebytych zabiegów czy choroby towarzyszące. Kwestia ta zostanie szerzej przeanalizowana w rozdziale „Dyskusja”. Wyniki osiągane przez sieci neuronowe dla grupy II okazały się zdecydowanie lepsze niż dla grupy I. Najefektywniejszy model, scharakteryzowany szczegółowo w tabeli 58, osiągnął jakość testowania 91%.

**Tabela 58. Najlepszy z uzyskanych modeli dokonujących predykcji uciążliwości sływania wydzieliny po tylnej ścianie gardła po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 6-1-1	VAS, sływanie, kaszel, ból twarzy, ból głowy, ból ucha, zmęczenie	90%-83%-91%	0.3-0.9-0.6	tak	nie

### **6.6.3 Podsumowanie**

Podobnie jak w przypadku omawianej w poprzednim podrozdziale ilości wydzieliny z nosa - przewidywanie uciążliwości sływania wydzieliny po tylnej ścianie gardła w I grupie pacjentów okazało się bardzo trudne. Większość sieci nie osiągnęła jakości testowania powyżej 65%. W szczególności dla wektorów wejściowych zbudowanych na podstawie analiz statystycznych nie osiągnięto jakości testowania powyżej 60%, mimo że zmienna przewidywana była istotnie statystycznie zależna od licznych zmiennych wejściowych. Najlepszy model dokonywał prawidłowych predykcji u 70% pacjentów zaliczonych do zbioru testowego. Analizy z wykorzystaniem losowych podziałów na podzbiory uczący, walidacyjny i testowy sugerowały jednak, że wartości te można traktować jako zaniżone, zapewne o około 5-10%. Mimo to predykcje uciążliwości sływania wydzieliny po tylnej ścianie gardła w okresie 3-6 miesięcy trudno uznać za zadowalające. Najprawdopodobniej na nasilenie tej dolegliwości zgłaszane przez pacjentów na tak wczesnym etapie leczenia mają wpływ czynniki oddziałujące na jamę nosową i zatoki już po operacji. Można do nich zaliczyć na przykład jakość toalety jamy pooperacyjnej czy wystąpienie dodatkowych zakażeń w trakcie

gojenia. Z oczywistych przyczyn modele predykcyjne bazujące na informacjach sprzed zabiegu nie są w stanie uwzględnić tego rodzaju zmiennych.

Co zaskakujące, zdecydowanie lepsze wyniki uzyskano dla II grupy pacjentów. Jak się okazało, prawidłowe prognozowanie było możliwe u około 90% chorych. Taki wynik może oczywiście być skutkiem przypadkowego nagromadzenia łatwych do przewidzenia przypadków w tej grupie pacjentów. Tym niemniej istnieje sugestia, że w okresie 12 miesięcy po zabiegu subiektywne nasilenie badanego objawu jest zdecydowanie bardziej przewidywalne i mniej zależne od przypadkowych indywidualnych różnic w przebiegu pooperacyjnym niż ma to miejsce we wcześniejszym okresie po operacji. Zwraca uwagę fakt, że wysokiej jakości predykcje uzyskały modele korzystające praktycznie wyłącznie z danych na temat subiektywnego nasilenia kilku objawów, podanego przez pacjentów przed operacją, podczas gdy informacje o charakterze obiektywnym zostały usunięte z wektora wejściowego bez zmniejszenia skuteczności prognozowania.

## **6.7 Zaburzenia węchu**

Nasilenie zaburzeń węchu ocenianych przez pacjentów w skali 4-stopniowej opisano przy pomocy następujących zmiennych:

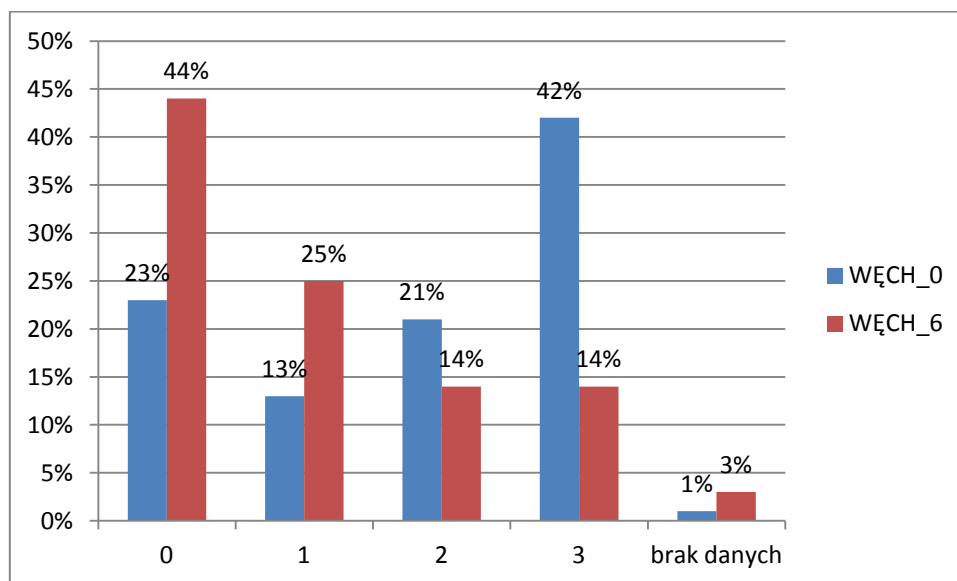
- WĘCH\_0 – nasilenie zaburzeń węchu w skali 4-stopniowej przed zabiegiem,
- WĘCH\_6 – nasilenie zaburzeń węchu w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- WĘCH\_12 – nasilenie zaburzeń węchu w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

### **6.7.1 Analiza statystyczna**

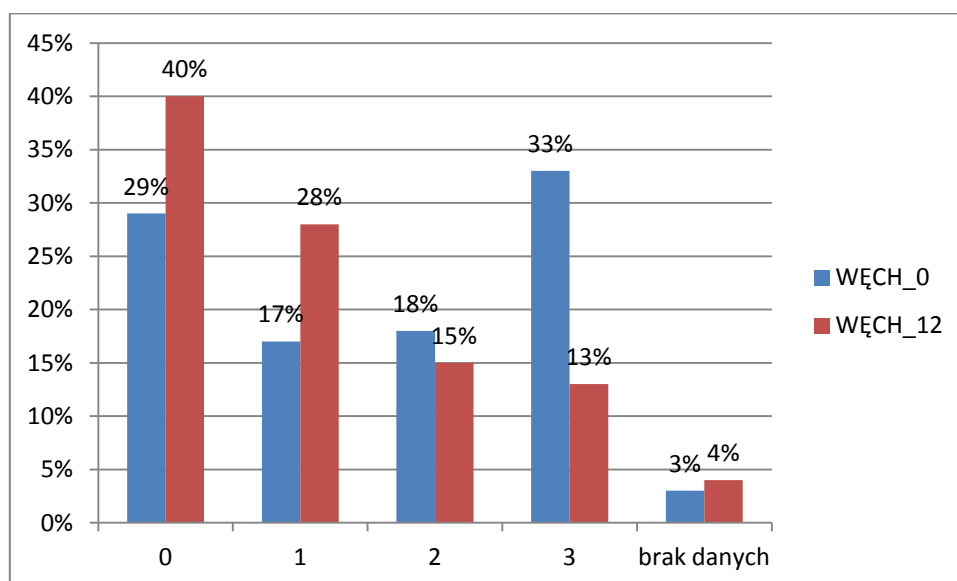
Wartości zmiennej WĘCH zanotowane przed zabiegiem i po zabiegu dla grupy I przedstawiono na rysunku 38, a dla grupy II na rysunku 39. Tabela 59 zawiera informacje o statystycznie istotnych zależnościach pomiędzy zmienną WĘCH a zmiennymi wejściowymi sieci neuronowych.



**Rysunek 38. Wartości zmiennej WĘCH w grupie I – przed zabiegiem (WĘCH\_0) i 3-6 miesięcy po zabiegu (WĘCH\_6)**



**Rysunek 39. Wartości zmiennej WĘCH w grupie II – przed zabiegiem (WĘCH\_0) i 12 miesięcy po zabiegu (WĘCH\_12)**



**Tabela 59. Zależności między nasileniem zaburzeń węchu po operacji (wyrażonym w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	WĘCH_6 (grupa I)	WĘCH_12 (grupa II)
pleć	-	-
wiek	R = 0.319256 p = 0.000161	R = 0.292458 p = 0.014747
polipy	p = 0.029406	p = 0.018811
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	R = 0.303962 p = 0.000395	-
astma	p = 0.000828	-
AIA	p = 0.000307	p = 0.012137
alergia	-	-
refluks	-	-
przebyte polipektomie	p = 0.028177	-
liczba przebytych polipektomii	R = 0.217263 p = 0.011367	-
przebyte ESS	p = 0.018058	-
liczba przebytych ESS	R = 0.231414 p = 0.006923	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	-	-
kaszel	-	-
kichanie	-	R = 0.248167 p = 0.042877
ból twarzy	-	-

Cecha	WEŁCH_6 (grupa I)	WEŁCH_12 (grupa II)
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	R = 0.243336 p = 0.047228
zaburzenia węchu	R = 0.537080 p < 0.000001	R = 0.460855 p = 0.000077
zmęczenie	-	-
nieprzyjemny zapach z ust	-	R = 0.337948 p = 0.004824
ból zębów	-	-

W grupie I (po uwzględnieniu braków danych) po zabiegu 45% pacjentów nie zgłaszało zaburzeń węchu. W grupie II prawidłowy węch po zabiegu uzyskało 42% pacjentów. W związku z tym zmienną WEŁCH przekształcono w zmienną binarną według następujących założeń:

- jeżeli WEŁCH = 0 to WEŁCH\_bin = 0 (tzn. po zabiegu brak zaburzeń węchu, co można interpretować jako normosmię),
- jeżeli WEŁCH > 0 to WEŁCH\_bin = 1 (tzn. po zabiegu utrzymują się mało, średnio lub bardzo nasilone zaburzenia węchu, czyli hiposmia lub anosmia).

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (ocenianej 3-6 miesięcy po zabiegu): WEŁCH\_bin\_6,
- dla grupy II (ocenianej 12 miesięcy po zabiegu): WEŁCH\_bin\_12.

**Model przewidujący wartość zmiennej WEŁCH\_bin odpowiada na pytanie pacjenta:  
„Czy po operacji będę miał prawidłowy węch?”**

**Tabela 60. Zależności pomiędzy nasileniem zaburzeń węchu po operacji wyrażonym za pomocą zmiennej WEŁCH\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	WEŁCH_bin_6 (grupa I)	WEŁCH_bin_12 (grupa II)
pleć	-	-

Cecha	WEŁCH_bin_6 (grupa I)	WEŁCH_bin_12 (grupa II)
wiek	p = 0.006725	p = 0.042963
polipy	(blisko granicy istotności) p=0.07021	p=0.02028
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	p = 0.007571	-
astma	p=0.00084	-
AIA	p=0.00188	(blisko granicy istotności) p=0.08692
alergia	-	-
refluks	-	-
przebyte polipektomie	(blisko granicy istotności) p=0.06940	-
liczba przebytych polipektomii	-	-
przebyte ESS	p=0.04333	-
liczba przebytych ESS	-	-
nasilenie dolegliwości przed zabiegiem:		
VAS_0	-	-
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	-	-
kaszel	-	-
kichanie	-	p = 0.012946
ból twarzy	-	-
ból czoła	-	-
ból głowy	-	-
dolegliwości uszne	-	-
zaburzenia węchu	p < 0.000001	p = 0.000118
zmęczenie	-	-

Cecha	WEŁCH_bin_6 (grupa I)	WEŁCH_bin_12 (grupa II)
nieprzyjemny zapach z ust	-	p = 0.004016
ból zębów	-	-

## 6.7.2 Modele neuronowe przewidujące ilość wydzieliny z nosa po zabiegu operacyjnym

### 6.7.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa WEŁCH\_bin\_6)

#### ➤ ETAP I:

#### (a) Poszukiwanie wektora wejściowego

Testując kolejne wektory wejściowe wykorzystywano podział zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 96:20:19. Ze względu na wyniki eksperymentów dotyczących poprzednio omawianych zmiennych wyjściowych zrezygnowano z testowania wektora wejściowego złożonego z 29 zmiennych, ponieważ jego użycie nieodmiennie prowadziło do tworzenia modeli o zbyt dużej liczbie połączeń w stosunku do ilości przypadków uczących. W związku z tym testowano następujące wektory wejściowe (por. tab. 61):

- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 8 zmiennych, od których istotnie statystycznie zależna była zmienna WEŁCH\_bin\_6 lub wynik był bliski granicy istotności statystycznej (por. tab. 60).  
**Wynik testu:** Uzyskano modele o stosunkowo wysokich jakościach uczenia, walidacji i testowania, wykazujące jednak niewielką tendencję do przeuczenia się.
  - Wektor „statystyczny 2” – 6 zmiennych, od których zmienna przewidywana była statystycznie istotnie zależna (por. tab. 60).  
**Wynik testu:** Usunięcie z wektora wejściowego 2 zmiennych pozwoliło uzyskać sieci neuronowe o niższych błędach walidacji niż w przypadku wektora „statystyczny 1”.
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 7 zmiennych (hist-pat, czas objawów, wdech, zmęczenie, astma, polipektomie, ESS).  
**Wynik testu:** Modele wykorzystujące ten wektor wejściowy wykazywały gorszą zdolność do generalizacji niż omawiane wcześniej.
  - Wektor „krokowa wsteczna” – 6 zmiennych (hist-pat, czas objawów, niedrożność, ból ucha, wdech, zmęczenie).  
**Wynik testu:** Sieci neuronowe wykazywały stosunkowo wysokie błędy testowania.

- Wektor „algorytm genetyczny” – 7 zmiennych (ból ucha, węż, astma, alergia, liczba polipektomii, polipektomie, ESS).

**Wynik testu:** Nie uzyskano tak wysokich jakości testowania, jak dla poprzednio testowanych wektorów.

**Tabela 61. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 1” (8)	78-92 (88)	55-80 (65)	58-85 (79)	0.4-0.6 (0.5)	0.8-3.3 (1.4)	1.0-2.5 (1.3)	MLP 8-9-1 88%-80%-79% (0.5-0.8-1.4)
wektor „statystyczny 2” (6)	75-85 (80)	60-80 (70)	63-84 (74)	0.5-0.7 (0.6)	0.7-0.9 (0.7)	1.1-1.7 (1.3)	<b>MLP 6-8-1</b> <b>81%-80%-84%</b> <b>(0.6-0.7-1.4)</b>
wektor „krokowa postępująca” (7)	76-90 (85)	45-75 (60)	52-79 (68)	0.4-0.6 (0.5)	0.7-1.4 (1.0)	1.1-2.2 (1.6)	MLP 7-10-1 91%-75%-73% (0.5-0.8-2.0)
wektor „krokowa wsteczna” (6)	83-94 (88)	60-80 (70)	63-84 (74)	0.4-0.5 (0.4)	0.8-1.3 (1.0)	1.2-2.8 (2.2)	MLP 6-12-1 88%-75%-79% (0.5-1.0-2.2)
wektor „algorytm genetyczny” (7)	77-90 (85)	50-80 (60)	47-79 (63)	0.4-0.6 (0.5)	0.9-2.2 (1.4)	0.8-1.4 (1.0)	MLP 5-9-1 79%-80%-74% (0.6-1.0-1.0)

### **(b) Ocena liczby neuronów ukrytych**

W kolejnym etapie eksperymentów wykorzystano wektor wejściowy „statystyczny 2”, który uznano za najbardziej obiecujący. Wyniki testów sieci neuronowych posiadających różną

liczbę neuronów w warstwie ukrytej przedstawiono w tabeli 62. Najlepsze modele posiadały 6 neuronów w warstwie ukrytej. Zauważono przy tym, że większość najskuteczniejszych sieci nie wykorzystuje całego wektora „statystyczny 2” (tj. wszystkich 6 zmiennych), a jedynie od 2 do 5 zmiennych wejściowych. To spostrzeżenie potraktowano jako punkt wyjścia dla drugiego etapu eksperymentów.

**Tabela 62. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 2”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
<b>4</b>	73-84 (77)	50-80 (65)	57-89 (79)	0.6-0.7 (0.6)	0.6-1.1 (0.7)	0.9-1.5 (1.3)	MLP 4-4-1 79%-80%-84% (0.6-0.7-1.2)
<b>6</b>	73-84 (77)	50-80 (70)	63-84 (79)	0.6-0.7 (0.6)	0.6-0.9 (0.7)	1.1-1.9 (1.3)	<b>MLP 5-6-1</b> <b>81%-80%-84%</b> <b>(0.6-0.8-1.1)</b>
<b>8</b>	73-84 (78)	60-80 (70)	63-89 (79)	0.5-0.7 (0.6)	0.6-0.9 (0.7)	1.1-1.5 (1.3)	MLP 4-8-1 80%-80%-84% (0.6-0.7-1.2)
<b>10</b>	73-86 (78)	55-80 (70)	63-89 (74)	0.6-0.7 (0.6)	0.6-0.9 (0.7)	1.2-1.6 (1.3)	MLP 4-10-1 80%-80%-84% (0.6-0.7-1.4)

## ➤ ETAP II

W drugim etapie eksperymentów przetestowano modele wykorzystujące od 2 do 5 zmiennych wejściowych, które zostały wyselekcjonowane przez *Automatycznego projektanta sieci* w trakcie poszukiwań optymalnej liczby neuronów ukrytych. Oceniano w związku z tym sieci o następujących architekturach:

- Wektor wejściowy złożony z 2 zmiennych (wiek, astma), 4-8 neuronów ukrytych. Najlepsze z modeli uzyskiwały jakość uczenia 62% i jakość testowania zaledwie 52%.
- Wektor wejściowy złożony z 3 zmiennych (wiek, astma, AIA), 4-8 neuronów ukrytych.

Ponownie nie uzyskano modeli o jakości testowania przekraczającej 53%.

- Wektor wejściowy złożony z 4 zmiennych (wiek, czas objawów, wdech, astma), 4-8 neuronów ukrytych.  
Najlepszy z modeli osiągnął jakości uczenia, walidacji i testowania odpowiednio 79%-80%-84%. Błędy uczenia, walidacji i testowania wynosiły 0.7-0.8-1.2. Ze względu na utrzymujące się nieproporcjonalnie wysokie błędy testowania przeprowadzono analogiczny test korzystając z losowych podziałów na podzbiory uczący, walidacyjny i testowy. Potwierdzono, że mediana błędów testowania popełnianych przez modele wykorzystujące wylosowane zbiory testowe wynosi 0.9, podczas gdy dla używanego standardowo zbioru testowego mediana błędów jest równa 1.2. W odniesieniu do jakości testowania nie stwierdzono tak dużych różnic. Przyjęto w związku z tym, że błąd testowania dla zbioru testowego wyodrębnionego na podstawie analizy skupień może być nieco zawyżony. Zdecydowano zatem, że błędy testowania w niewielkim stopniu przekraczające błędy walidacji i uczenia nie będą stanowiły bezwzględnego kryterium uznania modeli za nieefektywne.
- Wektor wejściowy złożony z 5 zmiennych (wiek, czas objawów, wdech, astma, AIA), 4-8 neuronów ukrytych.  
Dla tego wektora wejściowego uzyskano najlepszy model predykcyjny, którego charakterystykę przedstawiono w tabeli 63. Podjęto również próbę zbudowania przy pomocy *Automatycznego projektanta* 1000 sieci o takiej samej architekturze (5 zmiennych wejściowych i 6 neuronów w warstwie ukrytej), jednak żadna z uzyskanych sieci nie okazała się lepsza od opisanej wcześniej. Nie uzyskano również sieci neuronowej, dla której analiza wrażliwości nie wykazywałaby obecności potencjalnie zbędnych zmiennych. Sugeruje to, że w rzeczywistości wszystkie wejścia sieci były istotne dla działania modelu, być może ze względu na istniejące między nimi interakcje, których nie uwzględnia analiza wrażliwości.

**Tabela 63. Najlepszy z uzyskanych modeli dokonujących predykcji zaburzeń węchu po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 5-6-1	wiek, czas objawów, wdech, astma, AIA	83%-80%-84%	0.6-0.8-1.2	tak	tak

### 6.7.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa WĘCH\_bin\_12)

W analizach dotyczących zmiennej wyjściowej WĘCH\_bin\_12 w II grupie pacjentów proporcje przypadków w podzbiórach uczącym, walidacyjnym i testowym wynosiły 47:12:10.



Niemal wszystkie metody selekcji zmiennych wejściowych (tj. zarówno metody statystyczne, jak i algorytmy doboru cech) prowadziły do wyodrębnienia bardzo podobnych wektorów wejściowych. Najlepsze wyniki uzyskano dla wektora wejściowego uzyskanego dzięki zastosowaniu metody krokowej postępującej. Liczył on 7 zmiennych (wiek, hist-pat, czas objawów, wydzielina, kichanie, węch, zmęczenie). Przy pomocy *Automatycznego projektanta* uzyskano liczne sieci neuronowe o jakości testowania około 80%. Niektóre z modeli uzyskiwały jakość testowania 90%, jednak w tych przypadkach jakości uczenia i walidacji nie przekraczały 83%, więc wartość 90% należy przypuszczalnie uznać za nieco zawyżoną. Najlepszą sieć scharakteryzowano w tabeli 64.

**Tabela 64. Najlepszy z uzyskanych modeli dokonujących predykcji zaburzeń węchu po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 6-2-1	wiek, hist-pat, wydzielina, kichanie, węch, zmęczenie	83%-83%-90%	0.6-0.5-0.4	tak	nie

### **6.7.3 Podsumowanie**

W obu analizowanych grupach pacjentów normosmię (prawidłowy węch) po zabiegu uzyskało nieco ponad 40% chorych. Zastosowanie sieci neuronowych pozwoliło prawidłowo przewidywać wynik operacji w odniesieniu do zmysłu powonienia u około 83% chorych. Oznacza to, że zastosowanie neuronowych modeli predykcyjnych pozwala prognozować efekty leczenia znacznie lepiej niż prosta analiza danych historycznych, zarówno pod względem jakościowym, jak i ilościowym. Dobre wyniki uzyskane dla modeli o różnych architekturach świadczą o tym, że funkcję węchu po operacji u większości chorych można uznać za stosunkowo łatwo przewidywalną, mimo że rokowanie dotyczące tego objawu po zabiegu jest powszechnie uważane za bardzo niepewne z powodu słabej podatności na leczenie oraz złożonych uwarunkowań analizowanych zaburzeń [3, 109].

## **6.8 Ból twarzy**

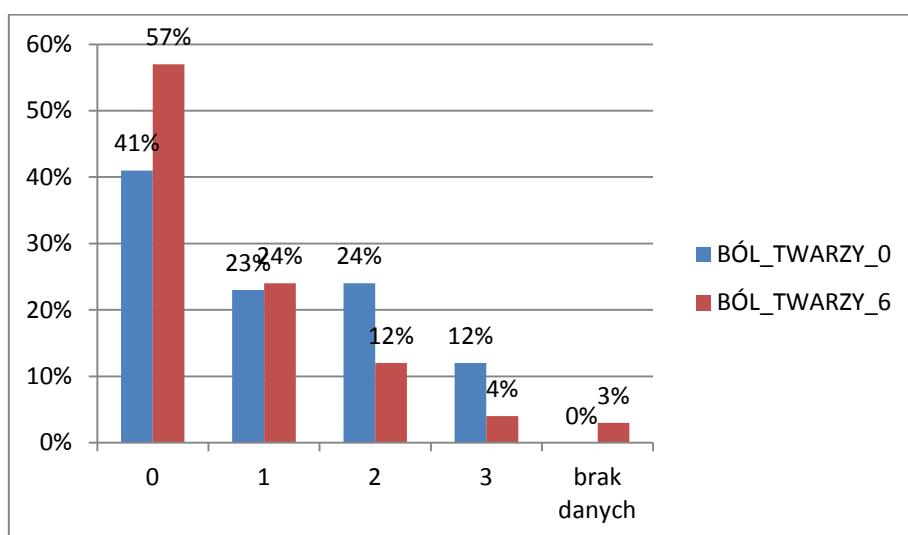
Ból twarzy, traktowany przez wielu chorych jako charakterystyczny objaw potwierdzający obecność stanu zapalnego zatok, był oceniany w skali 4-stopniowej. Podane przez pacjentów wartości zakodowano przy pomocy następujących zmiennych:

- BÓL\_TWARZY\_0 – nasilenie bólu twarzy w skali 4-stopniowej przed zabiegiem,
- BÓL\_TWARZY\_6 – nasilenie bólu twarzy w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- BÓL\_TWARZY\_12 – nasilenie bólu twarzy w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

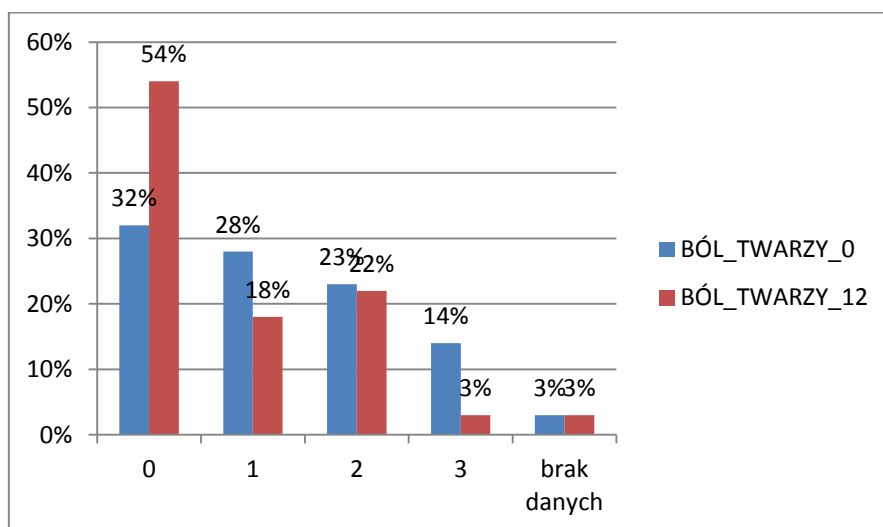
### 6.8.1 Analiza statystyczna

Wartości zmiennej BÓL\_TWARZY podane przez pacjentów z grup I i II przed zabiegiem oraz ich porównanie z wartościami zanotowanymi po zabiegu przedstawiono na rysunkach 40 i 41.

**Rysunek 40. Wartości zmiennej BÓL\_TWARZY w grupie I – przed zabiegiem (BÓL\_TWARZY\_0) i 3-6 miesięcy po zabiegu (BÓL\_TWARZY\_6)**



**Rysunek 41. Wartości zmiennej BÓL\_TWARZY w grupie II – przed zabiegiem (BÓL\_TWARZY\_0) i 12 miesięcy po zabiegu (BÓL\_TWARZY\_12)**



Analiza statystyczna zależności między nasileniem bólu twarzy a pozostałymi zmiennymi przedstawiona w tabeli 65 wykazała, że objaw ten często utrzymuje się u chorych skarżących się również na inne dolegliwości bólowe (ból głowy, ucha, zębów).

**Tabela 65. Zależności między nasileniem bólu twarzy po operacji (wyrażonym w skali 4-stopniowej), a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	BÓL_TWARZY_6 (grupa I)	BÓL_TWARZY_12 (grupa II)
pleć	-	(blisko granicy istotności) p = 0.076025
wiek	-	-
polipy	-	-
masywne polipy	p = 0.033905	(blisko granicy istotności) p = 0.078102
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	(blisko granicy istotności) R = -0.230918 p = 0.062120
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	p = 0.037666	-
alergia	-	-
refluks	(blisko granicy istotności) p = 0.065193	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	-	-
liczba przebytych ESS	-	R = 0.238010 p = 0.047246
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	-	-

Cecha	BÓL_TWARZY_6 (grupa I)	BÓL_TWARZY_12 (grupa II)
wydzielina z nosa	-	-
splywanie wydzieliny	-	(blisko granicy istotności) R = 0.221189 p = 0.069881
kaszel	-	R = 0.334706 p = 0.005273
kichanie	-	-
ból twarzy	R = 0.321652 p = 0.000142	R = 0.493884 p = 0.000016
ból czoła	R = 0.277101 p = 0.001190	R = 0.480604 p = 0.000044
ból głowy	R = 0.291916 p = 0.000651	R = 0.456249 p = 0.000092
dolegliwości uszne	R = 0.310287 p = 0.000250	R = 0.371218 p = 0.001984
zaburzenia węchu	-	-
zmęczenie	R = 0.313876 p = 0.000247	R = 0.488157 p = 0.000018
nieprzyjemny zapach z ust	-	R = 0.315491 p = 0.008276
ból zębów	R = 0.362699 p = 0.000015	R = 0.375263 p = 0.001487

Ze względu na fakt, że w każdej z grup po zabiegu około połowa chorych podawała całkowity brak bólu twarzy (po uwzględnieniu braków danych w grupie I było to 59% chorych, a w grupie II 56% chorych), zdecydowano przekształcić zmienną BÓL\_TWARZY w zmienną binarną BÓL\_TW\_bin w następujący sposób:

- jeżeli BÓL\_TWARZY = 0 to BÓL\_TW\_bin = 0 (tzn. po zabiegu ból twarzy nie występuje),
- jeżeli BÓL\_TWARZY > 0 to BÓL\_TW\_bin = 1 (tzn. po zabiegu występuje ból twarzy o małej, średniej lub dużej uciążliwości).

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (oceniaanej 3-6 miesięcy po zabiegu): BÓL\_TW\_bin\_6,
- dla grupy II (oceniaanej 12 miesięcy po zabiegu): BÓL\_TW\_bin\_12.

Model przewidujący wartość zmiennej BÓL\_TW\_bin odpowiada na pytanie pacjenta:  
*„Czy po operacji będę odczuwał ból twarzy?”*

Tabela 66. Zależności pomiędzy nasileniem zaburzeń węchu po operacji wyrażonym za pomocą zmiennej BÓL\_TW\_bin, a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.

Cecha	BÓL_TW_bin_6 (grupa I)	BÓL_TW_bin_12 (grupa II)
pleć	-	p=0.03718
wiek	-	-
polipy	p = 0.01585	-
masywne polipy	p=0.01094	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	-	-
AIA	-	-
alergia	-	-
refluks	p=0.03106	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	-	-
liczba przebytych ESS	-	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	-	-
kaszel	-	p = 0.008833
kichanie	-	-
ból twarzy	p = 0.001344	p = 0.000047
ból czoła	p = 0.005334	p = 0.000090

Cecha	BÓL_TW_bin_6 (grupa I)	BÓL_TW_bin_12 (grupa II)
ból głowy	p = 0.005828	p = 0.000159
dolegliwości uszne	p = 0.003369	p = 0.002139
zaburzenia węchu	-	-
zmęczenie	p = 0.001337	p = 0.000073
nieprzyjemny zapach z ust	-	p = 0.006174
ból zębów	p = 0.000372	p = 0.004584

## 6.8.2 Modele neuronowe przewidujące nasilenie bólu twarzy po zabiegu operacyjnym

### 6.8.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa BÓL\_TW\_bin\_6)

#### ➤ ETAP I:

#### (a) Poszukiwanie wektora wejściowego

Liczby przypadków w podzbiorach uczącym, walidacyjnym i testowym wynosiły odpowiednio 95:20:20. Oceniono następujące wektory wejściowe (tab. 67):

- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 9 zmiennych, od których statystycznie istotnie zależna była zmienna BÓL\_TW\_bin\_6 (por. tab. 66).  
**Wynik testu:** Sieci neuronowe wykorzystujące ten wektor wejściowy wykazywały niedostateczną zdolność do generalizacji. Z tego względu podjęto próby zmniejszenia liczby zmiennych wejściowych poprzez identyfikację i usunięcie zmiennych wzajemnie skorelowanych.
  - Wektor „statystyczny 2” – zestaw 8 zmiennych uzyskany po usunięciu z wektora „statystycznego 1” zmiennej „ból czoła” jako najsilniej skorelowanej z innymi zmiennymi wejściowymi („ból głowy” i „ból twarzy”), a słabiej od nich korelującej ze zmienną BÓL\_TWARZY\_6.  
**Wynik testu:** Zmniejszenie wektora wejściowego nie poprawiło zdolności sieci do generalizacji.
  - Wektor „statystyczny 3” – zestaw 8 zmiennych otrzymanych po usunięciu z wektora „statystycznego 1” zmiennej „masywne polipy” ze względu na jej silną korelację ze zmienną „polipy”.  
**Wynik testu:** Ponownie nie uzyskano poprawy efektywności modeli.
  - Wektor „statystyczny 4” – 7 zmiennych otrzymanych po usunięciu z poprzedniego wektora wejściowego zmiennej „ból zębów”, która w analizie wrażliwości dla większości modeli okazywała się najmniej istotna.  
**Wynik testu:** Uzyskano niewielką redukcję przeciętnych błędów walidacji i testowania bez znaczącej poprawy jakości predykcji.

- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca” – 5 zmiennych (ból czoła, ból głowy, ból ucha, ból zębów, refluks).  
**Wynik testu:** Warto zauważyć, że wektor metoda krokowa postępująca doprowadziła do wyodrębnienia tych samych zmiennych, co analizy statystyczne, jednak pozwoliła zredukować ich liczbę. Modele wykorzystujące ten wektor wejściowy wykazywały lepszą zdolność do generalizacji, jednak tak mała liczba zmiennych wejściowych okazała się niewystarczająca do osiągnięcia wysokiej jakości uczenia.
  - Wektor „krokowa wsteczna”- 8 zmiennych (polipy, hist-pat, ból twarzy, ból czoła, ból głowy, ból ucha, ból zębów, refluks).  
**Wynik testu:** Ponownie doszło do wyodrębnienia bardzo podobnego zestawu zmiennych jak w wektorach utworzonych na podstawie analiz statystycznych. Błędy testowania były jednak wyższe niż dla tej samej wielkości wektorów „statystycznych”.
  - Wektor „algorytm genetyczny” – 8 zmiennych (skrzywienie przegrody, sływanie, kichanie, ból głowy, zmęczenie, fetor, ból zębów, refluks).  
**Wynik testu:** Zastosowanie algorytmu genetycznego doprowadziło do wyboru zestawu zmiennych różniącego się od rozważanych uprzednio. Jakości uczenia i walidacji kazały się jednak gorsze niż w przypadku poprzednio testowanych wektorów.

**Tabela 67. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 1” (9)	85-99 (93)	45-80 (60)	50-80 (65)	0.3-0.5 (0.4)	0.9-2.3 (1.4)	1.0-2.6 (1.6)	MLP 9-11-1 93%-75%-70% (0.4-1.1-1.3)
wektor „statystyczny 2” (8)	82-96 (89)	35-70 (50)	45-70 (60)	0.3-0.5 (0.4)	1.1 – 2.6 (1.6)	0.9-2.2 (1.5)	MLP 8-11-1 86%-70%-60% (0.4-1.4-1.2)

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 3” (8)	83-98 (92)	45-75 (60)	50-80 (60)	0.2-0.5 (0.4)	1.0-3.5 (1.5)	1.2-3.3 (1.8)	MLP 8-10-1 91%-75%-70% (0.4-1.5-1.2)
wektor „statystyczny 4” (7)	79-98 (87)	45-75 (60)	50-80 (70)	0.2-0.6 (0.4)	0.8-3.0 (1.3)	0.9-3.7 (1.3)	MLP 7-11-1 89%-75%-70% (0.4-1.2-1.2)
wektor „krokowa postępująca” (5)	65-88 (81)	40-70 (55)	50-80 (65)	0.5-0.8 (0.6)	0.8-1.5 (1.0)	0.7-1.3 (1.0)	MLP 5-9-1 71%-70%-75% (0.6-0.8-0.8)
wektor „krokowa wsteczna” (8)	90-98 (95)	55-85 (70)	40-80 (60)	0.2-0.4 (0.3)	0.7-2.3 (1.4)	1.1-2.9 (1.9)	MLP 8-11-1 96%-85%-70% (0.4-0.9-1.8)
wektor „algorytm genetyczny” (7)	74-94 (83)	45-75 (56)	40-75 (60)	0.4-0.6 (0.5)	1.0-2.0 (1.4)	0.7-1.5 (0.9)	MLP 7-8-1 81%-70%-70% (0.5-1.0-0.8)

### **(b) Ocena liczby neuronów ukrytych**

Ponieważ prawie wszystkie testowane wektory wejściowe zawierały zestaw zmiennych bardzo podobny jak w wektorze „statystyczny 1”, zdecydowano przeprowadzić dalsze eksperymenty z użyciem tego właśnie wektora. W poprzednim etapie badań modele wykorzystujące wektor „statystyczny 1” cechowały się niewystarczającą zdolnością do generalizacji, jednak posiadały rozbudowaną warstwę ukrytą. Podjęto więc próbę zmniejszenia sieci i poprawy ich zdolności do generalizacji metodą redukcji liczby neuronów ukrytych. Wyniki zestawiono w tabeli 68. Zrezygnowano natomiast z testów z użyciem



wektora „krokowa postępująca”, ponieważ – mimo uzyskania dla niego sieci o najniższych błędach – wydawał się on nie zawierać dostatecznej ilości informacji, aby możliwe było uzyskanie dostatecznie wysokich jakości uczenia.

**Tabela 68. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „statystyczny 1”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	68-87 (76)	40-80 (60)	45-80 (65)	0.5-0.7 (0.7)	0.7-1.7 (1.0)	0.6-1.4 (0.9)	MLP 6-2-1 76%-75%-75% (0.7-0.9-0.7)
3	71-89 (82)	40-75 (60)	40-80 (65)	0.4-0.7 (0.6)	0.8-1.7 (1.1)	0.7-1.9 (1.2)	<b>MLP 9-3-1</b> <b>78%-75%-80%</b> <b>(0.5-0.8-0.9)</b>
4	75-93 (83)	45-75 (60)	45-80 (65)	0.4-0.6 (0.5)	0.8-1.8 (1.1)	0.7-2.1 (1.2)	MLP 9-4-1 83%-75%-75% (0.5-1.0-1.4)
6	79-97 (88)	45-75 (60)	40-80 (65)	0.3-0.6 (0.5)	0.9-2.7 (1.3)	0.9-2.2 (1.3)	MLP 9-6-1 88%-75%-75% (0.5-0.9-1.0)

## ➤ ETAP II

Po przeanalizowaniu wyników zestawionych w tabeli 68 zdecydowano w drugim etapie eksperymentów przetestować sieci o następującej architekturze:

- Wektor wejściowy „statystyczny 1”, 3 neurony ukryte.

Przy pomocy *Automatycznego projektanta* zbudowano 1000 sieci neuronowych o powyższej architekturze i uzyskano model o wyższych jakościach uczenia i walidacji niż w pierwszym etapie badań. Jego charakterystykę zawiera tabela 69.

Ponieważ analiza wrażliwości przeprowadzona dla tego modelu wykazała obecność potencjalnie zbędnych zmiennych wejściowych, przy pomocy *Automatycznego projektanta sieci* oraz *Edytora modelu* przetestowano również sieci neuronowe wykorzystujące wektor wejściowy zubożony o potencjalnie najmniej istotną zmienną

(ból zębów) - jednak doprowadziło to do pogorszenia predykcji.

Ponieważ parametr  $\rho$  przyjmuje optymalne wartości dla 4 neuronów ukrytych, podjęto również próbę zwiększenia rozmiarów warstwy ukrytej o 1 neuron, jednak w ten sposób również nie uzyskano równie efektywnych modeli jak opisany w tabeli 69.

**Tabela 69. Najlepszy z uzyskanych modeli dokonujących predykcji nasilenia bólu twarzy po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 9-3-1	polipy, masywne polipy, ból twarzy, ból głowy, ból czoła, ból ucha, zmęczenie, ból zębów, refluks	79%-80%-80%	0.6-0.7-0.9	tak	nie

### 6.8.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa BÓL\_TW\_bin\_12)

Jak wykazały analizy zestawione w tabeli 66, istnieje wysoce istotna statystycznie zależność między zmienną BÓL\_TW\_bin\_12 a kilkoma innymi zmiennymi, zwłaszcza opisującymi dolegliwości bólowe twarzy, czoła, głowy, ucha, zębów, a także zmęczenie i nieprzyjemny zapach z ust ( $p < 0.0001$ ). Poszukiwania najbardziej obiecującego wektora wejściowego prowadziły w większości przypadków do wyselekcjonowania tego właśnie zestawu zmiennych jako najbardziej korzystnego, niezależnie od zastosowanej metody doboru cech. Najlepszy z modeli predykcyjnym uzyskanych dla grupy II wykorzystywał 7 wymienionych wyżej zmiennych wejściowych oraz posiadał 6 neuronów w warstwie ukrytej. Charakterystykę tego modelu przedstawiono w tabeli 70. Mniejsze modele, dla których wartość parametru  $\rho$  mieściła się w optymalnych granicach, wykazywały niezadowalającą efektywność. Również próby modyfikacji wektora wejściowego przeprowadzone na podstawie analizy wrażliwości doprowadziły do pogorszenia działania sieci.

**Tabela 70. Najlepszy z uzyskanych modeli dokonujących predykcji nasilenia bólu twarzy po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 7-6-1	ból twarzy, ból czoła, ból głowy, ból ucha, ból zębów, zmęczenie, fetor	88%-92%-90%	0.5-0.8-0.6	tak	nie

### **6.8.3 Podsumowanie**

Analiza danych historycznych dotyczących bólu twarzy u chorych operowanych z powodu PZZP pozwalała przewidywać, że 3-6 miesięcy po operacji 59% pacjentów będzie wolnych od tej dolegliwości, a 12 miesięcy po operacji odsetek ten może być nieco niższy. Zastosowanie sztucznych sieci neuronowych pozwoliło dokonać prawidłowych predykcji nasilenia bólu twarzy u 80% chorych w okresie 3-6 miesięcy po operacji oraz u blisko 90% pacjentów 12 miesięcy po operacji.

W obu analizowanych grupach pacjentów stwierdzono, że istnieje pewien zestaw zmiennych wejściowych, który wykazuje wyraźny związek z rokowaniem dotyczącym bólu twarzy. Zmienne te opisywały przede wszystkim nasilenie dolegliwości bólowych twarzy, ale także czoła, głowy, ucha czy zębów, w okresie poprzedzającym zabieg operacyjny. Uzyskane wyniki sugerują po raz kolejny, że prognoza dotycząca nasilenia poszczególnych objawów PZZP może być dokonywana przede wszystkim na podstawie występowania u pacjentów pewnych charakterystycznych „konstelacji” dolegliwości przed zabiegiem (np. duże nasilenie niektórych objawów przy niewielkim nasileniu pozostałych), co najprawdopodobniej pozwala podzielić chorych na podgrupy różniące się między sobą przebiegiem choroby i rokowaniem. Ponadto szczegółowa analiza objawów klinicznych przypuszczalnie może dawać podstawy do zidentyfikowania tych chorych, u których ból twarzy ma podłoże pozarynologiczne i w związku z tym okazuje się niepodatny na leczenie operacyjne zatok przynosowych. Znaczenie tych spostrzeżeń zostanie szerzej omówione w rozdziale „Dyskusja”.

## 6.9 Ból głowy

Kolejną dolegliwością, dla której zbudowano modele predykcyjne, był ból głowy zlokalizowany w innych miejscach niż twarz lub czoło, czyli nie mający charakteru typowego bólu uznawanego powszechnie za „zatokopochodny”. Zdecydowano podjąć próbę prognozowania nasilenia także tej dolegliwości po zabiegu operacyjnym, ponieważ szczególnie interesująca wydawała się ocena, na ile jest ono przewidywalne w porównaniu z innymi objawami, wynikającymi bezpośrednio (w sposób przyczynowy) z obecności stanu zapalnego w jamie nosowej i zatokach przynosowych.

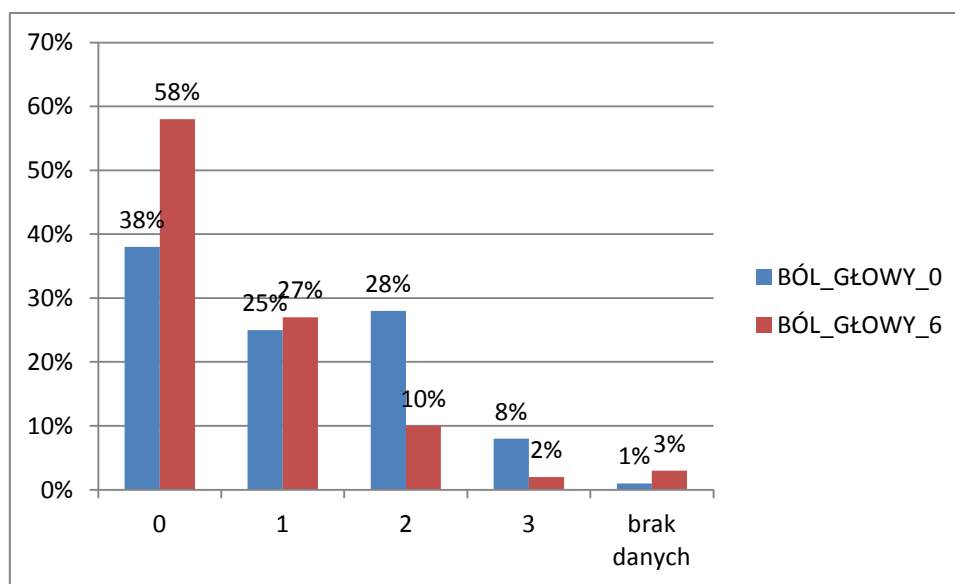
Przyjęto następujące nazwy zmiennych opisujących nasilenie bólu głowy:

- BÓL\_GŁOWY\_0 – nasilenie bólu głowy w skali 4-stopniowej przed zabiegiem,
- BÓL\_GŁOWY\_6 – nasilenie bólu głowy w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- BÓL\_GŁOWY\_12 – nasilenie bólu głowy w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

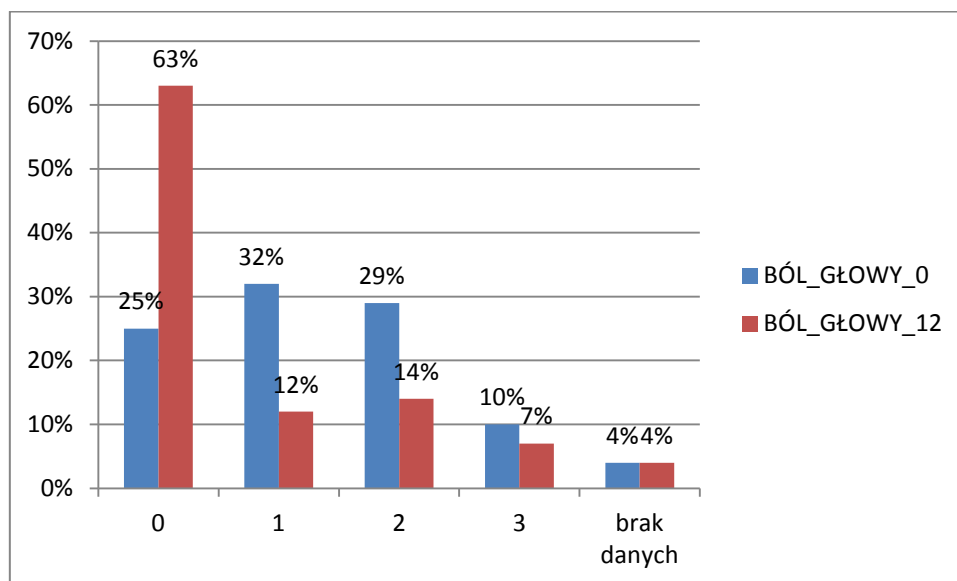
### 6.9.1 Analiza statystyczna

Na rysunkach 42 i 43 przedstawiono, w jakim stopniu operacja endoskopowa zatok przynosowych wywarła pozytywny wpływ na nasilenie bólu głowy w obu badanych grupach chorych. Wyniki analiz statystycznych zestawionych w tabeli 71 wykazały, że podobnie jak w przypadku bólu twarzy, utrzymującego się bólu głowy po zabiegu operacyjnym należy oczekiwać zwłaszcza u tych pacjentów, którzy już przed operacją skarżyli się na uciążliwe dolegliwości bólowe o różnej lokalizacji.

**Rysunek 42. Wartości zmiennej BÓL\_GŁOWY w grupie I – przed zabiegiem (BÓL\_GŁOWY\_0) i 3-6 miesięcy po zabiegu (BÓL\_GŁOWY\_6)**



**Rysunek 43. Wartości zmiennej BÓL\_GŁOWY w grupie II – przed zabiegiem (BÓL\_GŁOWY\_0) i 3-6 miesięcy po zabiegu (BÓL\_GŁOWY\_12)**



**Tabela 71. Zależności między nasileniem bólu głowy po operacji (wyrażonym w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	BÓL_GŁOWY_6 (grupa I)	BÓL_GŁOWY_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	R = 0.210441 p = 0.015437	-
astma	-	-

Cecha	BÓL_GŁOWY_6 (grupa I)	BÓL_GŁOWY_12 (grupa II)
AIA	-	(blisko granicy istotności) p = 0.059958
alergia	-	-
refluks	-	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	-	-
liczba przebytych ESS	-	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	R = 0.260063 p = 0.030923
niedrożność nosa	-	-
wydzielina z nosa	-	-
spływanie wydzieliny	R = 0.211540 p = 0.013779	-
kaszel	-	-
kichanie	R = 0.177870 p = 0.039770	-
ból twarzy	R = 0.287479 p = 0.000723	R = 0.365190 p = 0.002198
ból czoła	R = 0.422340 p < 0.000001	R = 0.461567 p = 0.000109
ból głowy	R = 0.469584 p < 0.000001	R = 0.556731 p = 0.000001
dolegliwości uszne	R = 0.242167 p = 0.004658	R = 0.273509 p = 0.026277
zaburzenia węchu	-	-
zmęczenie	R = 0.271035 p = 0.001670	R = 0.355446 p = 0.002725
nieprzyjemny zapach z ust	-	-
ból zębów	R = 0.245943 p = 0.004036	R = 0.251724 p = 0.038381

Po uwzględnieniu braków danych wyliczono, że w grupie I całkowity brak bólu głowy po zabiegu podało 59% pacjentów, a w grupie II 65% pacjentów. W związku z tym zmienną BÓL\_GŁOWY przekształcono w zmienną binarną BÓL\_GŁ\_bin zgodnie z następującymi założeniami:

- jeżeli BÓL\_GŁOWY = 0 to BÓL\_GŁ\_bin = 0 (tzn. po zabiegu ból głowy nie występuje),
- jeżeli BÓL\_GŁOWY > 1 to BÓL\_GŁ\_bin = 1 (tzn. po zabiegu ból głowy jest mało, średnio lub bardzo uciążliwy).

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (oceniaanej 3-6 miesięcy po zabiegu): BÓL\_GŁ\_bin\_6,
- dla grupy II (oceniaanej 12 miesięcy po zabiegu): BÓL\_GŁ\_bin\_12.

**Model przewidujący wartość zmiennej BÓL\_GŁ\_bin odpowiada na pytanie pacjenta:  
„Czy po zabiegu będę cierpiał z powodu bólu głowy?”**

**Tabela 72. Zależności pomiędzy nasileniem bólu głowy po operacji wyrażonym za pomocą zmiennej BÓL\_GŁ\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	BÓL_GŁ_bin_6 (grupa I)	BÓL_GŁ_bin_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	p = 0.03085	-
masywne polipy	p = 0.04934	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu		-
histopatologia – dominacja eozynofili	-	-
czas objawów	p = 0.029880	-
astma	-	-
AIA	-	-
alergia	-	-
refluks	-	-
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-

Cecha	BÓL_GŁ_bin_6 (grupa I)	BÓL_GŁ_bin_12 (grupa II)
przebyte ESS	-	-
liczba przebytych ESS	-	-
nasilenie dolegliwości przed zabiegiem:		
VAS_0	-	p = 0.024520
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	p = 0.038348	-
kaszel	-	-
kichanie	p = 0.039026	-
ból twarzy	p = 0.001899	p = 0.004348
ból czoła	p = 0.000001	p = 0.000212
ból głowy	p = 0.000002	p = 0.000018
dolegliwości uszne	p = 0.016892	p = 0.040807
zaburzenia węchu	-	-
zmęczenie	p = 0.008868	p = 0.009278
nieprzyjemny zapach z ust	-	-
ból zębów	p = 0.016789	(blisko granicy istotności) p = 0.073639

## 6.9.2 Modele neuronowe przewidujące nasilenie bólu głowy po zabiegu operacyjnym

### 6.9.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa BÓL\_GŁ\_bin\_6)

#### ➤ ETAP I:

#### (a) Poszukiwanie wektora wejściowego

Testy przeprowadzono wykorzystując podzbiory uczący, walidacyjny i testowy o liczebnościach odpowiednio 96:20:19. Oceniono następujące wektory wejściowe (tab. 73):

- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 11 zmiennych, od których statystycznie istotnie zależna była zmienna BÓL\_GŁ\_bin\_6 (por. tab. 72).  
**Wynik testu:** Mimo redukcji wektora wejściowego dokonanej przez *Automatycznego projektanta sieci*, modele nie wykazywały dostatecznej



zdolności do generalizacji, w związku z czym zdecydowano dokonać próby manualnego zmniejszenia liczby zmiennych wejściowych.

- Wektor „statystyczny 2” – zestaw 9 zmiennych uzyskany po usunięciu z wektora „statystycznego 1” zmiennych silnie skorelowanych z pozostałymi zmiennymi („ból twarzy” i „masywne polipy”).

**Wynik testu:** Zmniejszenie wektora wejściowego nie spowodowało oczekiwanej redukcji błędów walidacji i testowania.

- Wektor „statystyczny 3” – zestaw 7 zmiennych uzyskanych dzięki dalszej redukcji wektora „statystycznego 2” poprzez usunięcie zmiennych „polipy” oraz „spływanie”, nie wykorzystywanych przez najlepsze z modeli uzyskanych w poprzednim etapie badań.

**Wynik testu:** Dalsze zmniejszenie wektora wejściowego spowodowało pewne obniżenie błędów walidacji i testowania, jednak bez znaczącej poprawy ich jakości.

- Wektory uzyskane dzięki algorytmom doboru cech:

- Wektor „krokowa postępująca” – 3 zmienne (ból czoła, ból głowy, ból zębów).

**Wynik testu:** Wykorzystanie zaledwie 3 zmiennych wejściowych pozwoliło otrzymać modele o jakościach walidacji i testowania porównywalnych z modelami wykorzystującymi wektory „statystyczne” oraz zmniejszyć błędy walidacji i testowania, jednak poprawie zdolności do generalizacji uzyskanej dzięki zmniejszeniu modeli towarzyszyło obniżenie jakości uczenia do maksymalnie 75%.

- Wektor „krokowa wsteczna” – 11 zmiennych (płeć, polipy, skrzywienie przegrody, ból twarzy, ból czoła, ból głowy, ból ucha, węch, ból zębów, astma, alergia).

**Wynik testu:** Podobnie jak dla wektora wejściowego „statystyczny 1”, również liczącego 11 zmiennych, stwierdzono niedostateczną zdolność do generalizacji.

- Wektor „algorytm genetyczny” – 9 zmiennych (skrzywienie przegrody, kaszel, ból twarzy, ból czoła, ból głowy, ból ucha, węch, zmęczenie, ból zębów).

**Wynik testu:** Dla tego wektora uzyskano sieci neuronowe o stosunkowo wysokich jakościach uczenia i testowania oraz niskich błędach testowania w porównaniu z modelami wykorzystującymi omawiane wcześniej wektory wejściowe.

**Tabela 73. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 1” (11)	80-91 (85)	45-80 (65)	42-79 (65)	0.1-0.6 (0.4)	1.6-4.1 (2.5)	1.2-3.4 (2.0)	MLP 9-7-1 80%-70%-68% (0.5-2.0-1.7)
wektor „statystyczny 2” (9)	77-89 (82)	35-75 (55)	47-73 (63)	0.1-0.6 (0.5)	1.6-5.5 (2.3)	1.2-3.8 (2.0)	MLP 8-6-1 78%-70%-73% (0.5-1.7-1.4)
wektor „statystyczny 3” (7)	40-86 (80)	20-75 (60)	47-68 (63)	0.4-0.8 (0.5)	0.7-2.9 (1.8)	0.7-2.6 (1.5)	MLP 7-5-1 78%-70%-68% (0.7-1.5-1.2)
wektor „krokowa postępująca” (3)	36-75 (69)	45-75 (45)	42-75 (68)	0.6-0.8 (0.7)	0.7-1.1 (0.7)	0.6-0.9 (0.7)	MLP 3-6-1 75%-75%-68% 0.7-1.1-0.9
wektor „krokowa wsteczna” (11)	86-98 (92)	45-80 (65)	37-74 (58)	0.05-0.5 (0.2)	1.5-4.0 (2.6)	1.0-3.9 (2.0)	MLP 11-11-1 89%-75%-74% (0.1-3.0-1.9)
wektor „algorytm genetyczny” (9)	79-93 (86)	45-80 (65)	53-84 (74)	0.09-0.6 (0.4)	1.0-4.6 (2.0)	0.6-2.2 (1.0)	<b>MLP 9-12-1</b> <b>91%-75%-79%</b> <b>(0.2-1.8-1.1)</b>

Ze względu na utrzymujące się dla niemal wszystkich wektorów wejściowych wysokie wartości błędów walidacji i testowania przeprowadzono porównanie wyników uzyskanych dla stosowanego standardowo podziału na podzbiory uczący, walidacyjny i testowy z podziałem losowym dokonany w tych samych proporcjach (96:20:19). Wyniki uzyskane dla sieci neuronowych wykorzystujących wektor wejściowy „statystyczny 1” zestawiono w tabeli 74. Stwierdzono, że przeciętne błędy walidacji i testowania dla losowo wyodrębnionych

podzbiorów były znacznie niższe niż dla podzbiorów wyróżnionych na podstawie zastosowanej w tej pracy analizy skupień. Nie zanotowano natomiast tak znaczących różnic dotyczących jakości walidacji i testowania. Założono w związku z tym, że wartości błędów walidacji i testowania uzyskiwane dla standardowo używanych podzbiorów będą traktowane jako nieco zawyżone, natomiast jakości walidacji i testowania będą uważane za wiarygodne.

**Tabela 74. Porównanie wyników uzyskiwanych przez sieci neuronowe wykorzystujące wektor wejściowy „statystyczny 1” przy podziale na podzbiory uczący, walidacyjny i testowy wyodrębnione na podstawie analizy skupień oraz losowo (oddzielnie dla każdej sieci).**

Sposób podziału na podzbiory uczący, walidacyjny i testowy	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)					
wg analizy skupień	80-91 (85)	45-80 (65)	42-79 (65)	0.1-0.6 (0.4)	1.6-4.1 (2.5)	1.2-3.4 (2.0)
losowo (oddzielnie dla każdej sieci)	78-100 (91)	35-80 (60)	31-84 (63)	0.08-0.6 (0.4)	0.8-3.7 (1.7)	0.6-3.5 (1.5)

### **(b) Ocena liczby neuronów ukrytych**

Badania dotyczące oceny liczby neuronów ukrytych przeprowadzono dla wektora wejściowego „algorytm genetyczny”. Dla wektora tej wielkości parametr  $\rho$  uzyskuje optymalną wartość, gdy liczba neuronów ukrytych wynosi około 4. Najlepszy uzyskany wcześniej model wykorzystujący wektor wejściowy „algorytm genetyczny” (wyróżniony w tabeli 73) posiadał aż 12 neuronów w warstwie ukrytej i zgodnie z oczekiwaniami wykazywał tendencję do przeuczania się. Porównano w związku z tym modele zawierające od 2 do 12 neuronów ukrytych. Wyniki testów zestawiono w tabeli 75. Najefektywniejsza z otrzymanych sieci neuronowych posiadała 6 neuronów w warstwie ukrytej. Analiza wrażliwości nie wykazała dla tej sieci obecności potencjalnie zbędnych zmiennych wejściowych, ale wartość parametru  $\rho$  wynosiła 1.57, tj. nieco poniżej dolnej granicy przedziału optymalnego.

**Tabela 75. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „algorytm genetyczny”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
2	66-84 (76)	55-75 (65)	47-74 (63)	0.5-0.7 (0.7)	0.8-2.1 (1.3)	0.8-1.5 (1.1)	MLP 7-2-1 78%-75%-68% (0.7-1.0-0.8)
4	71-89 (79)	55-85 (65)	47-79 (68)	0.5-0.7 (0.6)	0.8-2.8 (1.4)	0.6-1.6 (1.1)	MLP 9-4-1 79%-75%-74% (0.6-1.4-0.9)
6	72-94 (81)	55-85 (65)	47-89 (68)	0.3-0.7 (0.5)	0.7-2.9 (1.5)	0.6-2.7 (1.0)	<b>MLP 8-6-1</b> <b>86%-75%-79%</b> <b>(0.5-0.8-0.7)</b>
8	74-94 (83)	50-80 (65)	47-89 (68)	0.2-0.7 (0.5)	0.9-3.1 (1.7)	0.5-2.1 (1.1)	MLP 9-8-1 88%-75%-79% (0.4-1.1-0.9)
10	75-93 (84)	50-80 (65)	52-84 (68)	0.2-0.6 (0.4)	0.9-4.3 (1.8)	0.7-1.7 (1.1)	MLP 9-10-1 81%-75%-79% (0.5-1.5-0.9)
12	77-93 (86)	50-85 (65)	47-89 (68)	0.1-0.6 (0.4)	0.9-4.2 (1.9)	0.5-1.9 (1.1)	MLP 9-12-1 88%-80%-79% (0.4-1.5-1.3)

## ➤ ETAP II

Na podstawie wyników uzyskanych w I etapie badań zdecydowano ocenić sieci neuronowe wykorzystujące wektor wejściowy „algorytm genetyczny” i posiadające od 4 neuronów w warstwie ukrytej (ponieważ dla tej liczby parametr  $\rho$  uzyskuje optymalną wartość) do 6 neuronów w warstwie ukrytej (gdyż dla tej liczby uzyskano wcześniej najlepsze modele). Testowano w związku z tym sieci neuronowe o następujących architekturach:

- Wektor wejściowy „algorytm genetyczny”, 4 neurony ukryte.  
Żadna z 1000 sieci wygenerowanych przez *Automatycznego projektanta* nie uzyskała równocześnie tak wysokich jakości uczenia i testowania jak wcześniej otrzymane modele.

- Wektor wejściowy „algorytm genetyczny”, 5 neuronów ukrytych.  
Dla tej architektury uzyskano model, który uznano ostatecznie za najlepszy. Był to perceptron o architekturze MLP 9-5-1, o jakościach uczenia, walidacji i testowania odpowiednio 89%-75%-79% i błędach 0.4-0.9-0.8. Analiza wrażliwości nie wykazała w przypadku tego modelu obecności zbędnych zmiennych wejściowych, a wartość parametru  $\rho$  była bliższa optimum niż dla innych sieci z podobną efektywnością prognozujących wartość zmiennej BÓL\_GŁ\_bin\_6 ( $\rho=1.7$ ). Charakterystykę tego modelu przedstawiono w tabeli 76.
- Wektor wejściowy „algorytm genetyczny”, 6 neuronów ukrytych.  
Wśród sieci o tej architekturze uzyskano model o wyższej jakości walidacji niż opisany powyżej (MLP 9-6-1, jakości odpowiednio: 88%-80%-79%), jednak wyższe błędy walidacji i testowania (0.5-1.1-1.2) oraz bardziej odległa od optymalnej wartość parametru  $\rho$  świadczyły, że w rzeczywistości zdolność tej bardziej rozbudowanej sieci neuronowej do generalizacji była niższa niż modelu posiadającego 5 neuronów ukrytych.

**Tabela 76. Najlepszy z uzyskanych modeli dokonujących predykcji nasilenia bólu głowy po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 9-5-1	skrzywienie przegrody, kaszel, ból twarzy, ból czoła, ból głowy, ból ucha, wdech, zmęczenie, ból zębów	89%-75%-79%	0.4-0.9-0.8	nie	nie

#### 6.9.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa BÓL\_GŁ\_bin\_12)

Modele predykcyjne dla zmiennej BÓL\_GŁ\_bin\_12 opracowano wykorzystując podział zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 47:12:10. W trakcie poszukiwań najbardziej obiecującego wektora wejściowego stwierdzono, że algorytmy doboru cech w większości przypadków dokonywały wyboru wektorów wejściowych zawierających te same zmienne, od których zmienna wyjściowa wykazywała statystycznie istotną zależność (por. tab. 72). Zmienne te opisywały przede wszystkim dolegliwości bólowe zgłaszane przed operacją (ból twarzy, głowy, czoła, ucha, zębów). Najlepszy zestaw zmiennych wejściowych uzyskano dzięki zastosowaniu metody krokowej wstecznej, a

następnie dalszej redukcji wektora wejściowego przy pomocy *Automatycznego projektanta sieci*. Posługując się wybranym w ten sposób zestawem zmiennych uzyskano liczne modele a architekturach MLP 8-2-1 lub MLP 7-3-1, wykazujące bardzo podobną efektywność i jakości testowania około 80%. Charakterystykę sieci neuronowej, którą ostatecznie uznano za najlepszą, przedstawiono w tabeli 77.

**Tabela 77. Najlepszy z uzyskanych modeli dokonujących predykcji nasilenia bólu głowy po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 8-2-1	płeć, wydzielina, spływanie, ból twarzy, ból czoła, ból głowy, węch, zmęczenie	87%-92%-80%	0.4-0.3-1.3	tak	tak

### **6.9.3 Podsumowanie**

Przedstawione powyżej wyniki dotyczące predykcji nasilenia bólu głowy okazały się - wbrew oczekiwaniom - bardzo podobne do omawianych w poprzednim podrozdziale, a dotyczących bólu twarzy. Chociaż ból głowy nie jest objawem ściśle powiązanim z występowaniem PZZP i może wynikać z wielu innych przyczyn, u wielu chorych obserwowano jego ustępowanie po operacji zatok. Efekt ten był zauważalny nawet 12 miesięcy po operacji. Analizując dane historyczne stwierdzono, że po zabiegu 59% pacjentów z grupy I i 65% pacjentów z grupy II nie zgłaszało bólu głowy. W przypadku bólu twarzy odsetek chorych wolnych od dolegliwości po operacji był bardzo podobny. Nasilenie bólu głowy po operacji udało się prawidłowo przewidzieć przy pomocy sieci neuronowych u około 80% chorych w obu grupach. Oznacza to, że występowanie tego objawu po operacji jest przewidywalne w podobnym stopniu, jak obecność bólu twarzy. Co więcej, prognozowanie nasilenie bólu głowy okazało się łatwiejsze niż predykcja niektórych dolegliwości wynikających bezpośrednio z obecności stanu zapalnego w jamie nosa i zatokach przynosowych, takich jak na przykład obecność wydzieliny z nosa. Ponadto zarówno w przypadku bólu głowy, jak i bólu twarzy, zauważono, że istotne znaczenie dla predykcji mają przede wszystkim zmienne opisujące nasilenie wszelkich dolegliwości bólowych przed zabiegiem operacyjnym oraz uczucie przewlekłego zmęczenia. Uzyskane wyniki po raz kolejny sugerują, że odczuwanie dolegliwości bólowych (w tym również bólu głowy) po operacji jest w większym stopniu zależne od indywidualnych cech chorego wyrażających się w jego subiektywnej ocenie własnego samopoczucia, niż od obiektywnej charakterystyki schorzenia dotyczącego jego zatok przynosowych.

## 6.10 Przewlekłe zmęczenie

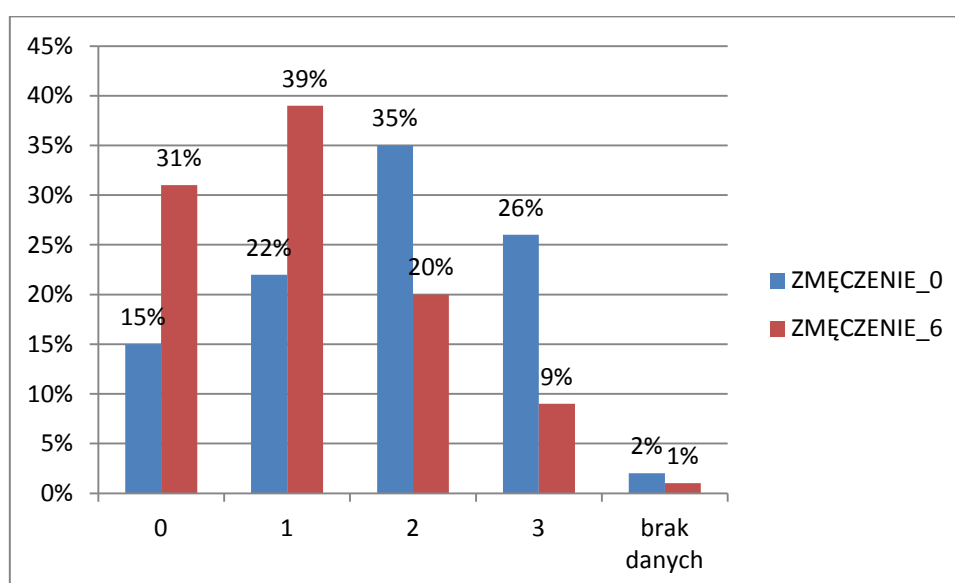
Przewlekłe zmęczenie jest objawem często zgłaszanym przez pacjentów z PZZP, jednak – jak wykazał przegląd piśmiennictwa – jego uwarunkowania w przebiegu tego schorzenia pozostają niejasne (por. podrozdział 3.6.9). W związku z tym ocenę możliwości przewidywania nasilenia tej dolegliwości przy pomocy sztucznych sieci neuronowych uznano za szczególnie interesującą. Pacjenci oceniali nasilenie przewlekłego zmęczenia w skali 4-stopniowej. Przyjęto następujące nazwy zmiennych:

- ZMĘCZENIE\_0 – nasilenie przewlekłego zmęczenia w skali 4-stopniowej przed zabiegiem,
- ZMĘCZENIE\_6 – nasilenie przewlekłego zmęczenia w skali 4-stopniowej po 3-6 miesiącach (dotyczy grupy I),
- ZMĘCZENIE\_12 – nasilenie przewlekłego zmęczenia w skali 4-stopniowej po 12 miesiącach (dotyczy grupy II).

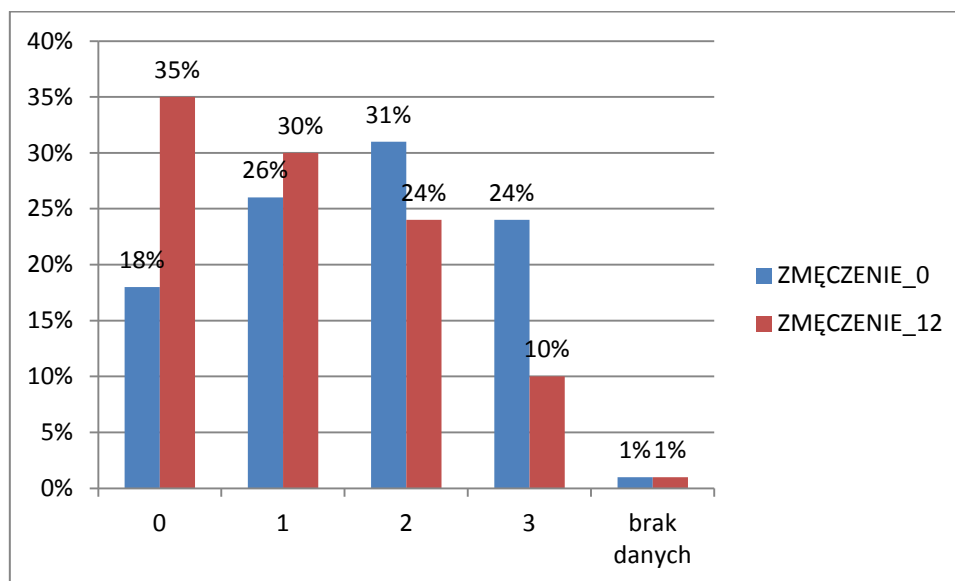
### 6.9.1 Analiza statystyczna

Rysunki 44 i 45 pokazują, jak w obu badanych grupach po zabiegu operacyjnym zmniejszył się odsetek pacjentów skarżących się na przewlekłe zmęczenie. Zidentyfikowano również zmienne, które w sposób statystycznie istotny wpływały na wartość zmiennej ZMĘCZENIE, przy czym najsilniejszy związek występował między pooperacyjnym nasileniem przewlekłego zmęczenia, a jego uciążliwością zgłaszaną jeszcze przed operacją (tab. 78).

**Rysunek 44. Wartości zmiennej ZMĘCZENIE w grupie I – przed zabiegiem (ZMĘCZENIE\_0) i 3-6 miesięcy po zabiegu (ZMĘCZENIE\_6)**



**Rysunek 45. Wartości zmiennej ZMĘCZENIE w grupie II – przed zabiegiem (ZMĘCZENIE \_0) i 3-6 miesięcy po zabiegu (ZMĘCZENIE \_12)**



**Tabela 78. Zależności między uciążliwością przewlekłego zmęczenia po operacji (wyrażoną w skali 4-stopniowej) a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych, a w przypadku istotnych statystycznie korelacji podano wartość współczynnika korelacji rang Spearmana (R).**

Cecha	ZMĘCZENIE _6 (grupa I)	ZMĘCZENIE _12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	-
czas objawów	-	-
astma	p = 0.007594	-
AIA	p = 0.001228	-
alergia	(blisko granicy istotności) p = 0.074395	-
refluks	p = 0.025211	p = 0.033025



Cecha	ZMĘCZENIE _6 (grupa I)	ZMĘCZENIE _12 (grupa II)
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	p = 0.029066	p = 0.043575
liczba przebytych ESS	R = 0.194693 p = 0.022619	R = 0.257670 p = 0.030052
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	-	-
wydzielina z nosa	-	-
splýwanie wydzieliny	R = 0.245575 p = 0.003822	R = 0.268466 p = 0.025719
kaszel	-	R = 0.387209 p = 0.001013
kichanie	-	-
ból twarzy	R = 0.213068 p = 0.012426	R = 0.347895 p = 0.003167
ból czoła	R = 0.217926 p = 0.010813	R = 0.348455 p = 0.003856
ból głowy	-	R = 0.468414 p = 0.000049
dolegliwości uszne	R = 0.261875 p = 0.001993	R = 0.319536 p = 0.007904
zaburzenia węchu	-	-
zmęczenie	R = 0.527639 p < 0.000001	R = 0.672549 p < 0.000001
nieprzyjemny zapach z ust	R = 0.227187 p = 0.007590	R = 0.349500 p = 0.003021
ból zębów	R = 0.196669 p = 0.021256	R = 0.273189 p = 0.022125

Rozważając, w jaki sposób dokonać przekształcenia zmiennej ZMĘCZENIE w zmienną binarną napotkano następujący problem: w obu grupach chorych brak uczucia zmęczenia po zabiegu (wartość zmiennej ZMĘCZENIE równą 0) zgłaszało zaledwie około 30% chorych. W

związku z tym wyróżnienie podgrupy chorych nie zgłaszających tej dolegliwości oraz podgrupy chorych zgłaszających mało, średnio lub bardzo nasiloną dolegliwość prowadziłyby do nierównomiernego podziału w proporcjach około 30% : 70%. Co więcej, przyjęcie wyższego progu binaryzacji (wyróżnienie podgrupy zgłaszającej brak lub niewielkie nasilenie zmęczenia oraz podgrupy zgłaszającej średnie lub duże nasilenie zmęczenia) skutkowałoby podobną nierównowagą (po uwzględnieniu braków danych byłyby to proporcje 70% : 30 w grupie I i 66% : 34% w grupie II). Z punktu widzenia klinicznego bardziej zasadne wydaje się wyróżnienie grupy chorych, którzy po zabiegu wcale nie będą odczuwać przewlekłego zmęczenia, niż łączne potraktowanie pacjentów nie odczuwających zmęczenia oraz odczuwających je w małym stopniu. W związku z tym zdecydowano następująco zdefiniować zmienną binarną odnoszącą się do przewlekłego zmęczenia, godząc się na to, że grupy odpowiadające wartości zmiennej binarnej odpowiednio wynoszącej 0 i 1 jeden będą w tym przypadku nierówno liczne:

- jeżeli ZMĘCZENIE = 0 to ZMĘCZ\_bin = 0 (tzn. po zabiegu przewlekłe zmęczenie nie występuje),
- jeżeli ZMĘCZENIE > 0 to ZMĘCZ\_bin = 1 (tzn. po zabiegu przewlekłe zmęczenie występuje w stopniu małym, średnim lub znacznym).

Nazwy nowych zmiennych binarnych ustalono następująco:

- dla grupy I (ocenianej 3-6 miesięcy po zabiegu): ZMĘCZ\_bin\_6,
- dla grupy II (ocenianej 12 miesięcy po zabiegu): ZMĘCZ\_bin\_12.

W tabeli 79 przedstawiono zależności między tak zdefiniowanymi zmiennymi binarnymi a zmiennymi wejściowymi budowanych modeli.

**Model przewidujący wartość zmiennej ZMĘCZ\_bin odpowiada na pytanie pacjenta:  
„Czy po operacji będę odczuwać przewlekłe zmęczenie?”**

**Tabela 79. Zależności pomiędzy uciążliwością przewlekłego zmęczenia po operacji wyrażoną za pomocą zmiennej ZMĘCZ\_bin a zmiennymi wejściowymi. Wartości poziomu istotności p podano dla zależności statystycznie istotnych.**

Cecha	ZMĘCZ_bin_6 (grupa I)	ZMĘCZ_bin_12 (grupa II)
pleć	-	-
wiek	-	-
polipy	-	-
masywne polipy	-	-
skrzywienie przegrody nosa	-	-
TK (skala Lund-Mackay)	-	-

Cecha	ZMĘCZ_bin_6 (grupa I)	ZMĘCZ_bin_12 (grupa II)
radykałność zabiegu	-	-
histopatologia – dominacja eozynofili	-	p=.01542
czas objawów	-	-
astma	-	-
AIA	p= 0.04644	-
alergia	-	-
refluks	p= 0.03141	(blisko granicy istotności) p=0.08480
przebyte polipektomie	-	-
liczba przebytych polipektomii	-	-
przebyte ESS	(blisko granicy istotności) p=.06698	-
liczba przebytych ESS	-	-
<b>nasilenie dolegliwości przed zabiegiem:</b>		
VAS_0	-	-
niedrożność nosa	-	-
wydzielina z nosa	-	-
splywanie wydzieliny	p = 0.019769	-
kaszel	-	p = 0.013445
kichanie	-	-
ból twarzy	(blisko granicy istotności) p = 0.054830	p = 0.020196
ból czoła	p = 0.033412	p = 0.038662
ból głowy	-	p = 0.004170
dolegliwości uszne	p = 0.020264	-
zaburzenia węchu	-	p = 0.044416
zmęczenie	p = 0.000006	p < 0.000001
nieprzyjemny zapach z ust	(blisko granicy istotności) p = 0.067631	p = 0.038895
ból zębów	-	-

## **6.10.2 Modele neuronowe przewidujące występowanie przewlekłego zmęczenia po zabiegu operacyjnym**

### **6.10.2.1 Grupa I – 3-6 miesięcy po operacji (zmienna wyjściowa ZMĘCZ\_bin\_6)**

#### **➤ ETAP I:**

#### **(a) Poszukiwanie wektora wejściowego**

Budując modele predykcyjne dla grupy I posługiwano się podziałem bazy danych na podzbiory uczący, walidacyjny i testowy w proporcjach 97:20:20. Oceniono następujące wektory wejściowe (por. tab. 80):

- Wektory opracowane na podstawie analizy statystycznej:
  - Wektor „statystyczny 1” – 9 zmiennych, od których zmienna ZMĘCZ\_bin\_6 była statystycznie istotnie zależna lub wynik był bliski granicy istotności statystycznej (por. tab. 79).  
**Wynik testu:** Ze względu na wysokie błędy i niskie jakości testowania zdecydowano zmniejszyć wektor wejściowy.
  - Wektor „statystyczny 2” – 6 zmiennych, od których statystycznie istotnie zależna była zmienna ZMĘCZ\_bin\_6 (por. tab. 79).  
**Wynik testu:** Mimo obniżenia przeciętnych błędów testowania, jakości testowania pozostały bardzo niskie (do 65%).
- Wektory uzyskane dzięki algorytmom doboru cech:
  - Wektor „krokowa postępująca – 3 zmienne (czas objawów, kichanie, zmęczenie).  
**Wynik testu:** Sieci neuronowe wykorzystujące tak mały wektor wejściowy wykazywały lepszą zdolność do generalizacji, ale jakości uczenia okazały się niższe niż dla poprzednio ocenianych wektorów, a jakości testowania nadal nie przekraczały 70%.
  - Wektor „krokowa wsteczna” – 6 zmiennych (masywne polipy, zabieg, niedrożność, wdech, zmęczenie, refluks).  
**Wynik testu:** Modele wykorzystujące ten wektor wejściowy osiągnęły najwyższe jakości testowania (do 75%).
  - Wektor „algorytm genetyczny” – 6 zmiennych (sPLYwanie, ból głowy, ból ucha, zmęczenie, AIA, refluks).  
**Wynik testu:** Ponownie uzyskano sieci neuronowe o jakości testowania nie przekraczającej 65%.

**Tabela 80. Porównanie jakości i błędów sieci neuronowych uzyskanych dla różnych wektorów wejściowych**

Wektor wejściowy (liczba zmiennych)	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
wektor „statystyczny 1” (9)	82-98 (90)	40-70 (55)	25-60 (45)	0.3-0.6 (0.5)	0.7-1.8 (1.1)	1.3-4.5 (2.1)	MLP 9-12-1 88%-60%-60% (0.5-0.9-1.9)
wektor „statystyczny 2” (6)	75-91 (86)	50-80 (55)	35-65 (45)	0.5-0.8 (0.6)	0.6-1.4 (0.9)	1.0-1.9 (1.4)	MLP 6-6-1 80%-80%-65% (0.7-0.7-1.1)
wektor „krokowa postępująca” (3)	53-77 (72)	25-80 (75)	45-70 (65)	0.6-0.7 (0.7)	0.6-0.8 (0.7)	0.6-1.0 (0.8)	MLP 3-8-1 75%-75%-65% (0.7-0.7-0.9)
wektor „krokowa wsteczna” (6)	70-82 (73)	65-80 (70)	55-75 (65)	0.5-0.6 (0.6)	0.8-1.2 (1.0)	1.2-1.8 (1.4)	<b>MLP 6-11-1</b> <b>74%-75%-70%</b> <b>(0.5-0.9-1.2)</b>
wektor „algorytm genetyczny” (6)	73-80 (77)	60-85 (75)	50-65 (60)	0.6-0.8 (0.6)	0.6-0.7 (0.7)	0.8-1.4 (1.1)	MLP 4-7-1 76%-85%-65% (0.7-0.6-1.0)

### **(b) Ocena liczby neuronów ukrytych**

W kolejnym etapie badań zdecydowano korzystać z wektora wejściowego „krokowa wsteczna”, dla którego uzyskano najwyższe jakości testowania. Najlepszy model uzyskany dla tego wektora w I etapie badań, wyróżniony w tabeli 80, posiadał 11 neuronów ukrytych. Ponieważ optymalne wartości parametr  $\rho$  przyjmuje dla 6 neuronów ukrytych, przetestowano modele posiadające 6-11 neuronów w warstwie ukrytej. Wyniki eksperymentów zestawiono w tabeli 81. Najlepsze modele osiągały jakość testowania do 75%, jednak utrzymywały się stosunkowo wysokie błędy walidacji i testowania.

**Tabela 81. Porównanie jakości i błędów sieci neuronowych przy zmieniającej się liczbie neuronów ukrytych dla wektora wejściowego „krokowa wsteczna”**

Liczba neuronów ukrytych	Jakość uczenia [%]	Jakość walidacji [%]	Jakość testowania [%]	Błąd uczenia	Błąd walidacji	Błąd testowania	Najlepsze uzyskane modele: jakości (błędy)
	zakres wartości od najmniejszej do największej (w nawiasie: mediana)						
6	68-81 (72)	60-85 (70)	55-75 (65)	0.6-0.7 (0.6)	0.7-1.2 (1.0)	0.7-1.8 (1.3)	<b>MLP 6-6-1</b> 73%-75%-75% <b>(0.6-1.0-1.4)</b>
8	69-80 (72)	60-85 (70)	60-75 (65)	0.5-0.6 (0.6)	0.8-1.2 (1.0)	1.0-1.7 (1.3)	MLP 6-8-1 74%-75%-70% (0.6-1.0-1.2)
10	70-82 (72)	65-85 (70)	55-75 (65)	0.5-0.6 (0.6)	0.8-1.2 (1.0)	1.1-1.7 (1.4)	MLP 6-10-1 74%-75%-70% (0.6-1.0-1.6)
11	69-81 (73)	60-85 (70)	55-75 (65)	0.5-0.6 (0.6)	0.7-1.2 (1.0)	1.1-2.0 (1.4)	<b>MLP 6-11-1</b> 80%-75%-75% <b>(0.5-1.1-1.7)</b>

## ➤ ETAP II

Testowano sieci neuronowe o następujących architekturach:

- Wektor wejściowy „krokowa wsteczna”, 6 neuronów ukrytych.  
Przy pomocy *Automatycznego projektanta* zbudowano 1000 sieci o tej architekturze. Najlepszym z uzyskanych w ten sposób modeli była sieć o architekturze MLP 4-6-1 o jakościach uczenia, walidacji i testowania 71%-75%-75% i błędach odpowiednio 0.6-0.7-1.1. Ze względu na niską jakość uczenia nie uznano tego modelu za lepszy niż otrzymany na wcześniejszym etapie model MLP 6-6-1, wyróżniony w tabeli 81. Dla tego ostatniego modelu analiza wrażliwości wykazywała, że zmienna wejściowa „zmęczenie” potencjalnie może pogarszać działanie modelu, chociaż intuicyjnie wydawało się to mało prawdopodobne ze względu na wysoce istotną statystycznie zależność między tą zmienną a zmienną przewidywaną. Przetestowano jednak 1000 modeli wykorzystujących wektor wejściowy „krokowa wsteczna” bez zmiennej „zmęczenie” i otrzymano sieci neuronowe o maksymalnej jakości testowania 50%.
- Wektor wejściowy „krokowa wsteczna”, 11 neuronów ukrytych.  
Najlepszą z 1000 wygenerowanych przez *Automatycznego projektanta* sieci był model MLP 6-11-1 o jakościach uczenia, walidacji i testowania 71%-75%-75% i błędach

odpowiednio 0.6-0.9-1.5. Również tego modelu nie uznano za lepszy od otrzymanego wcześniej MLP 6-6-1 ze względu na niską jakość uczenia oraz zdecydowanie gorszą wartość parametru  $\rho$ .

W wyniku przeprowadzonych eksperymentów za najlepszy model uznano sieć neuronową uzyskaną na etapie doboru liczby neuronów ukrytych, której charakterystykę przedstawiono w tabeli 82.

**Tabela 82. Najlepszy z uzyskanych modeli dokonujących predykcji uciążliwości przewlekłego zmęczenia po zabiegu w I grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 6-6-1	masywne polipy, zabieg, niedrożność, węch, zmęczenie, refluks	73%-75%-75%	0.6-1.0-1.4	tak	tak

#### 6.10.2.2 Grupa II – 12 miesięcy po operacji (zmienna wyjściowa ZMĘCZ\_bin\_12)

Prowadząc obliczenia dla grupy II posługiwano się podziałem zbioru danych na podzbiory uczący, walidacyjny i testowy w proporcjach 48:12:11. Najlepsze wyniki uzyskano dzięki wykorzystaniu wektora wejściowego liczącego zaledwie 4 zmienne, wygenerowanego dzięki zastosowaniu algorytmów doboru cech (metody krokowej wstecznej) i dalszej redukcji liczby zmiennych wejściowych do 3 przy pomocy *Automatycznego projektanta sieci*. Najefektywniejsze modele posiadały 4 neurony w warstwie ukrytej. Jakkolwiek parametr  $\rho$  przyjmuje optymalne wartości dla 5 neuronów ukrytych, zwiększenie rozmiarów warstwy ukrytej o 1 neuron prowadziło do pogorszenia zdolności do generalizacji. Ostatecznie za najlepszy model uznano sieć neuronową scharakteryzowaną w tabeli 83.

Obserwacja zachowania testowanych powyżej sieci neuronowych doprowadziła do wniosku, że najistotniejszą zmienną dla predykcji zmęczenia zgłaszanego po zabiegu jest subiektywna uciążliwość zmęczenia przed zabiegiem. Podjęto w związku z tym próbę zbudowania modeli korzystających tylko z jednej zmiennej wejściowej („zmęczenie”). Uzyskano w ten sposób sieci neuronowe o jakościach testowania sięgających 91%, jakkolwiek jakości walidacji nie przekraczały 67%. Zakładając, że zbiór walidacyjny może być niereprezentatywnie trudny, przeprowadzono podobny eksperyment dla wylosowanych jednorazowo podzbiorów uczącego, walidacyjnego i testowego. Uzyskano sieci neuronowe o jakościach uczenia do 81% oraz jakościach walidacji i testowania do 88%. Wyniki tych dodatkowo

przeprowadzonych testów sugerują, że najważniejszym wyznacznikiem poziomu zmęczenia odczuwanego przez chorego po zabiegu jest subiektywna ocena tej dolegliwości dokonana już w okresie przedoperacyjnym. Chorzy, którzy przed operacją skarżą się na bardzo nasilone zmęczenie, z dużym prawdopodobieństwem nadal będą odczuwać tę dolegliwość mimo przeprowadzonego leczenia. Informacja wnoszona przez dodatkowe zmienne opisujące pacjentów nie ma natomiast bardzo istotnego znaczenia dla dokonywanych prognoz.

**Tabela 83. Najlepszy z uzyskanych modeli dokonujących predykcji uciążliwości przewlekłego zmęczenia po zabiegu w II grupie pacjentów**

Architektura sieci	Wektor wejściowy	Jakość: uczenia-walidacji-testowania	Błąd: uczenia-walidacji-testowania	Analiza wrażliwości: potencjalnie zbędne zmienne?	Wartość parametru $\rho$ w przedziale optymalnym?
MLP 3-4-1	ból głowy, zmęczenie, refluks	92%-83%-82%	0.6-1.1-1.0	nie	nie

### **6.10.3 Podsumowanie**

Opisane powyżej wyniki dotyczące tworzenia modeli prognozujących przewlekłe zmęczenie u pacjentów operowanych z powodu PZZP okazały się szczególnie interesujące w kontekście danych z literatury cytowanej w podrozdziale 3.6.9 niniejszej pracy. Jak podaje piśmiennictwo, ponad połowa pacjentów z PZZP skarży się na przewlekłe zmęczenie. Spośród chorych badanych w tej pracy jedynie kilkanaście procent nie zgłaszało tego objawu przed operacją. Po zabiegu wolna od uczucia przewlekłego zmęczenia pozostawała zaledwie około jedna trzecia pacjentów. Dzięki zastosowaniu sztucznych sieci neuronowych nasilenie omawianego objawu udało się prawidłowo przewidywać u około 75% chorych z grupy I i aż 82% z grupy II. Analizując dostępne piśmiennictwo zidentyfikowano tylko 1 zmienną warunkującą uciążliwość przewlekłego zmęczenia po operacji. Było to nasilenie tego samego objawu zgłaszane przed operacją, przy czym pacjenci skarżący się na szczególnie nasilone zmęczenie przed zabiegiem podawali w tym zakresie większą poprawę niż pozostali chorzy. W niniejszej pracy stwierdzono z kolei, że utrzymywanie się tego objawu w okresie pooperacyjnym zgłaszają na ogół te osoby, które już przed operacją skarżyły się na jego znaczną uciążliwość. Wszystkie najskuteczniejsze modele predykcyjne zbudowane w ramach niniejszej pracy wykorzystywały w ramach swoich wektorów wejściowych zmienną opisującą uciążliwość zmęczenia przed operacją. Co więcej, modele bazujące wyłącznie na tym jednym wejściu również okazały się bardzo skuteczne. Oznacza to, że nasilenie przewlekłego zmęczenia w okresie pooperacyjnym jest przewidywalne u tak dużego odsetka chorych przede wszystkim dlatego, że wykazuje silną zależność od subiektywnej samooceny tego objawu przed zabiegiem. Pozostałe informacje dotyczące na przykład obiektywnie ocenianego zaawansowania choroby lub schorzeń towarzyszących okazały się niewiele wnosić do ostatecznej prognozy.



## 7. Dyskusja

### 7.1. Ocena zasadności podjętych badań

Przegląd dostępnego piśmiennictwa oraz analiza statystyczna danych zebranych na potrzeby niniejszej pracy dają podstawy do stwierdzenia, że **podjęta w ramach prezentowanych badań próba komputerowego wspomaganie procesu przewidywania wyników leczenia operacyjnego chorych z PZZP jest uzasadniona i istotna z klinicznego punktu widzenia.**

Jak podaje literatura, większość pacjentów zgłasza poprawę w odniesieniu do każdego z objawów subiektywnych po operacji endoskopowej zatok przynosowych, jednak wszyscy badacze potwierdzają, że u pewnego odsetka pacjentów nie uzyskuje się zadowalających wyników mimo prawidłowo przeprowadzonego leczenia [26, 27]. Podobną prawidłowość zaobserwowano w obu grupach chorych badanych w ramach niniejszej pracy. Jeżeli zatem u niektórych pacjentów operacja nie przynosi zadowalających efektów, to jej przeprowadzenie u tych osób mogłoby być uznane za nieuzasadnione – pod warunkiem, że już na etapie kwalifikacji do zabiegu byłaby znana wiarygodna prognoza dotycząca spodziewanego niepowodzenia leczenia. Obecnie nie istnieją narzędzia, które pozwalałyby z dostatecznie dużym prawdopodobieństwem dokonywać tego rodzaju predykcji. W związku z tym uzasadnione wydaje się podjęcie badań mających na celu przewidywanie, u których chorych można się spodziewać pozytywnego wyniku operacji, a u których uzyskanie poprawy jest mało prawdopodobne.

### 7.2. Zagadnienie przewidywalności wyników leczenia chorych z PZZP w kontekście przeglądu piśmiennictwa i w świetle własnych badań

Zarówno przegląd piśmiennictwa, jak i analizy statystyczne przeprowadzone w niniejszej pracy, pozwalają zauważyć, że nasilenie każdej z dolegliwości związanych z PZZP jest zależne od złożonego wpływu rozmaitych czynników klinicznych.

Analizując dostępną literaturę stwierdzono, że większość prac badawczych dotyczyła wpływu rozmaitych zmiennych na nasilenie poszczególnych objawów PZZP głównie przed leczeniem operacyjnym, a nie po jego zakończeniu. Tymczasem czynniki warunkujące uciążliwość objawów już po zabiegu – czyli takie, które mogą służyć do prognozowania wyników leczenia - często różnią się od tych, które miały istotny wpływ na nasilenie dolegliwości przed operacją. Na przykład obecność polipów nosa w oczywisty sposób powoduje jego niedrożność przed zabiegiem, ale po usunięciu polipów na drodze operacyjnej drożność nosa u chorych, którzy wcześniej mieli polipy, nie musi wcale być gorsza niż u pozostałych operowanych pacjentów [46, 94-97]. Dlatego przeprowadzana przez wielu autorów ocena, którzy pacjenci z PZZP mają bardziej nasiloną niedrożność nosa przed zabiegiem, może się wydawać trywialna (gorsza drożność nosa u pacjentów z polipami nosa jest oczywista), podczas gdy przewidywanie, u których z nich upośledzona drożność utrzyma się mimo

przeprowadzenia zabiegu, musi się opierać na zupełnie innych przesłankach i stanowi poważne wyzwanie naukowe i praktyczne.

W przypadku niektórych dolegliwości związanych z PZZP (np. zaburzeń węchu) przegląd literatury dostarczył informacji o licznych czynnikach ryzyka warunkujących ich nasilenie po operacji [96, 107, 108, 111, 117, 119, 121, 125-127], natomiast dla innych objawów (np. przewlekłego zmęczenia) istnienie tego rodzaju prostych zależności nie zostało wykazane więc ich przyczynowe uwarunkowanie nie było oczywiste [100, 179]. Intuicyjnie można oczekiwać, że najłatwiej przewidywalne powinno być pooperacyjne nasilenie tych dolegliwości, których utrzymywanie się jest zależne od wielu znanych czynników.

Przeprowadzone w tej pracy eksperymenty wykazały, że dla możliwości prognozowania najważniejsze jest istnienie **możliwych do przedstawienia w formie modelu zależności między zmiennymi wejściowymi a rozważaną zmienną wyjściową**. Należy zauważyć, że zależności te wcale nie muszą w jasny sposób ujawniać się w prostych analizach statystycznych, które stanowią podstawę identyfikacji czynników ryzyka niepowodzenia leczenia stosowaną w większości tradycyjnie prowadzonych prac badawczych. Zależności możliwe do przedstawienia w formie modelu (na przykład przy pomocy stosowanych w tej pracy sieci neuronowych) **mogą mieć złożony, nieliniowy charakter**, co jednak w żaden sposób nie komplikuje pracy badacza, gdyż model formowany jest przez automatyczny proces uczenia sieci. Okoliczność ta wskazuje na kolejną przewagę użycia sieci neuronowych w stosunku do metod statystycznych, ponieważ przy tworzeniu wielowymiarowego nieliniowego modelu regresyjnego metodami statystycznymi użytkownik tych metod musi sam podać wiele założeń dotyczących kształtu tworzonego modelu. Założenia te są zwykle trudne do sformułowania i jeszcze trudniejsze do uzasadnienia. Tymczasem badacz korzystający z sieci neuronowych osiąga cel bez konieczności pokonywania tych trudności.

Ponadto w przypadku niewielkiej liczebności badanych grup pacjentów osiągnięcie prognozy istotności statystycznej może nie być możliwe. Powoduje to, że modelu statystycznego nie można skutecznie zbudować, mimo obiektywnego istnienia określonych zależności w populacji. Natomiast modele oparte na sieciach neuronowych są od tych ograniczeń wolne.

Analizując wyniki eksperymentów opisanych w tej pracy można na przykład zauważyć, że zmienna „niedrożność nosa” w I grupie chorych była statystycznie istotnie zależna jedynie od nielicznych zmiennych wejściowych, a mimo to modele neuronowe osiągnęły 85% poprawnych predykcji w zbiorze testowym. Dla grupy II dzięki zastosowaniu odpowiednich algorytmów doboru cech uzyskano 70% prawidłowych odpowiedzi **mimo braku jakichkolwiek statystycznie istotnych zależności między dostępnymi zmiennymi wejściowymi a zmienną wyjściową**. Tymczasem na przykład w odniesieniu do prognozowania ilości wydzieliny z nosa większość sieci neuronowych nie osiągnęła jakości testowania powyżej 65%, co oznacza, że w tym przypadku w zbiorze danych uczących naprawdę brakowało istotnej informacji o tym, w jaki sposób dostępne zmienne wejściowe warunkują wartość tej właśnie zmiennej wyjściowej.

Warto zauważyć, że skuteczne prognozowanie nie okazuje się prostsze, jeśli analizowany objaw łatwo ustępuje po leczeniu operacyjnym. Można nawet stwierdzić, że **jeżeli dolegliwość ma tendencję do uporczywego utrzymywania się mimo leczenia u tych chorych, u których już przed zabiegiem była bardzo nasiloną, to przewidywanie wyniku bywa mniej skomplikowane**. Jako przykład można podać predykcję zaburzeń węchu. Zaburzenia te, mimo złożonych uwarunkowań patofizjologicznych, okazały się możliwe do przewidzenia u niemal 90% chorych. Przewidywalność tego objawu wynikała z jednej strony z istnienia licznych zależności między upośledzeniem węchu a różnymi dobrze zdefiniowanymi czynnikami klinicznymi, z drugiej zaś strony z faktu, że funkcja węchu po zabiegu była tym gorsza, im większe zaburzenia w tym zakresie notowano już przed operacją, więc model mógł wykorzystać tę prawidłowość.

Podsumowując należy zaznaczyć, że **nasilenie większości dolegliwości związanych z PZZP udało się z powodzeniem prognozować przy pomocy sztucznych sieci neuronowych** u 70-94% pacjentów (por. tab. 84), przy czym przewidywalność poszczególnych zmiennych była zależna od tego, czy przykłady zawarte w zbiorze uczącym zawierały informacje o zależnościach między zmiennymi wejściowymi a zmienną wyjściową. Trudność prognozowania nie zawsze była możliwa do wstępnej oceny wyłącznie na podstawie danych z literatury i analizy statystycznej zależności między zmiennymi wejściowymi a zmienną wyjściową, ponieważ sztuczne sieci neuronowe były w stanie wymodelować również takie zależności, których nie udało się odkryć przy użyciu prostych testów statystycznych.

**Tabela 84. Jakość testowania osiągnięta przez najlepsze modele neuronowe dla poszczególnych zmiennych wyjściowych.**

Zmienna wyjściowa	grupa I	grupa II
VAS	79%	82%
dVAS	94%	82%
niedrożność	85%	70%
wydzielina	75%	70%
splywanie	70%	91%
węch	84%	90%
ból twarzy	80%	90%
ból głowy	79%	80%
zmęczenie	75%	82%

### 7.3. Porównanie modeli opracowanych w ramach niniejszej pracy z modelami predykcyjnymi opisywanymi w literaturze

W rozdziale dotyczącym przeglądu literatury zaznaczono, że w dostępnym piśmiennictwie udało się odnaleźć wzmianki jedynie o trzech modelach mających z założenia służyć do prognozowania wyników leczenia operacyjnego chorych z PZZP.

Pierwszym z nich był **model regresyjny**, który opracowali Sil, Mackay i Rowe-Jones [37]. Narzędzie to dokonywało prawidłowych predykcji wyników leczenia operacyjnego u 81.7% chorych z PZZP. Istnieją obiektywne trudności w porównaniu tego wyniku z efektywnością sztucznych sieci neuronowych opisywanych w ramach niniejszej pracy. Przede wszystkim niepowodzenie leczenia w publikacji Sila i wsp. było definiowane w zupełnie odmienny sposób, to jest jako nawrót dolegliwości wymagający zastosowania antybiotyków i sterydów doustnych. Sukces terapii zanotowano w związku z tym aż u około 90% chorych. Co więcej, Sil, Mackay i Rowe-Jones oceniali wyniki po 5 latach, w związku z czym brak danych o efektach leczenia opisywanej przez nich grupy chorych po 3-6 miesiącach i po 12 miesiącach, co było przedmiotem prognozowania w niniejszej pracy. Ponadto podstawę predykcji w omawianej publikacji stanowiły tylko dwie zmienne, to jest obecność polipów nosa oraz zakres zmian zapalnych w obrazie tomografii komputerowej. W prezentowanej tu rozprawie rozważano znacznie szerszy zbiór zmiennych wejściowych. Najistotniejszy jest jednak fakt, że odsetek prawidłowych predykcji wynoszący 81.7% został podany dla grupy chorych analogicznej do zbioru uczącego, **nie wyodrębniano natomiast zbioru testowego**. Zdolność modelu do generalizacji oraz jego skuteczność w prognozowaniu wyniku u nowych pacjentów pozostaje w związku z tym nieznana.

Mając świadomość wymienionych wyżej różnic można dokonać jedynie orientacyjnego porównania opisanego wyżej modelu regresyjnego z sieciami neuronowymi przewidującymi nasilenie dolegliwości wyrażone w skali VAS (zmienna VAS\_bin), ponieważ jest to zmienna najbardziej zbliżona do przewidywanej przez model Sila i wsp. W toku eksperymentów prowadzonych w ramach niniejszej pracy uzyskiwano modele dokonujące nawet 100% prawidłowych predykcji zmiennej VAS\_bin w zbiorze uczącym, jednak nie były one brane pod uwagę, ponieważ analizując błędy popełniane przez te same modele w zbiorze walidacyjnym i testowym dowiedziono ich słabej zdolności do generalizacji. Brak wyodrębnienia zbioru testowego i ograniczenie eksperymentów jedynie do zbioru uczącego nieuchronnie doprowadziłby do błędnego wniosku, że sieci neuronowe są w stanie prawidłowo przewidzieć wynik leczenia u 100% chorych. Modele uznane ostatecznie za najefektywniejsze, a przy tym wystarczająco wiarygodne, wykazywały natomiast jakość testowania około 80% - i to jest wynik, jakiego można oczekiwać także u całkiem nowych pacjentów. Podsumowując ten fragment dyskusji można stwierdzić, że **efektywność sieci neuronowych była podobna jak w przypadku modelu regresyjnego, ale dodatkowo dowiedziono dla tych sieci, że wykazywały one cechy świadczące o tym, iż podobnej jakości predykcje mogą być dokonywane nie tylko dla grupy badanej, ale również dla pacjentów operowanych w przyszłości.**

Kolejne dwa modele, które opracowali Katotomichelakis i wsp. [182] służyły do prognozowania zaburzeń węchu notowanych 6 miesięcy po ESS w populacji greckiej. Pierwszy z nich, wykorzystujący **regresję logistyczną**, można stosunkowo łatwo porównać z opracowanymi w ramach niniejszej pracy klasyfikacyjnymi sieciami neuronowymi przewidującymi funkcję węchu po zabiegu w I grupie chorych. Zarówno model regresyjny, jak i sieci neuronowe, miały przewidywać, którzy z pacjentów zgłoszą prawidłowy węch po zabiegu. Katotomichelakis i wsp. posługiwali się w tym celu obiektywnymi testami olfaktorycznymi, a nie subiektywną oceną dokonywaną przez pacjenta, co nieco utrudnia porównanie, ponieważ wyniki badań subiektywnych i obiektywnych badań węchu nie zawsze bywają zgodne [10].

Model regresyjny prawidłowo zaklasyfikował 76% pacjentów, jednak podobnie jak w przypadku publikacji omawianej wyżej - dotyczyło to wyłącznie zbioru uczącego, ponieważ **zbiór testowy nie został przez autorów pracy wyodrębniony**. Dla porównania, najlepsza z sieci neuronowych rozważanych w tej pracy dokonywała 83% prawidłowych predykcji w zbiorze uczącym i 84% w zbiorze testowym. Warto zauważyć, że oba modele wykorzystywały bardzo podobne zestawy zmiennych. W przypadku sieci neuronowej były to: wiek, czas trwania objawów, stopień upośledzenia węchu przed zabiegiem, astma i nadwrażliwość na niesteroidowe leki przeciwzapalne. Model regresyjny brał pod uwagę dokładnie te same czynniki (poza astmą) oraz dodatkowo płeć, alergię, przebyte zabiegi oraz dysfunkcję zmysłu smaku. Porównanie efektywności obu typów modeli pozwala stwierdzić, że **sieci neuronowe wykazały w tym przypadku przewagę nad regresją logistyczną**. Ponadto wydaje się, że **istnieje pewien zespół czynników klinicznych, które pozwalają prognozować funkcję węchu u chorych z PZZP niezależnie od tego, jakiej populacji dotyczy badanie**.

Drugi z modeli opisanych w cytowanej wyżej publikacji, wykorzystujący **wieloczynnikową regresję liniową**, służył do przewidywania wyniku obiektywnych badań węchu wyrażonego w skali od 0 do 48. Ponieważ działanie tego modelu polegało na przewidywaniu konkretnej wartości liczbowej, a nie na klasyfikacji, jest on trudny do porównania z klasyfikacyjnymi sieciami neuronowymi zbudowanymi w ramach tej pracy. Warto jednak zaznaczyć, że przy jego pomocy udało się wyjaśnić około 70% zmienności wyników testów olfaktometrycznych. Wykorzystywane zmienne były takie same jak w opisanym wcześniej modelu wykorzystującym regresję logistyczną.

#### **7.4. Identyfikacja nie opisywanych dotychczas czynników warunkujących nasilenie dolegliwości u chorych z PZZP**

We wprowadzeniu do niniejszej pracy dotyczącej sztucznych sieci neuronowych zaznaczono, że **działanie tego rodzaju modeli nie poddaje się łatwo interpretacji** ze względu na trudności w prześledzeniu procesów przetwarzania danych w ich wnętrzu [221]. W związku z tym zasadniczo nie służą one do wyjaśniania natury modelowanych zjawisk. Wykorzystanie przez model określonego zestawu zmiennych wejściowych nie stanowi dowodu na to, że między opisywanymi przez nie czynnikami a zmienną wyjściową istnieje

rzeczywisty związek przyczynowo-skutkowy. Z tego powodu wszelkie próby interpretacji zachowań sieci, na przykład w celu identyfikacji czynników ryzyka mających wpływ na wyniki leczenia określonej choroby, wymagają daleko idącej ostrożności.

W toku badań przeprowadzonych dla potrzeb niniejszej pracy obserwowano jednak bardzo często, że w trakcie powtarzanych wielokrotnie eksperymentów polegających na poszukiwaniu optymalnych modeli, pewne koincydencje wejść i wyjść powtarzały się wyraźnie częściej niż inne. W szczególności obserwowano, że najefektywniejszych predykcji określonej zmiennej wyjściowej dokonywały modele wykorzystujące pewien konkretny (za każdym razem ten sam) zestaw zmiennych wejściowych. Te wybór najkorzystniejszych zmiennych wejściowych powtarzał się, mimo stosowania różnych metod doboru cech. Dzięki zastosowaniu metody testowania grup liczących pod 100 lub 1000 sieci, a nie pojedynczych modeli, możliwe było obserwowanie tego rodzaju tendencji w sposób do pewnego stopnia niezależny od efektów losowych, z założenia także wpływających na działanie sieci neuronowych.

Analiza opisanych powyżej zachowań sieci neuronowych upoważnia do wnioskowania, że wyodrębnione (często wybierane jako optymalne) zestawy zmiennych wejściowych miały szczególnie istotne znaczenie dla predykcji określonych zmiennych wyjściowych. Warto na marginesie tych rozważań zauważyć, że wykorzystanie sztucznych sieci neuronowych do przewidywania efektów leczenia stwarza szanse zauważenia również takich zależności, które byłyby trudne do odkrycia tradycyjnie stosowanymi metodami statystycznymi, (sprowadzającymi się w większości badań do prostych porównań grupy pacjentów narażonych na określony pojedynczy czynnik ryzyka z grupą nie wykazującą takiego narażenia). Obserwacja działania modeli korzystających z różnych wektorów wejściowych może mianowicie wskazywać na **istnienie pewnych zależności nieliniowych lub relacji wieloczynnikowych**, chociaż oczywiście nie wyjaśnia ich mechanizmu ani charakteru.

Poszukując najefektywniejszych modeli neuronowych dokonujących predykcji efektów leczenia chorych z PZZP zauważono, że w większości przypadków **istotne znaczenie prognostyczne ma oceniane subiektywnie nasilenie dolegliwości, zgłaszane przez pacjentów przed zabiegiem**. Wielokrotnie w toku prac uzyskiwano modele, które preferencyjnie wykorzystywały tego rodzaju dane, niekiedy nawet zupełnie pomijając informacje o charakterze obiektywnym (na przykład płeć, wiek, obecność chorób towarzyszących lub liczbę przebytych zabiegów).

Odnotowując tę obserwację należy zaznaczyć, że w dostępnym piśmiennictwie nie odnaleziono wzmianek o wykorzystywaniu przez innych autorów do celów predykcyjnych zmiennych opisujących subiektywną samoocenę dolegliwości dokonywaną przez pacjenta przed zabiegiem. W niektórych pracach autorzy analizowali jedynie zależność między nasileniem konkretnego objawu przed zabiegiem i po nim. Oceniano na przykład, czy stan węchu przed operacją wpływa na jego funkcjonowanie po leczeniu [111, 121]. Zauważono również, że przewlekłe zmęczenie po operacji zgłaszają głównie ci chorzy, którzy już wcześniej zgłaszali znaczną uciążliwość tego objawu [179] Prognozując nasilenie określonej dolegliwości nie brano natomiast pod uwagę innych objawów.

W niniejszej pracy zdecydowano natomiast podjąć próbę oceny wzajemnych związków między poszczególnymi objawami. Istnienie tego rodzaju zależności w odniesieniu do każdej z analizowanych zmiennych wyjściowych potwierdziły zarówno analizy statystyczne, jak i zachowanie testowanych modeli neuronowych. Wszystkie najefektywniejsze modele, zbudowane w ramach prezentowanych badań, wykorzystywały dane o nasileniu dolegliwości przez zabiegiem. Wielokrotnie stwierdzano wysoce istotne statystycznie zależności nawet pomiędzy takimi objawami, które nie wydawały się logicznie ze sobą powiązane (na przykład między nasileniem dolegliwości usznych lub zmęczeniem zgłaszanym przed zabiegiem, a szansą na ustąpienie spływania wydzieliny po tylnej ścianie gardła po operacji). Poniżej przedstawiono próbę wyjaśnienia tego zjawiska, które można uznać za dodatkowe spostrzeżenie, dokonane niejako przy okazji budowy predykcyjnych modeli neuronowych.

#### **7.4.1 Przewlekłe zapalenie zatok jako grupa schorzeń o różnej etiologii, patogenezie i manifestacji klinicznej**

Postępy w zakresie badań dotyczących PZZP dowiodły, że nie jest to pojedyncze schorzenie, jak sądzono wcześniej, ale grupa chorób o różnej etiologii i patogenezie. Z tego względu obecnie często mówi się o „przewlekłych zapaleniach zatok” w liczbie mnogiej. Nieustannie odkrywano nowe czynniki powodujące lub modulujące występowanie procesów zapalnych w jamie nosowej i zatokach przynosowych. Okazuje się, że u różnych pacjentów, których łączy wspólne rozpoznanie PZZP, w rzeczywistości mogą występować bardzo różne schorzenia, odmienne pod względem przebiegu klinicznego, a co za tym idzie również rokowania.

W obliczu pojawiających się nieustannie nowych doniesień naukowych dotyczących uwarunkowań PZZP trudno o jednolitą klasyfikację, która uwzględniałaby wszystkie znane warianty tej choroby. Najogólniej PZZP dzieli się na **dwie typy** o odmiennej manifestacji klinicznej: PZZP z polipami nosa i PZZP bez polipów nosa.

Warto zauważyć, że te dwie zasadnicze „odmiany” choroby cechują pewne charakterystyczne zespoły objawów. Pacjenci cierpiący z powodu PZZP z polipami nosa często skarżą się na trudności z oddychaniem przez nos, zaburzenia węchu oraz znaczną ilość wydzieliny z nosa, rzadko natomiast odczuwają bóle głowy. Tymczasem u chorych z PZZP bez polipów nosa właśnie bóle głowy i twarzy mogą stanowić najbardziej uciążliwe objawy choroby.

Wyróżnione wyżej podgrupy chorych różnią się również rokowaniem odnośnie wyników leczenia operacyjnego. U pacjentów, u których stwierdza się polipy nosa, częściej dochodzi do nawrotu objawów po operacji [28, 45, 46], jednak na ogół zgłaszają oni większą satysfakcję z efektów zabiegu niż chorzy bez polipów nosa [25, 29, 30, 46]. Po leczeniu operacyjnym chorzy, u których stwierdzono polipy, podają bardziej zauważalną poprawę drożności nosa, gorsze jest u nich natomiast rokowanie dotyczące na przykład poprawy węchu po ESS [108]. Z kolei u pacjentów bez polipów nosa można się spodziewać utrzymywania się bólu twarzy mimo przebytego zabiegu [46].

Przedstawione powyżej prawidłowości dowodzą, że występowanie u pacjenta pewnej „konstelacji” objawów pozwala zaklasyfikować go do podgrupy chorych o określonym

rokowaniu. Oznacza to, że **rodzaj dominujących objawów klinicznych może stanowić podstawę do prognozowania wyników leczenia.**

Zależności zaobserwowane w ramach prezentowanych w tej pracy badań sugerują, że na podstawie objawów klinicznych **wśród chorych z PZZP w rzeczywistości można zidentyfikować więcej niż dwie wyróżniane obecnie zasadnicze podgrupy pacjentów.** Modele neuronowe dokonywały prawidłowych predykcji nasilenia objawów klinicznych po zabiegu na podstawie informacji o nasileniu kilkunastu dolegliwości przed zabiegiem. Wynika z tego, że sieci neuronowe „dokonały odkrycia” istnienia w zbiorze przypadków uczących pewnych klas różniących się rokowaniem, możliwych do wyodrębnienia właśnie na podstawie różnic w nasileniu dolegliwości przed zabiegiem.

Powyższa obserwacja może stanowić podstawę do podjęcia dalszych badań w tym zakresie. W ramach niniejszej pracy odkryte zależności z założenia posłużyły wyłącznie do przewidywania wyników leczenia. W dalszym etapie prac celowa wydaje się próba szczegółowego zdefiniowania nieznanych dotąd, a różniących się rokowaniem podgrup pacjentów. Zbudowane dotychczas modele nie są najwłaściwszym narzędziem do tego typu analiz ze względu na wspomniane wyżej trudności w interpretacji ich działania (brak możliwości bezpośredniego prześledzenia, w jaki sposób w przypadku ogólnym na podstawie danych wejściowych ustalane są wartości danych wyjściowych). W celu wyodrębnienia poszukiwanych podtypów PZZP wskazana wydaje się próba zastosowania metod należących do dziedziny **analizy skupień**, na przykład samoorganizujących się sieci neuronowych (*Self Organizing Maps*, SOM). Tego rodzaju modele, zwane sieciami Kohonena, mają zdolność samoczynnego dostosowywania się do struktury danych i są z powodzeniem wykorzystywane właśnie w zadaniach automatycznego wyróżniania istniejących w zbiorze danych klas, nawet jeżeli użytkownik nie zna zasady, według której ma być dokonywany ten podział.

Po wyróżnieniu odpowiednich klas pacjentów na podstawie występujących u nich zestawów objawów klinicznych oraz rokowania, wskazane byłoby przeprowadzenie analizy, jakie inne cechy odróżniają od siebie wyodrębnione wcześniej podgrupy (etiologia PZZP, lokalizacja i rozległość zmian zapalnych w zatokach, choroby towarzyszące). Tego rodzaju postępowanie ma szansę doprowadzić do zidentyfikowania czynników obiektywnych, warunkujących odmienny przebieg choroby u różnych pacjentów. Zaproponowany kierunek badań stanowi w pewnym sensie odwrócenie przyjętego powszechnie schematu rozumowania, w którym jako punkt wyjścia najczęściej przyjmuje się wyniki badań obiektywnych. W ten sposób **wykorzystanie modeli neuronowych ma szansę przyczynić się do wyjaśnienia patogenezy zespołu chorób określanych obecnie wspólnym mianem PZZP.**

#### **7.4.2 Potencjalne pozarynologiczne uwarunkowania dolegliwości**

Fakt wykorzystywania przez skuteczne modele predykcyjne danych o objawach subiektywnych zgłaszanych przed zabiegiem może również świadczyć o tym, że występowanie pewnych zespołów dolegliwości potencjalnie sugeruje pozarynologiczne pochodzenie objawów. Najnowsze piśmiennictwo podaje na przykład, że u wielu chorych



bóle twarzy i głowy związane z obecnością stanu zapalnego w zatokach okazują się wynikać z przyczyn neurologicznych [10, 142]. U tych pacjentów przeprowadzenie operacji zatok przynosowych nie przynosi spodziewanej poprawy. Obserwacja działania sieci neuronowych przewidujących nasilenie bólu twarzy i bólu głowy po operacji sugeruje, że modele te mogły automatycznie „odkryć”, **w jaki sposób odróżnić osoby o pozarynologicznych uwarunkowaniach objawów od pozostałych**. Najistotniejsze dla predykcji okazały się informacje o uciążliwości przewlekłego zmęczenia oraz różnego rodzaju dolegliwości bólowych przed zabiegiem operacyjnym. Przepuszczalnie w trakcie uczenia sieci doszło do wyróżnienia pewnych wzorcowych zespołów objawów, mogących charakteryzować osoby, u których ból nie wynika z obecności stanu zapalnego w zatokach i w związku z tym może pozostawać niepodatny na leczenie rynologiczne.

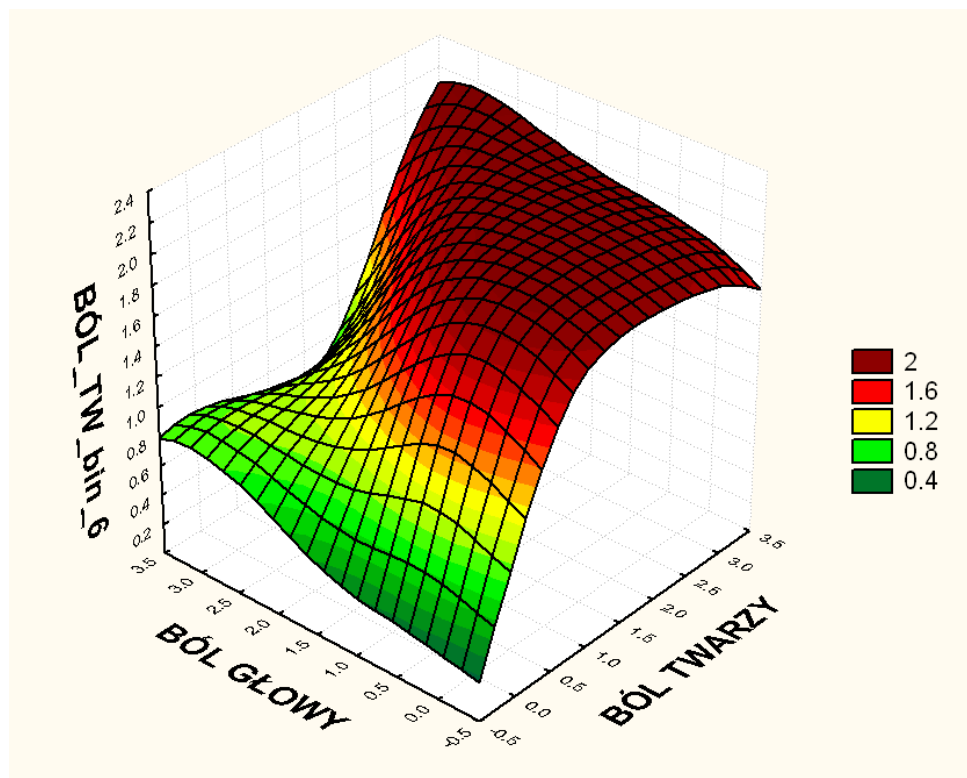
Należy przy tym dodać, że u każdego z chorych w badanych grupach wyniki badań obiektywnych potwierdzały obecność PZZP, ponieważ tylko tacy pacjenci zostali zakwalifikowani do ESS i włączeni do niniejszych badań. Można jednak zakładać, że u części z nich zmiany zapalne w obrębie zatok przynosowych jedynie współwystępowały z dolegliwościami bólowymi innego pochodzenia.

#### **7.4.3 Wpływ indywidualnych cech pacjenta na subiektywną ocenę nasilenia dolegliwości związanych z PZZP**

Silne zależności między skargami zgłaszanymi przed operacją a nasileniem objawów po zabiegu mogą wynikać ze złożonych uwarunkowań związanych z odczuwaniem dolegliwości przez konkretnego chorego. Uważa się, że obiektywnie występujące czynniki patofizjologiczne stanowią tylko jeden z elementów wpływających na uciążliwość ocenianych subiektywnie objawów. Niemniej istotna wydaje się interakcja między tymi czynnikami a układem nerwowym chorego oraz indywidualne cechy pacjenta, w tym również uwarunkowania psychologiczne. Przepuszczalnie jest to jedna z przyczyn obserwowanych często niezgodności między wynikami badań obiektywnych a subiektywną samooceną dolegliwości (na przykład między rzeczywistą drożnością nosa ocenioną w rynomanometrii a poczuciem niedrożności nosa zgłaszanym przez pacjenta) [10].

Na samoocenę dolegliwości wpływają najprawdopodobniej liczne czynniki charakteryzujące w sposób indywidualny poszczególnych pacjentów. Chorzy mogą między innymi różnić się między sobą progiem odczuwania bólu. W ramach niniejszej pracy obserwowano, że pacjenci skarżący się już przed zabiegiem na uciążliwe dolegliwości bólowe w rozmaitych lokalizacjach (głowa, twarz, czoło, ale także ucho czy zęby) na ogół odczuwali silniejsze bóle głowy i twarzy także po zabiegu. Dla zilustrowania tego zjawiska na rysunku 46 przedstawiono powierzchnię odpowiedzi modelu dokonującego predykcji bólu twarzy po zabiegu. Wykres ten pokazuje, że sieć neuronowa przewidywała największe nasilenie bólu twarzy u tych osób, które zgłaszały silny ból głowy i twarzy przed operacją, najmniejsze zaś u pacjentów, którzy przed zabiegiem nie skarżyli się na żadną z tych dolegliwości. Obecność wyłącznie silnego bólu głowy bez towarzyszącego bólu twarzy nie implikowała natomiast utrzymywania się omawianego objawu po leczeniu.

**Rysunek 46. Powierzchnia odpowiedzi modelu przewidującego nasilenie bólu twarzy po zabiegu operacyjnym w I grupie chorych (zależność zmiennej przewidywanej od nasilenia bólu twarzy i bólu głowy zgłaszanego przed zabiegiem).**



Można również zakładać, że niektórzy pacjenci wykazują szczególną tendencję do zwracania uwagi na pewne dolegliwości, które u większości osób nie powodują zauważalnego dyskomfortu. W praktyce klinicznej spotyka się na przykład chorych stale narzekających na sływanie wydzieliny po tylnej ścianie gardła, chociaż w świetle obiektywnych badań wydaje się, że ilość wydzieliny nie jest zwiększona w stosunku do stanu prawidłowego, a jedynie indywidualna wrażliwość chorego każe mu interpretować zjawisko fizjologiczne jako przyczynę trudnej do zniesienia dolegliwości. Z kolei chorzy, u których przed zabiegiem występowała znacznie obniżająca jakość życia niedrożność nosa, po jego udrożnieniu na drodze operacyjnej rzadko koncentrują się na mniej uciążliwych objawach, takich jak wspomniane wyżej sływanie wydzieliny. O wpływie uwarunkowań psychologicznych na samoocenę dolegliwości może także świadczyć tendencja do zgłaszania większego nasilenia objawów PZZP przez pacjentów z depresją [25, 80].

Istnienie opisanych powyżej zależności uzasadnia wykorzystanie danych dotyczących samooceny objawów przed operacją do predykcji ich pooperacyjnego nasilenia. Informacja o odpowiedziach udzielonych przez konkretnego pacjenta przed zabiegiem może stanowić podstawę do wnioskowania o występujących u niego indywidualnych uwarunkowaniach, mogących wpływać na odczuwanie dolegliwości związanych z PZZP również po operacji. Dane te warto zwłaszcza interpretować w kontekście badań obiektywnych, potwierdzających lub wykluczających występowanie klinicznych uwarunkowań zgłaszanych skarg. Celem

takiej analizy nie jest podważanie zaufania do odpowiedzi pacjenta, ale poznanie jego wrażliwości na pewne bodźce związane z chorobą, ponieważ od tej wrażliwości może w dużym stopniu zależeć efekt leczenia. Miarą sukcesu operacji z punktu widzenia pacjenta jest bowiem odczuwane przez niego usunięcie lub zmniejszenie nasilenia dolegliwości, a nie stopień normalizacji wyników badań obiektywnych.

Podsumowując powyższe rozważania można stwierdzić, że wprowadzenie nie wykorzystywanych dotychczas zmiennych wejściowych, dotyczących nasilenia dolegliwości zgłaszanego przed zabiegiem, a także potwierdzenie ich znaczenia w toku eksperymentów dotyczących predykcyjnych modeli neuronowych, jest oryginalnym elementem tej rozprawy, a ponadto daje podstawę do tego, aby zarekomendować ich rutynowe uwzględnianie do predykcji wyników leczenia PZZP.

## **7.5 Możliwości zastosowania opracowanych modeli neuronowych w praktyce klinicznej oraz perspektywy ich dalszego rozwoju**

Efektywność sieci neuronowych opracowanych w ramach tej pracy podsumowuje tabela 84, w której zestawiono jakości działania osiągnięte przez najlepsze modele dla każdej ze zmiennych wejściowych. Warto podkreślić, że w odróżnieniu od prac, których wyniki omawiano w kontekście przeglądu piśmiennictwa, wszystkie oceny jakości działania neuronowych modeli przedstawione w tej pracy oparte są na testowaniu działania modeli przy próbach prognozowania przypadków pochodzących ze zbioru testowego, rozłącznego w stosunku do zbiorów: uczącego i walidacyjnego, używanych wcześniej do formowania modelu.

Można założyć, że wartości ocen, uzyskane przez modele sieci przy ich testowaniu, nie są istotnie zawyżone, ponieważ w trakcie eksperymentów podjęto stosowne kroki mające zapobiegać selekcji modeli, dla których dane dotyczące efektywności były mało wiarygodne. W tym celu dokonano starannego wyboru zbioru testowego oraz zachowywano szczególną ostrożność w trakcie oceny modeli. W szczególności odrzucano te modele, których cechy wskazywały na niską rzeczywistą zdolność do generalizacji. Tam, gdzie osiągnięta jakość testowania wydawała się potencjalnie niereprezentatywna, wyraźnie to zaznaczono w tekście odpowiednich podrozdziałów. Można zatem założyć, że próby zastosowania opracowanych modeli dla nowych przypadków (tj. w celu prognozowania wyników leczenia u chorych operowanych w przyszłości) nie powinny dać wyników gorszych, niż to wynika z jakości testowania modeli przedstawionych w tej pracy.

Jednocześnie jednak należy podkreślić, że przedstawione wyniki można w pewnym sensie traktować zaledwie jako **dolną granicę** możliwości sztucznych sieci neuronowych w odniesieniu do analizowanego zagadnienia. Ze względu na specyfikę funkcjonowania tego rodzaju modeli, istnieją możliwości dalszego rozwijania opracowanych w ramach tej pracy sieci i potencjalnie otrzymania modeli o **znacznie wyższej efektywności**. Dzięki przeprowadzonym w tej pracy eksperymentom można też wskazać najbardziej obiecujące

kierunki kontynuacji badań. Najistotniejsze wnioski dotyczące tego zagadnienia omówiono poniżej.

### **7.5.1 Najlepsze uzyskane modele a modele optymalne**

Jak zaznaczono w rozdziale 4.13, prowadząc badania z użyciem sztucznych sieci neuronowych nigdy nie można mieć pewności, że uzyskany model jest najlepszym z możliwych do zbudowania dla określonego zbioru danych. Nie da się (w ogólnym przypadku) udowodnić, że zostało odnalezione minimum globalne funkcji błędu, a nie tylko pewne minimum lokalne. W trakcie poszukiwań możliwie najbardziej efektywnej sieci neuronowej wiele zależy od przypadku, na przykład od wstępnego rozlosowania wartości parametrów sieci przed rozpoczęciem uczenia. Można zatem wyobrazić sobie sytuację, w której na skutek korzystnego zbiegu okoliczności w trakcie kolejnych eksperymentów prowadzonych w dokładnie ten sam sposób i dla tej samej bazy danych, otrzymuje się przypadkowo znacznie lepszy model. Co więcej, ze względu na nieprzewidywalną, trudną do bezpośredniej interpretacji i kontrolowania specyfikę funkcjonowania sieci neuronowych, nawet drobna zmiana w poczynionych założeniach (na przykład odmienny sposób kodowania danych wejściowych) może nieoczekiwanie skutkować znaczącymi zmianami w działaniu modeli. Z tego względu poszukiwań najlepszego możliwego do uzyskania modelu w praktyce nigdy nie można uznać za zakończone.

W niniejszej pracy oceniano efektywność sieci neuronowych nie na podstawie jakości pojedynczych modeli, ale całych grup liczących po 100 lub 1000 sieci, dzięki czemu można było obserwować pewne generalne tendencje do poprawy lub pogarszania ich efektywności w zależności od dokonywanych zmian architektury. Na podstawie tych obserwacji ukierunkowywano dalsze działania. Tego rodzaju postępowanie zwiększało prawdopodobieństwo budowania stopniowo coraz lepszych modeli, jednak oczywiście nie dawało gwarancji otrzymania modelu optymalnego.

### **7.5.2 Wielkość zbioru uczącego**

Oceniając możliwość zastosowania sieci neuronowych do przewidywania wyników leczenia chorych z PZZP w praktyce klinicznej trzeba potraktować opisane w tej pracy badania jako eksperymenty prowadzone w niewielkiej skali – to znaczy jako swego rodzaju „doświadczenie laboratoryjne”. Dla porównania, modele rzeczywiście używane na co dzień przez lekarzy w praktyce klinicznej stanowią efekt prac prowadzonych „w skali przemysłowej”. Podstawą do ich stworzenia nie były bazy danych zebrane przez pojedynczą osobę na potrzeby badań naukowych (na przykład dla napisania artykułu w czasopiśmie czy pracy doktorskiej), ale niejednokrotnie doświadczenia całych towarzystw naukowych, dotyczące setek tysięcy pacjentów. Jako przykład można przypomnieć, że w przypadku wspomnianego we wstępie systemu SCORE do jego zbudowania potrzebne były to aż 3 miliony osobolat obserwacji [6]. Celem zaprezentowanych w tej pracy badań nie było w związku z tym przygotowanie sieci neuronowych gotowych do zastosowania w praktyce, ale **opracowanie metod budowania takich modeli** oraz przedstawienie dowodów na to, że wykorzystanie tego rodzaju narzędzi do realizacji założonych celów jest zasadne i efektywne. Uzyskane wyniki jasno wskazują, że wyniki leczenia operacyjnego chorych z PZZP są w

dużym stopniu przewidywalne. Ponadto dowiedziono, że wystarczy dysponować zaledwie podstawowymi informacjami na temat pacjenta, aby dokonać stosunkowo trafnej predykcji. Oznacza to, że **istnieje** możliwa do wymodelowania zależność między danymi, traktowanymi w tej pracy jako zmienne wejściowe, a przewidywanymi zmiennymi wyjściowymi.

Specyfika procesu uczenia sieci neuronowych powoduje, że na ogół uzyskiwane wyniki są tym lepsze, im większym zbiorem uczącym dysponuje badacz. Dowodem na poprawę działania testowanych w tej pracy sieci neuronowych wraz ze zwiększaniem zbioru uczącego były eksperymenty opisane w rozdziale 6.2.2.1, dotyczące zmiany proporcji w podzbiórach uczącym, walidacyjnym i testowym. Wykazano w nich, że zwiększenie zbioru uczącego zaledwie o kilka przypadków powodowało zauważalną poprawę efektywności modeli neuronowych. Podobne wnioski można wyciągnąć porównując przebieg eksperymentów dla liczniejszej grupy I i dla mniej licznej grupy II. Dla tej ostatniej grupy często szczególnie trudne było zbudowanie sieci neuronowych spełniających równocześnie wszystkie założone kryteria.

Zbiór danych zawierający informacje o tysiącach pacjentów prawdopodobnie nie wymagałby również stosowania wykorzystanych w tej pracy skomplikowanych metod wyodrębnienia podzbiorów uczącego, walidacyjnego i testowego. W przypadku tak obszernej bazy danych przypuszczalnie można byłoby uzyskać reprezentatywne podzbiory stosując po prostu podział losowy. Aby ocenić, czy losowanie rzeczywiście daje akceptowalne wyniki, należałoby jednak porównać działanie modeli kilkukrotnie rozlosowując przypadki do podzbiorów uczącego, walidacyjnego i testowego. Jeżeli dla każdego losowania modele uzyskiwałyby podobną efektywność, można by było założyć, że istnieje tendencja do odpowiednio „równomiernego” rozdzielenia przypadków między podzbiory. Natomiast duże różnice np. w jakości testowania zależnie od sposobu rozlosowania przypadków dowodziłyby, że zastosowana metoda niekiedy prowadzi do uzyskania niereprezentatywnych podzbiorów.

### **7.5.3 Liczba zmiennych wejściowych**

Wykorzystanie w przyszłych badaniach odpowiednio liczego zbioru przypadków uczących pozwoli na prawidłowe nauczenie modelu o większych rozmiarach, w szczególności posiadającego większą liczbę neuronów wejściowych. Dzięki temu będzie możliwe wykorzystanie większej liczby zmiennych wejściowych bez straty zdolności do generalizacji. Umożliwi to na modelowanie znacznie bardziej subtelnych aspektów zależności między zmiennymi wejściowymi a wyjściowymi.

Trzeba jednak zauważyć, że w praktyce klinicznej najchętniej wykorzystywane są takie modele, które nie wymagają podawania bardzo dużej liczby zmiennych wejściowych [222]. Jeżeli sieć neuronowa wymagałaby od użytkownika zgromadzenia bardzo licznych informacji na temat pacjenta, najprawdopodobniej nie znalazłaby zastosowania w praktyce, gdyż jej uruchomienie byłoby zbyt pracochłonne. Z tego powodu budując modele predykcyjne należy zachować umiar w odniesieniu do liczby ostatecznie wykorzystywanych zmiennych wejściowych, nawet jeśli ograniczenie ich ilości skutkuje pewnym pogorszeniem jakości predykcji.

#### **7.5.4 Nowe zmienne wejściowe**

W niniejszej pracy założono, że budowane modele neuronowe mają bazować wyłącznie na takich zmiennych, których pozyskanie jest łatwe dla każdego lekarza opiekującego się pacjentami przygotowywanymi do operacji endoskopowej zatok przynosowych. Z tego powodu, jak podkreślano w rozdziale „Materiał i metoda”, na przykład w celu określenia obecności chorób towarzyszących posługiwano się wyłącznie danymi z wywiadu. Potencjalnie dokładniejszych predykcji można by było dokonywać wykorzystując bardziej wiarygodne informacje (wyniki testów alergicznych, spirometrii, pH-metrii itd.) Modele wymagające podania takich danych byłyby jednak niemożliwe do zastosowania u większości chorych, ponieważ aktualnie w warunkach polskich uzyskanie wyników wszystkich wspomnianych badań jest zbyt czasochłonne i kosztowne, w związku z czym nie u każdego pacjenta są one rutynowo wykonywane przed kwalifikacją do ESS.

Analizując piśmiennictwo zauważono, że istotną wartość rokowniczą może mieć natomiast informacja dotycząca lokalizacji zmian zapalnych w zatokach przynosowych, stwierdzanych w tomografii komputerowej, a nie tylko ocena ich sumarycznej rozległości [150]. Tego rodzaju dane są dostępne u każdego pacjenta przygotowywanego do operacji endoskopowej zatok, ponieważ wykonanie tomografii komputerowej przed operacją jest obligatoryjne. Włączenie zmiennych wejściowych opisujących, w których zatokach stwierdza się obecność polipów lub pogrubiałej śluzówki, mogłoby umożliwić bardziej szczegółowe wyodrębnienie podgrup chorych różniących się pod względem przebiegu klinicznego PZZP. Aby wykorzystać te zmienne niezbędne byłoby jednak wprowadzenie co najmniej kilku dodatkowych neuronów wejściowych, co prowadziłoby do zwiększenia liczby połączeń w sieci i nieodzownie wymagałoby liczniejszego zbioru uczącego.

#### **7.5.5 Uwzględnienie doświadczeń innych ośrodków**

Baza danych wykorzystana na potrzeby niniejszej pracy zawierała wyłącznie informacje o pacjentach operowanych w Klinice Otolaryngologii w Krakowie. Zbudowane modele neuronowe powinny zatem efektywnie przewidywać wyniki leczenia kolejnych chorych, którzy trafią do tego samego ośrodka, zostaną zoperowani tą samą metodą, a następnie będą pozostawali w analogicznie prowadzonej opiece pooperacyjnej. Trzeba jednak zaznaczyć, że pacjenci leczeni w innych ośrodkach mogą uzyskiwać inne wyniki operacji, na przykład ze względu na różnice w doświadczeniu chirurgów lub odmienne zasady farmakoterapii pooperacyjnej. W przypadku chorych z PZZP nie obowiązują ściśle zasady prowadzenia leczenia, zwłaszcza po przebytych zabiegach, a szczegóły terapii zależą od indywidualnego doświadczenia lekarzy oraz od tradycji panującej w poszczególnych szpitalach.

Sieci neuronowe nabywają zdolności rozwiązywania wyłącznie zadań podobnych do tych, z którymi zostały zapoznane w procesie uczenia. W efekcie modele, które prawidłowo przewidują wyniki leczenia prowadzonego według pewnego schematu, mogą wykazywać mniejszą skuteczność prognozowania w odniesieniu do pacjentów leczonych w nieco odmienny sposób. Zbudowanie modeli mających zdolność prawidłowego przewidywania

wyników leczenia PZZP niezależnie od tego, gdzie konkretny pacjent będzie operowany, wymaga badań wielośrodkowych. Nie można wykluczyć, że wśród zmiennych wejściowych niezbędnych dla tych modeli będą musiały się znaleźć dane dotyczące specyfiki leczenia w konkretnym ośrodku (na przykład kompetencji operującego chirurga).

### **7.5.6 Inne typy modeli neuronowych**

W ramach prezentowanych badań do rozwiązania analizowanego problemu wykorzystywano wyłącznie sieci neuronowe jednego typu, tj. perceptrony trójwarstwowe, którym stawiano zadania o charakterze klasyfikacji. Zdecydowano ograniczyć poszukiwania do tak wąskiego zakresu, aby przedstawiony w pracy wywód nie tracił przejrzystości. Zrezygnowano z omawiania testów dotyczących innych rodzajów modeli, ponieważ w tym przypadku prezentacja wyników musiałaby być bardzo pobieżna lub znacznie zwiększyłyby się rozmiary pracy. Tym niemniej w dziedzinie sztucznych sieci neuronowych istnieje jeszcze wiele obszarów, które wydają się warte eksploracji w kontekście prognozowania wyników leczenia chorych z PZZP.

W odróżnieniu od opisanych w tej pracy sieci klasyfikacyjnych, sieci regresyjne mają za zadanie przewidywać konkretną wartość zmiennej wyjściowej (na przykład nasilenie objawów w skali VAS), a nie tylko przyporządkować przypadki do jednej z zadanych klas. Ich zastosowanie nie wymaga też przekształcania zmiennych wyjściowych w zmienne binarne. Można dzięki temu uniknąć arbitralnych decyzji dotyczących „progu odcięcia”. Jak opisano wcześniej, tego rodzaju narzucone z góry założenia zmniejszają zaufanie użytkowników do modelu ze względu na świadomość, że bardzo podobni do siebie chorzy mogą być przyporządkowywani do różnych klas jeżeli występują u nich wartości analizowanej zmiennej bliskie progu odcięcia [222]. Jak wskazuje doświadczenie autorki, predykcja wyników leczenia przy pomocy sieci regresyjnych jest jednak znacznie bardziej skomplikowana niż opisane w tej pracy zadania klasyfikacyjne. Ponadto uzyskane wyniki są nieco trudniejsze do oceny i interpretacji (dla sieci klasyfikacyjnych łatwo przedstawić jakość testowania jako odsetek prawidłowo zaklasyfikowanych przypadków).

Nie można również wykluczyć, że analizowane zadania mogłyby być z powodzeniem rozwiązywane przez sieci o innych architekturach niż perceptrony wielowarstwowe. Wstępne doświadczenia z użyciem sieci RBF (*Radial Basis Function*) oraz sieci probabilistycznych nie okazały się jednak zachęcające.

### **7.5.7 Odmienne techniki modelowania**

W niniejszej pracy do modelowania analizowanych zależności wykorzystano sztuczne sieci neuronowe, ponieważ ich cechy wydawały się dobrze odpowiadać stawianemu zadaniu. Za istotne zalety sieci neuronowych uznano przede wszystkim brak konieczności dokonywania wstępnych założeń na temat charakteru modelowanej zależności (a ten charakter w odniesieniu do wyników leczenia PZZP pozostaje nieznanym) oraz możliwości modelowania zależności wielowymiarowych i nieliniowych (których spodziewano w przypadku badanej zależności na podstawie przeglądu literatury).

We wprowadzeniu wspomniano o istnieniu wielu innych metod modelowania, takich jak na przykład modele statystyczne (w szczególności modele regresyjne). Ich skuteczność była porównywana z efektywnością sieci neuronowych przez wielu badaczy, przy czym w zależności od analizowanego zadania wykazywano przewagę jednej lub drugiej metody, bądź też stwierdzano ich równoważność [219, 221]. W ramach tej pracy przedstawiano efektywność sieci neuronowych jedynie na tle prostej analizy danych historycznych, natomiast nie dokonywano porównań z narzędziami statystycznymi służącymi do analiz wielowymiarowych. Dobre wyniki prognozowania uzyskane z zastosowaniem sztucznych sieci neuronowych zachęcają do podjęcia prób budowania również innych typów modeli, wykorzystujących zestawy zmiennych wejściowych, których wartość predykcyjną wykazano w niniejszej pracy.

## **7.6 Obiektywne ograniczenia możliwości modeli predykcyjnych w PZZP**

Powyżej omówiono, w jaki sposób wyniki uzyskane w ramach prezentowanych badań mogą ukierunkować dalszy rozwój modeli neuronowych, tak aby możliwe było ich powszechne użytkowanie w praktyce klinicznej jako narzędzi wspomagających decyzje terapeutyczne dotyczące chorych z PZZP. Dowodzone, że przeprowadzone eksperymenty mogą posłużyć jako punkt wyjścia do zbudowania sieci neuronowych dokonujących znacznie bardziej wiarygodnych predykcji, pod warunkiem zaangażowania w projekt dużych zespołów badawczych i zebrania informacji o dostatecznie licznej grupie pacjentów.

Należy jednak mieć świadomość, że zbudowanie modelu dokonującego trafnych predykcji w 100% przypadków nie jest wykonalne. Jak zawsze w naukach medycznych istnieje margines niepewności, którego nie da się usunąć niezależnie od stosowanych metod prognozowania. Poniżej wypunktowano najważniejsze czynniki zaburzające proces predykcji wyników leczenia chorych z PZZP.

### **7.6.1 Czynniki oddziałujące w okresie pooperacyjnym**

Aby model prognostyczny mógł wspomagać decyzje kliniczne dotyczące kwalifikacji do operacji, musi oczywiście dokonywać predykcji jeszcze przed przeprowadzeniem zabiegu. Z tego powodu niezbędne dla modelu dane wejściowe muszą być dostępne na etapie przedoperacyjnym, a wykorzystanie zmiennych opisujących przebieg pooperacyjny mija się z celem. Tym niemniej opieka pooperacyjna i stosowanie się pacjenta do zaleceń odgrywa istotną rolę w utrzymaniu efektów leczenia i zapobieganiu nawrotom dolegliwości. Również nieprzewidziane infekcje jamy nosowej i zatok przynosowych mogą prowadzić do nasilenia objawów PZZP. Z tego powodu otrzymanie modelu wykazującego 100% prawidłowych predykcji w zbiorze testowym nie jest dowodem osiągnięcia ostatecznego sukcesu, ale raczej powinno wzbudzić wątpliwości co do prawidłowości przeprowadzonego testowania.



### **7.6.2 Przypadki nietypowe**

Różnorodność sytuacji klinicznych spotykanych w praktyce jest tak duża, że zawsze należy liczyć się z możliwością napotkania przypadków zupełnie odmiennych od uwzględnionych w zbiorze uczącym. U niektórych chorych dolegliwości mogą być uwarunkowane czynnikami nie uwzględnianymi w rutynowej diagnostyce. Wśród chorych badanych w ramach niniejszej pracy znalazł się na przykład pacjent cierpiący z powodu PZZP, u którego występowały zaburzenia węchu nie będące skutkiem stanu zapalnego w zatokach, ale spowodowane przebyciem w przeszłości urazem głowy, który doprowadził do uszkodzenia włókien nerwowych. Był to jedyny taki pacjent w całej bazie danych. Należy oczekiwać, że dokonywane dla niego predykcje poprawy węchu po zabiegu muszą być nieadekwatne, ponieważ występujące u niego zaburzenia nie wynikają z uwarunkowań rynologicznych, które sieci neuronowe uczyły się modelować na podstawie danych dostępnych w zbiorze uczącym.

### **7.6.3 Zmienność dolegliwości w czasie**

Przypuszczalnie najistotniejsze ograniczenie możliwości predykcji wyniku leczenia chorych z PZZP stanowi fakt, że nasilenie poszczególnych dolegliwości po zabiegu operacyjnym nie jest wartością stałą. Objawy zmieniają się z upływem czasu w przebiegu gojenia oraz pod wpływem stosowanej farmakoterapii. U niektórych chorych stopniowo dochodzi do pogorszenia funkcji zatok i nawrotu dolegliwości po kilku miesiącach lub latach, mimo uzyskania poprawy bezpośrednio po zabiegu operacyjnym. W niniejszej pracy wykazano, że istnieją różnice między pacjentami badanymi 3-6 miesięcy po zabiegu i 12 miesięcy po zabiegu (choć wyciągając wnioski dotyczące porównywania tych wyników należy zachować ostrożność, ponieważ nie dotyczyły one tej samej grupy chorych). Z tego powodu prognozowanie nasilenia dolegliwości dotyczące pojedynczego punktu w czasie (na przykład dokładnie 12 miesięcy po operacji) może nie stanowić dostatecznie wiarygodnej podstawy do decyzji o przeprowadzeniu zabiegu lub odstąpieniu od niego.

Sugerowanym rozwiązaniem mogłaby być próba tworzenia modeli dokonujących predykcji nie dla pojedynczego punktu w czasie, ale na przykład dla kolejnych miesięcy czy lat po zabiegu. Tego typu modele powinny wykreślać krzywe ilustrujące przewidywaną u konkretnego chorego zmienność objawów w czasie, co pozwalałoby również prognozować ewentualne nawroty choroby. Zadanie to przypomina wykreślanie krzywych przeżycia, a tego rodzaju problemy były już z powodzeniem rozwiązywane przy pomocy sieci neuronowych w odniesieniu do innych zagadnień klinicznych [227].

Można założyć, że u wielu chorych byłoby możliwe przewidywanie pooperacyjnego przebiegu PZZP na podstawie czynników znanych już przed operacją. Sieci neuronowe powinny z powodzeniem modelować na przykład tendencję do nawrotów niedrożności nosa u pacjentów z astmą aspirynową [19, 53, 66, 67, 92]. Niestety, na nasilenie dolegliwości mogą mieć wpływ również czynniki niemożliwe do uwzględnienia w okresie przedoperacyjnym. W przebiegu przewlekłego stanu zapalnego może dochodzić do zaostrzeń spowodowanych na przykład przez wspomniane wyżej infekcje czy sezonowy kontakt z alergenami. Z tego powodu ten sam chory może podawać różne nasilenie dolegliwości nawet w odstępie kilku

dni, zależnie od okoliczności, w jakich przeprowadzana jest ocena. W efekcie nierealne jest oczekiwanie od modelu, aby predykcje dokonywane na podstawie takich danych wejściowych były stuprocentowo precyzyjne.

## **7.7 Podsumowanie – znaczenie modeli prognozujących wyniki leczenia operacyjnego chorych z PZZP**

Efektywne modele predykcyjne mogą być skutecznymi narzędziami wspomagającymi decyzje terapeutyczne lekarza i dającymi podstawy do wyrażania świadomej zgody na leczenie przez pacjenta. Obecnie przed podjęciem decyzji o poddaniu się operacji pacjenci na ogół są informowani, jaki odsetek chorych uzyskuje poprawę po zabiegu. Posługując się sztucznymi sieciami neuronowymi lekarz mógłby przedstawić choremu bardziej szczegółową prognozę uwzględniającą występujące u niego indywidualne uwarunkowania wpływające na spodziewany efekt terapii.

Większość pacjentów poddaje się leczeniu ze względu na jedną lub kilka dolegliwości, od której ma nadzieję się uwolnić [3]. Dzięki modelom predykcyjnym można by było ocenić, czy u danej osoby należy się spodziewać ustąpienia określonych objawów, czy też istnieje duże ryzyko ich utrzymywania się mimo przeprowadzenia operacji. Wprowadzenie tego typu modeli do codziennej praktyki klinicznej powinno ograniczyć liczbę zabiegów kończących się niepowodzeniem dzięki wspieraniu decyzji o odstąpieniu od operacji u chorych, u których prawdopodobieństwo osiągnięcia oczekiwanych rezultatów na skutek tego rodzaju terapii jest niewielkie. U tych osób preferowane powinno być odpowiednie leczenie zachowawcze.

Ponadto udzielenie choremu kwalifikowanemu do operacji rzetelnej informacji o tym, na jakie efekty powinien liczyć, a jakich nie ma podstaw się spodziewać, może pomóc uniknąć nieuzasadnionych roszczeń pacjentów, w opinii których wynik leczenia okazuje się niezgodny z wcześniejszymi wyobrażeniami. Odstąpienie od operacji u osób mających niewielkie szanse na wyniesienie z niej korzyści pozwoliłoby również na lepszą dystrybucję środków i łatwiejszy dostęp do leczenia zabiegowego dla tych chorych, u których rzeczywiście można uzyskać zadowalające efekty na drodze operacyjnej.

Neuronowe modele predykcyjne mają szansę znaleźć zastosowanie w codziennej praktyce wszystkich lekarzy opiekujących się chorymi z przewlekłym zapaleniem zatok przynosowych. Gotowe modele mogą zostać bez trudu udostępnione w formie prostych w obsłudze programów do zainstalowania na dowolnym komputerze lub nawet telefonie komórkowym. W związku z tym wydaje się, że najważniejszą barierą, jaka musi zostać pokonana zanim sztuczne sieci neuronowe rzeczywiście zaczną wspomagać pracę klinicystów, nie są potencjalne problemy merytoryczne czy techniczne, ale nieufność potencjalnych użytkowników w stosunku do nowego, nietypowego narzędzia.

## 8. Wnioski

- Operacja endoskopowa zatok przynosowych pozwala uzyskać poprawę u większości chorych z PZZP, jednak u niektórych pacjentów zabieg nie przynosi oczekiwanych efektów. Przewidywanie wyniku leczenia operacyjnego umożliwia indywidualizację terapii i dostosowanie jej do potrzeb konkretnego pacjenta. Modele predykcyjne mogą wspomagać decyzję lekarza, dostarczając danych o potencjalnej przewadze leczenia operacyjnego lub zachowawczego u określonego chorego, a także pomagają pacjentowi zrozumieć, jakie są spodziewane efekty leczenia. W pracy wykazano, że przewidywanie wyniku leczenia operacyjnego może być skutecznie przeprowadzone przy pomocy sztucznych sieci neuronowych.
- Przy pomocy sztucznych sieci neuronowych zbudowanych w ramach niniejszej pracy udało się z powodzeniem przewidywać, jakie będzie pooperacyjne nasilenie poszczególnych dolegliwości związanych z PZZP u 70-94% pacjentów.
- Niektóre z analizowanych objawów okazały się łatwiej przewidywalne od pozostałych. U niemal wszystkich pacjentów udało się prawidłowo przewidzieć, czy w ciągu kilku miesięcy po zabiegu nastąpi zauważalna poprawa w zakresie ogólnego nasilenia objawów PZZP. Najtrudniejsze okazało się prognozowanie, na ile uciążliwa może być po operacji obecność wydzieliny z nosa.
- Do skutecznego prognozowania wystarczające okazały się podstawowe, łatwo dostępne informacje dotyczące pacjentów, co pozytywnie rokuje jeśli chodzi o możliwości zastosowania tego typu modeli w codziennej praktyce klinicznej. W wyniku przeprowadzenia badań opisanych w pracy wskazano te zmienne, które wydają się najistotniejsze dla predykcji poszczególnych objawów PZZP.
- Sztuczne sieci neuronowe testowane w niniejszej pracy potwierdziły oczekiwaną od tego typu narzędzi zdolność modelowania wielowymiarowych zależności nieliniowych.
- Porównanie sztucznych sieci neuronowych z modelami regresyjnymi opisywanymi w literaturze sugeruje, że sieci neuronowe rozwiązują rozważane zadania z porównywalną lub większą efektywnością.
- Eksperymenty pozwoliły zidentyfikować nie wykorzystywane dotychczas zmienne, które wydają się mieć istotniejsze znaczenie dla predykcji ocenianych subiektywnie wyników ESS niż tradycyjnie stosowane dotychczas w celach prognostycznych dane o charakterze obiektywnym. Są to zmienne opisujące nasilenie dolegliwości subiektywnie oceniane przez pacjenta przed zabiegiem. Wykorzystanie tych zmiennych może być rekomendowane we wszelkich zadaniach dotyczących predykcji wyników leczenia chorych z PZZP.
- Podjęcie próby zbudowania modeli predykcyjnych analogicznych do przedstawionych w niniejszej pracy z wykorzystaniem większych zbiorów przypadków uczących najprawdopodobniej umożliwi dalsze udoskonalenie sieci neuronowych służących do predykcji wyników leczenia operacyjnego chorych z PZZP, jakkolwiek żaden model

predykcyjny wykorzystujący jedynie dane dostępne przed operacją nie zagwarantuje osiągnięcia 100% prawidłowych predykcji.

- Za najważniejsze kierunki dalszych badań uznano:
  - zebranie znacznie większej liczby przypadków uczących (najlepiej w ramach badań wieloośrodkowych) i dalsze doskonalenie uzyskanych w ramach tej pracy sieci neuronowych, przede wszystkim poprzez wykorzystanie większej liczby danych wejściowych oraz testowanie sieci neuronowych o różnych architekturach,
  - próbę przewidywania pooperacyjnych zmian nasilenia objawów PZZP dla wielu kolejnych punktów w czasie i kreślenia „krzywych przebiegu choroby”, umożliwiających między innymi prognozowanie zaostrzeń i nawrotów,
  - po uzyskaniu modeli o dostatecznie wysokiej efektywności rozpropagowanie ich wśród praktykujących lekarzy w formie łatwych w obsłudze programów komputerowych, możliwych do zainstalowania na dowolnym komputerze czy telefonie komórkowym, a następnie monitorowanie opinii o użyteczności tych aplikacji,
  - próbę wyodrębnienia podgrup chorych różniących się objawami i rokowaniem (np. przy pomocy samoorganizujących się sieci neuronowych), co może się przyczynić do lepszego zrozumienia mechanizmów PZZP.

## 9. Streszczenie

**Wykorzystanie sztucznych sieci neuronowych w prognozowaniu efektów leczenia operacyjnego chorych z przewlekłym zapaleniem zatok przynosowych.**

**Wprowadzenie.** Leczenie operacyjne prowadzi do zmniejszenia nasilenia objawów przewlekłego zapalenia zatok przynosowych (PZZP) u większości, ale nie u wszystkich chorych. Przewidywanie wyniku zabiegu może wspomagać wybór optymalnej metody leczenia u poszczególnych pacjentów.

**Cel pracy.** Celem pracy było zbudowanie modeli predykcyjnych pozwalających w sposób zindywidualizowany przewidywać wynik leczenia operacyjnego chorych z PZZP.

**Material i metoda.** Grupę badaną stanowiło 183 chorych operowanych metodą endoskopową z powodu PZZP. Ocena przed operacją obejmowała wywiad, badanie laryngologiczne, tomografię komputerową i kwestionariusz samooceny objawów. Nasilenie dolegliwości oceniano ponownie po 3-6 i 12 miesiącach po operacji. Do przewidywania wyników leczenia na podstawie danych uzyskanych przed zabiegiem wykorzystano sztuczne sieci neuronowe (samouczące się systemy informatyczne wzorowane na budowie i funkcji układu nerwowego).

**Wyniki.** Najlepsze modele prawidłowo przewidywały zmianę nasilenia dolegliwości po zabiegu operacyjnym u 94% chorych. W odniesieniu do poszczególnych objawów PZZP uzyskano 70-91% trafnych predykcji. Dla skutecznego prognozowania wystarczające były dane uzyskane w ramach rutynowej diagnostyki przedoperacyjnej.

**Wnioski.** Sztuczne sieci neuronowe są efektywnymi narzędziami do prognozowania wyników leczenia operacyjnego chorych z PZZP i mają szansę znaleźć zastosowanie w codziennej praktyce klinicznej.

## 10. Summary

### **Artificial neural networks for prediction of outcome after surgery for chronic rhinosinusitis.**

**Introduction.** Surgical treatment reduces the symptoms of chronic rhinosinusitis (CRS) in most, but not all patients. The choice of optimal treatment modality for each patient can be aided by outcome prediction.

**Aim.** The aim of the study was to build models for individual outcome prediction in patients with CRS.

**Material and methods.** The study group comprised of 183 patients who underwent endoscopic sinus surgery for CRS. The preoperative evaluation included anamnesis, laryngological examination, computed tomography and self-assessment of symptoms. The symptoms were reevaluated 3-6 and 12 months after surgery. The results of treatment were predicted using artificial neural networks (self-learning systems inspired by the structure and function of the nervous system).

**Results.** The best obtained models provided correct prognoses of the postoperative symptom reduction in 94% of patients. The individual symptoms were accurately predicted in 70-91% of patients. Effective prognostication did not require data exceeding the routine preoperative investigations.

**Conclusions.** Artificial neural networks are effective tools for the prediction of outcome after surgery for chronic rhinosinusitis and may be adopted in everyday clinical practice.

# 11. Piśmiennictwo

1. Burke HB. The power of prediction. *Cancer*. 2008 Sep 1;113(5):890-2.
2. Ohno-Machado L. Modeling medical prognosis: survival analysis techniques. *J Biomed Inform*. 2001 Dec;34(6):428-39.
3. Chester AC, Antisdell JL, Sindwani R. Symptom-specific outcomes of endoscopic sinus surgery: A systematic review. *Otolaryngol Head Neck*. 2009 May;140(5):633-9.
4. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010 Jun;17(6):1471-4.
5. Wierzbicka M, Szyfter W, Bień S, Maciejewski B, Składowski K, Milecki P. Zalecenia diagnostyczno-terapeutyczne dla wybranych nowotworów głowy i szyi. *Postępy w chirurgii głowy i szyi*. 2006;5 (supl. 1).
6. De Backer G, Ambrosioni E, Borch-Johnsen K, Brotons C, Cifkova R, Dallongeville J, Ebrahim S, Faergeman O, Graham I, Mancia G, Manger Cats V, Orth-Gomer K, Perk J, Pyörälä K, Rodicio JL, Sans S, Sansoy V, Sechtem U, Silber S, Thomsen T, Wood D. European guidelines on cardiovascular disease prevention in clinical practice. Third Joint Task Force of European and Other Societies on Cardiovascular Disease Prevention in Clinical Practice. *Eur Heart J*. 2003 Sep;24(17):1601-10.
7. Nicolaides KH, Węgrzyn P. Badanie ultrasonograficzne między 11(+0)-13(+6) tygodniem ciąży. Londyn: Fetal Medicine Foundation; 2004.
8. Lee ET, Wang JW. *Statistical methods for survival data analysis*: J. Wiley; 2003.
9. Cox DR, Oakes D. *Analysis of survival data*: Chapman & Hall; 1994.
10. Fokkens WJ, Lund VJ, Mullol J, Bachert C, Alobid I, Baroody F, Cohen N, Cervin A, Douglas R, Gevaert P, Georgalas C, Goossens H, Harvey R, Hellings P, Hopkins C, Jones N, Joos G, Kalogjera L, Kern B, Kowalski M, Price D, Riechelmann H, Schlosser R, Senior B, Thomas M, Toskala E, Voegels R, Wang de Y, Wormald PJ. European Position Paper on Rhinosinusitis and Nasal Polyps 2012. *Rhinol Suppl*. 2012 Mar(23):3 p preceding table of contents, 1-298.
11. Lim M, Lew-Gor S, Darby Y, Brookes N, Scadding G, Lund VJ. The relationship between subjective assessment instruments in chronic rhinosinusitis. *Rhinology*. 2007 Jun;45(2):144-7.
12. Lanza DC, Kennedy DW. Adult rhinosinusitis defined. *Otolaryngol Head Neck Surg*. 1997 Sep;117(3 Pt 2):S1-7.
13. Hastan D, Fokkens WJ, Bachert C, Newson RB, Bislimovska J, Bockelbrink A, Bousquet PJ, Brozek G, Bruno A, Dahlén SE, Forsberg B, Gunnbjörnsdóttir M, Kasper L, Krämer U, Kowalski ML, Lange B, Lundbäck B, Salagean E, Todo-Bom A, Tomassen P, Toskala E, van Drunen CM, Bousquet J, Zuberbier T, Jarvis D, Burney P. Chronic rhinosinusitis in Europe – an underestimated disease. A GA2LEN study. *Allergy*. 2011;66(9):1216-23.
14. Lund VJ, Mackay IS. Staging in rhinosinusitis. *Rhinology*. 1993 Dec;31(4):183-4.
15. Metson R, Gliklich RE, Stankiewicz JA, Kennedy DW, Duncavage JA, Hoffman SR, Ohnishi T, Terrell JE, White PS. Comparison of sinus computed tomography staging systems. *Otolaryngol Head Neck Surg*. 1997 Oct;117(4):372-9.
16. Oluwole M, Russell N, Tan L, Gardiner Q, White P. A comparison of computerized tomographic staging systems in chronic sinusitis. *Clin Otolaryngol Allied Sci*. 1996 Feb;21(1):91-5.

17. Lund VJ, Kennedy DW. Staging for rhinosinusitis. *Otolaryngol Head Neck Surg.* 1997 Sep;117(3 Pt 2):S35-40.
18. Krzeski A, Kennedy DW. Zapalenie zatok przynosowych. W: Krzeski A, Janczewski G, red. *Choroby nosa i zatok przynosowych.* Wrocław: Urban & Partner; 2003. s. 203-18.
19. Krzeski A, Gromek I. *Zapalenia zatok przynosowych.* Gdańsk: Via Medica; 2008.
20. Krzeski A. Patofizjologia zapaleń zatok przynosowych. W: Krzeski A, Janczewski G, red. *Choroby nosa i zatok przynosowych.* Wrocław: Urban & Partner; 2003. s. 195-202.
21. van Oene CM, van Reij EJ, Sprangers MA, Fokkens WJ. Quality-assessment of disease-specific quality of life questionnaires for rhinitis and rhinosinusitis: a systematic review. *Allergy.* 2007 Dec;62(12):1359-71.
22. Piccirillo JF, Merritt MG, Jr., Richards ML. Psychometric and clinimetric validity of the 20-Item Sino-Nasal Outcome Test (SNOT-20). *Otolaryngol Head Neck Surg.* 2002 Jan;126(1):41-7.
23. Gliklich RE, Metson R. Effect of sinus surgery on quality of life. *Otolaryngol Head Neck Surg.* 1997 Jul;117(1):12-7.
24. Benninger MS, Senior BA. The development of the Rhinosinusitis Disability Index. *Arch Otolaryngol Head Neck Surg.* 1997 Nov;123(11):1175-9.
25. Smith TL, Mendolia-Loffredo S, Loehrl TA, Sparapani R, Laud PW, Nattinger AB. Predictive factors and outcomes in endoscopic sinus surgery for chronic rhinosinusitis. *Laryngoscope.* 2005 Dec;115(12):2199-205.
26. Terris MH, Davidson TM. Review of published results for endoscopic sinus surgery. *Ear Nose Throat J.* 1994 Aug;73(8):574-80.
27. Lund VJ. Evidence-based surgery in chronic rhinosinusitis. *Acta Otolaryngol.* 2001 Jan;121(1):5-9.
28. Kim HY, Dhong HJ, Chung SK, Chung YJ, Min JY. Prognostic factors of pediatric endoscopic sinus surgery. *Int J Pediatr Otorhinolaryngol.* 2005 Nov;69(11):1535-9.
29. Hopkins C, Slack R, Lund V, Brown P, Copley L, Browne J. Long-term outcomes from the English national comparative audit of surgery for nasal polyposis and chronic rhinosinusitis. *The Laryngoscope.* 2009 Dec;119(12):2459-65.
30. Chambers DW, Davis WE, Cook PR, Nishioka GJ, Rudman DT. Long-term outcome analysis of functional endoscopic sinus surgery: correlation of symptoms with endoscopic examination findings and potential prognostic variables. *Laryngoscope.* 1997 Apr;107(4):504-10.
31. Wigand ME. *Endoscopic Surgery of the Paranasal Sinuses and Anterior Skull Base.* 2 ed. Stuttgart New York: Georg Thieme Verlag; 2008.
32. Marks SC, Shamsa F. Evaluation of prognostic factors in endoscopic sinus surgery. *Am J Rhinol.* 1997 May-Jun;11(3):187-91.
33. Baumann I, Blumenstock G, Zalaman IM, Praetorius M, Klingmann C, Sittel C, Plinkert PK, Piccirillo JF. Impact of gender, age, and comorbidities on quality of life in patients with chronic rhinosinusitis. *Rhinology.* 2007 Dec;45(4):268-72.
34. Colclasure JC, Gross CW, Kountakis SE. Endoscopic sinus surgery in patients older than sixty. *Otolaryngol Head Neck Surg.* 2004 Dec;131(6):946-9.
35. Jiang RS, Hsu CY. Endoscopic sinus surgery for the treatment of chronic sinusitis in geriatric patients. *Ear Nose Throat J.* 2001 Apr;80(4):230-2.



36. Lee JY, Lee SW. Influence of age on the surgical outcome after endoscopic sinus surgery for chronic rhinosinusitis with nasal polyposis. *Laryngoscope*. 2007 Jun;117(6):1084-9.
37. Sil A, Mackay I, Rowe-Jones J. Assessment of predictive prognostic factors for functional endoscopic sinus surgery in a 5-year prospective outcome study. *Am J Rhinol*. 2007 May-Jun;21(3):289-96.
38. Ramadan HH, VanMetre R. Endoscopic sinus surgery in geriatric population. *Am J Rhinol*. 2004 Mar-Apr;18(2):125-7.
39. Bonfils P, Halimi P, Le Bihan C, Nores JM, Avan P, Landais P. Correlation between nasosinus symptoms and topographic diagnosis in chronic rhinosinusitis. *Ann Otol Rhinol Laryngol*. 2005 Jan;114(1 Pt 1):74-83.
40. Bhattacharyya N. Radiographic stage fails to predict symptom outcomes after endoscopic sinus surgery for chronic rhinosinusitis. *Laryngoscope*. 2006 Jan;116(1):18-22.
41. Bradley DT, Kountakis SE. Correlation between computed tomography scores and symptomatic improvement after endoscopic sinus surgery. *The Laryngoscope*. 2005 Mar;115(3):466-9.
42. Kennedy DW. Prognostic factors, outcomes and staging in ethmoid sinus surgery. *Laryngoscope*. 1992 Dec;102(12 Pt 2 Suppl 57):1-18.
43. Wang P-C, Chu C-C, Liang S-C, Tai C-J. Outcome predictors for endoscopic sinus surgery. *Otolaryngol Head Neck Surg*. 2002;126(2):154-9.
44. Sobol SE, Wright ED, Frenkiel S. One-year outcome analysis of functional endoscopic sinus surgery for chronic sinusitis. *J Otolaryngol*. 1998 Oct;27(5):252-7.
45. Deal RT, Kountakis SE. Significance of nasal polyps in chronic rhinosinusitis: symptoms and surgical outcomes. *Laryngoscope*. 2004 Nov;114(11):1932-5.
46. Poetker DM, Mendolia-Loffredo S, Smith TL. Outcomes of endoscopic sinus surgery for chronic rhinosinusitis associated with sinonasal polyposis. *Am J Rhinol*. 2007 Jan-Feb;21(1):84-8.
47. Dursun E, Korkmaz H, Eryilmaz A, Bayiz U, Sertkaya D, Samim E. Clinical predictors of long-term success after endoscopic sinus surgery. *Otolaryngol Head Neck Surg*. 2003 Nov;129(5):526-31.
48. Wynn R, Har-El G. Recurrence rates after endoscopic sinus surgery for massive sinus polyposis. *Laryngoscope*. 2004 May;114(5):811-3.
49. Bhattacharyya N. Clinical outcomes after revision endoscopic sinus surgery. *Arch Otolaryngol Head Neck Surg*. 2004 Aug;130(8):975-8.
50. Smith TL, Litvack JR, Hwang PH, Loehrl TA, Mace JC, Fong KJ, James KE. Determinants of outcomes of sinus surgery: a multi-institutional prospective cohort study. *Otolaryngol Head Neck Surg*. 2010 Jan;142(1):55-63.
51. Lee JY, Lee SW, Lee JD. Comparison of the surgical outcome between primary and revision endoscopic sinus surgery for chronic rhinosinusitis with nasal polyposis. *Am J Otolaryngol*. 2008 Nov-Dec;29(6):379-84.
52. Litvack JR, Griest S, James KE, Smith TL. Endoscopic and quality-of-life outcomes after revision endoscopic sinus surgery. *The Laryngoscope*. 2007 Dec;117(12):2233-8.
53. Young J, Frenkiel S, Tewfik MA, Mouadeb DA. Long-term outcome analysis of endoscopic sinus surgery for chronic sinusitis. *Am J Rhinol*. 2007 Nov-Dec;21(6):743-7.
54. Ramadan HH, Hinerman RA. Outcome of endoscopic sinus surgery in children with allergic rhinitis. *Am J Rhinol*. 2006 Jul-Aug;20(4):438-40.

55. Nishioka GJ, Cook PR, Davis WE, McKinsey JP. Immunotherapy in patients undergoing functional endoscopic sinus surgery. *Otolaryngol Head Neck Surg.* 1994 Apr;110(4):406-12.
56. Jarvis D, Newson R, Lotvall J, Hastan D, Tomassen P, Keil T, Gjomarkaj M, Forsberg B, Gunnbjornsdottir M, Minov J, Brozek G, Dahlen SE, Toskala E, Kowalski ML, Olze H, Howarth P, Kramer U, Baelum J, Loureiro C, Kasper L, Bousquet PJ, Bousquet J, Bachert C, Fokkens W, Burney P. Asthma in adults and its association with chronic rhinosinusitis: the GA2LEN survey in Europe. *Allergy.* 2012 Jan;67(1):91-8.
57. Gromek I. Zapalenie dróg oddechowych - współzależności. *Magazy Otorynolaryngologiczny.* 2005;IV(4):123-7.
58. Gromek I, Krzeski A. *Rynologiczne aspekty astmy aspirynowej.* Gdańsk: Via Medica; 2007.
59. Dinis PB, Gomes A. Sinusitis and asthma: how do they interrelate in sinus surgery? *Am J Rhinol.* 1997 Nov-Dec;11(6):421-8.
60. Proimos E, Papadakis CE, Chimona TS, Kiagiadaki D, Ferekidis E, Yiotakis J. The effect of functional endoscopic sinus surgery on patients with asthma and CRS with nasal polyps. *Rhinology.* 2010 Sep;48(3):331-8.
61. Dhong HJ, Jung YS, Chung SK, Choi DC. Effect of endoscopic sinus surgery on asthmatic patients with chronic rhinosinusitis. *Otolaryngol Head Neck Surg.* 2001 Jan;124(1):99-104.
62. Ragab S, Scadding GK, Lund VJ, Saleh H. Treatment of chronic rhinosinusitis and its effects on asthma. *Eur Respir J.* 2006 Jul;28(1):68-74.
63. Kountakis SE, Bradley DT. Effect of asthma on sinus computed tomography grade and symptom scores in patients undergoing revision functional endoscopic sinus surgery. *Am J Rhinol.* 2003 Jul-Aug;17(4):215-9.
64. Dejima K, Hama T, Miyazaki M, Yasuda S, Fukushima K, Oshima A, Yasuda M, Hisa Y. A clinical study of endoscopic sinus surgery for sinusitis in patients with bronchial asthma. *Int Arch Allergy Immunol.* 2005 Oct;138(2):97-104.
65. Mendelsohn D, Jeremic G, Wright ED, Rotenberg BW. Revision rates after endoscopic sinus surgery: a recurrence analysis. *Ann Otol Rhinol Laryngol.* 2011 Mar;120(3):162-6.
66. Szczeklik A, Stevenson DD. Aspirin-induced asthma: advances in pathogenesis, diagnosis, and management. *J Allergy Clin Immunol.* 2003 May;111(5):913-21; quiz 22.
67. Amar YG, Frenkiel S, Sobol SE. Outcome analysis of endoscopic sinus surgery for chronic sinusitis in patients having Samter's triad. *J Otolaryngol.* 2000 Feb;29(1):7-12.
68. Fokkens W, Lund V, Mullol J. European position paper on rhinosinusitis and nasal polyps 2007. *Rhinol Suppl.* 2007(20):1-136.
69. Kim JE, Kountakis SE. The prevalence of Samter's triad in patients undergoing functional endoscopic sinus surgery. *Ear Nose Throat J.* 2007 Jul;86(7):396-9.
70. Gosepath J, Pogodsky T, Mann WJ. Characteristics of recurrent chronic rhinosinusitis after previous surgical therapy. *Acta Otolaryngol.* 2008 Jul;128(7):778-84.
71. Arcimowicz M. Polipy nosa. W: Janczewski G, red. *Otolaryngologia praktyczna.* Gdańsk: Via Medica; 2007. s. 294-304.
72. Awad OG, Lee JH, Fasano MB, Graham SM. Sinonasal outcomes after endoscopic sinus surgery in asthmatic patients with nasal polyps: a difference between aspirin-tolerant and aspirin-induced asthma? *The Laryngoscope.* 2008 Jul;118(7):1282-6.

73. DelGaudio JM. Direct nasopharyngeal reflux of gastric acid is a contributing factor in refractory chronic rhinosinusitis. *Laryngoscope*. 2005 Jun;115(6):946-57.
74. Ferguson BJ, Otto BA, Pant H. When surgery, antibiotics, and steroids fail to resolve chronic rhinosinusitis. *Immunol Allergy Clin North Am*. 2009 Nov;29(4):719-32.
75. Briggs RD, Wright ST, Cordes S, Calhoun KH. Smoking in chronic rhinosinusitis: a predictor of poor long-term outcome after endoscopic sinus surgery. *Laryngoscope*. 2004 Jan;114(1):126-8.
76. Krzeski A, Galewicz A, Chmielewski R, Kisiel M. Influence of cigarette smoking on endoscopic sinus surgery long-term outcomes. *Rhinology*. 2011 Dec;49(5):577-82.
77. Rudmik L, Mace JC, Smith TL. Smoking and endoscopic sinus surgery: does smoking volume contribute to clinical outcome. *Int Forum Allergy Rhinol*. 2011 May-Jun;1(3):145-52.
78. Das S, Becker AM, Perakis H, Prosser JD, Kountakis SE. The effects of smoking on short-term quality of life outcomes in sinus surgery. *The Laryngoscope*. 2007 Dec;117(12):2229-32.
79. Das S, Khichi SS, Perakis H, Woodard T, Kountakis SE. Effects of smoking on quality of life following sinus surgery: 4-year follow-up. *The Laryngoscope*. 2009 Nov;119(11):2284-7.
80. Davis GE, Yueh B, Walker E, Katon W, Koepsell TD, Weymuller EA. Psychiatric distress amplifies symptoms after surgery for chronic rhinosinusitis. *Otolaryngol Head Neck Surg*. 2005 Feb;132(2):189-96.
81. Wasan A, Fernandez E, Jamison RN, Bhattacharyya N. Association of anxiety and depression with reported disease severity in patients undergoing evaluation for chronic rhinosinusitis. *Ann Otol Rhinol Laryngol*. 2007 Jul;116(7):491-7.
82. Bhattacharyya N, Wasan A. Do anxiety and depression confound symptom reporting and diagnostic accuracy in chronic rhinosinusitis? *Ann Otol Rhinol Laryngol*. 2008 Jan;117(1):18-23.
83. Albu S, Baciut M. Failures in endoscopic surgery of the maxillary sinus. *Otolaryngol Head Neck Surg*. 2010 Feb;142(2):196-201.
84. Matsuwaki Y, Ookushi T, Asaka D, Mori E, Nakajima T, Yoshida T, Kojima J, Chiba S, Ootori N, Moriyama H. Chronic rhinosinusitis: risk factors for the recurrence of chronic rhinosinusitis based on 5-year follow-up after endoscopic sinus surgery. *Int Arch Allergy Immunol*. 2008;146 Suppl 1:77-81.
85. Tosun F, Arslan HH, Karslioglu Y, Deveci MS, Durmaz A. Relationship between postoperative recurrence rate and eosinophil density of nasal polyps. *Ann Otol Rhinol Laryngol*. 2010 Jul;119(7):455-9.
86. Bhattacharyya N. Symptom outcomes after endoscopic sinus surgery for chronic rhinosinusitis. *Arch Otolaryngol Head Neck Surg*. 2004 Mar;130(3):329-33.
87. Soler ZM, Mace J, Smith TL. Symptom-based presentation of chronic rhinosinusitis and symptom-specific outcomes after endoscopic sinus surgery. *Am J Rhinol*. 2008 May-Jun;22(3):297-301.
88. Damm M, Quante G, Jungehueling M, Stennert E. Impact of functional endoscopic sinus surgery on symptoms and quality of life in chronic rhinosinusitis. *The Laryngoscope*. 2002 Feb;112(2):310-5.
89. Ling FT, Kountakis SE. Important clinical symptoms in patients undergoing functional endoscopic sinus surgery for chronic rhinosinusitis. *The Laryngoscope*. 2007 Jun;117(6):1090-3.
90. Jakobsen J, Svendstrup F. Functional endoscopic sinus surgery in chronic sinusitis--a series of 237 consecutively operated patients. *Acta Otolaryngol Suppl*. 2000;543:158-61.

91. Giger R, Dulguerov P, Quinodoz D, Leuba D, Landis BN, Lacroix JS, Friedrich JP. Chronic paranasal sinusitis without nasal polyps: long-term outcome after functional endoscopic sinus surgery. *Otolaryngol Head Neck Surg.* 2004 Oct;131(4):534-41.
92. Garrel R, Gardiner Q, Khudjadze M, Demoly P, Vergnes C, Makeieff M, Guerrier B, Crampette L. Endoscopic surgical treatment of sinonasal polyposis-medium term outcomes (mean follow-up of 5 years). *Rhinology.* 2003 Jun;41(2):91-6.
93. Senior BA, Kennedy DW, Tanabodee J, Kroger H, Hassab M, Lanza D. Long-term results of functional endoscopic sinus surgery. *The Laryngoscope.* 1998 Feb;108(2):151-7.
94. Agius AM. Long-term follow-up of patients with facial pain in chronic rhinosinusitis - correlation with nasal endoscopy and CT. *Rhinology.* 2010 Mar;48(1):65-70.
95. Bugten V, Nordgard S, Romundstad P, Steinsvag S. Chronic rhinosinusitis and nasal polyposis; indicia of heterogeneity. *Rhinology.* 2008 Mar;46(1):40-4.
96. Banerji A, Piccirillo JF, Thawley SE, Levitt RG, Schechtman KB, Kramper MA, Hamilos DL. Chronic rhinosinusitis patients with polyps or polypoid mucosa have a greater burden of illness. *Am J Rhinol.* 2007 Jan-Feb;21(1):19-26.
97. Bhattacharyya N. Influence of polyps on outcomes after endoscopic sinus surgery. *The Laryngoscope.* 2007 Oct;117(10):1834-8.
98. Agius AM. Chronic sinusitis in Malta - correlation between symptoms and CT scan. *Rhinology.* 2010 Mar;48(1):59-64.
99. Rosbe KW, Jones KR. Usefulness of patient symptoms and nasal endoscopy in the diagnosis of chronic sinusitis. *Am J Rhinol.* 1998 May-Jun;12(3):167-71.
100. Kenny TJ, Duncavage J, Bracikowski J, Yildirim A, Murray JJ, Tanner SB. Prospective analysis of sinus symptoms and correlation with paranasal computed tomography scan. *Otolaryngol Head Neck Surg.* 2001 Jul;125(1):40-3.
101. Seybt MW, McMains KC, Kountakis SE. The prevalence and effect of asthma on adults with chronic rhinosinusitis. *Ear Nose Throat J.* 2007 Jul;86(7):409-11.
102. Busaba NY, Sin HJ, Salman SD. Impact of gender on clinical presentation of chronic rhinosinusitis with and without polyposis. *J Laryngol Otol.* 2008 Nov;122(11):1180-4.
103. DiBaise JK, Olusola BF, Huerter JV, Quigley EM. Role of GERD in chronic resistant sinusitis: a prospective, open label, pilot trial. *Am J Gastroenterol.* 2002 Apr;97(4):843-50.
104. Wong IW, Rees G, Greiff L, Myers JC, Jamieson GG, Wormald PJ. Gastroesophageal reflux disease and chronic sinusitis: in search of an esophageal-nasal reflex. *Am J Rhinol Allergy.* 2010 Jul-Aug;24(4):255-9.
105. Hanna BC, Wormald PJ. Gastroesophageal reflux and chronic rhinosinusitis. *Curr Opin Otolaryngol Head Neck Surg.* 2012 Feb;20(1):15-8.
106. Wise SK, Wise JC, DelGaudio JM. Association of nasopharyngeal and laryngopharyngeal reflux with postnasal drip symptomatology in patients with and without rhinosinusitis. *Am J Rhinol.* 2006 May-Jun;20(3):283-9.
107. de Loos DAED, Hopkins C, Fokkens WJ. Symptoms in Chronic Rhinosinusitis With and Without Nasal Polyps. *Laryngoscope.* 2013 Jan;123(1):57-63.
108. Perry BF, Kountakis SE. Subjective improvement of olfactory function after endoscopic sinus surgery for chronic rhinosinusitis. *Am J Otolaryngol.* 2003 Nov-Dec;24(6):366-9.

109. Doty RL, Mishra A. Olfaction and its alteration by nasal obstruction, rhinitis, and rhinosinusitis. *The Laryngoscope*. 2001 Mar;111(3):409-23.
110. Bramerson A, Nordin S, Bende M. Clinical experience with patients with olfactory complaints, and their quality of life. *Acta Otolaryngol*. 2007 Feb;127(2):167-74.
111. Litvack JR, Fong K, Mace J, James KE, Smith TL. Predictors of olfactory dysfunction in patients with chronic rhinosinusitis. *The Laryngoscope*. 2008 Dec;118(12):2225-30.
112. Kern RC. Chronic sinusitis and anosmia: pathologic changes in the olfactory mucosa. *The Laryngoscope*. 2000 Jul;110(7):1071-7.
113. Robinson AM, Conley DB, Shinnery MJ, Kern RC. Apoptosis in the aging olfactory epithelium. *The Laryngoscope*. 2002 Aug;112(8 Pt 1):1431-5.
114. Raviv JR, Kern RC. Przewlekłe zapalenia zatok przynosowych a zaburzenia węchu. W: Krzeski A, Gromek I, red. *Zapalenia zatok przynosowych*. Gdańsk: Via Medica; 2008. s. 311-22.
115. Lund VJ, Scadding GK. Objective assessment of endoscopic sinus surgery in the management of chronic rhinosinusitis: an update. *J Laryngol Otol*. 1994 Sep;108(9):749-53.
116. Downey LL, Jacobs JB, Lebowitz RA. Anosmia and chronic sinus disease. *Otolaryngol Head Neck Surg*. 1996 Jul;115(1):24-8.
117. Minovi A, Hummel T, Ural A, Draf W, Bockmuhl U. Predictors of the outcome of nasal surgery in terms of olfactory function. *Eur Arch Oto-Rhino-L*. 2008 Jan;265(1):57-61.
118. Min YG, Yun YS, Song BH, Cho YS, Lee KS. Recovery of nasal physiology after functional endoscopic sinus surgery: olfaction and mucociliary transport. *ORL J Otorhinolaryngol Relat Spec*. 1995 Sep-Oct;57(5):264-8.
119. Delank KW, Stoll W. Olfactory function after functional endoscopic sinus surgery for chronic sinusitis. *Rhinology*. 1998 Mar;36(1):15-9.
120. Delank KW, Stoll W. [Sense of smell before and after endonasal surgery in chronic sinusitis with polyps]. *Hno*. 1994 Oct;42(10):619-23.
121. Pade J, Hummel T. Olfactory Function Following Nasal Surgery. *Laryngoscope*. 2008 Jul;118(7):1260-4.
122. Kimmelman CP. The risk to olfaction from nasal surgery. *The Laryngoscope*. 1994 Aug;104(8 Pt 1):981-8.
123. Jiang RS, Lu FJ, Liang KL, Shiao JY, Su MC, Hsin CH, Chen WK. Olfactory function in patients with chronic rhinosinusitis before and after functional endoscopic sinus surgery. *Am J Rhinol*. 2008 Jul-Aug;22(4):445-8.
124. Klimek L, Moll B, Amedee RG, Mann WJ. Olfactory function after microscopic endonasal surgery in patients with nasal polyps. *Am J Rhinol*. 1997 Jul-Aug;11(4):251-5.
125. Litvack JR, Mace J, Smith TL. Does olfactory function improve after endoscopic sinus surgery? *Otolaryng Head Neck*. 2009 Mar;140(3):312-9.
126. Oka H, Tsuzuki K, Takebayashi H, Kojima Y, Daimon T, Sakagami M. Olfactory changes after endoscopic sinus surgery in patients with chronic rhinosinusitis. *Auris Nasus Larynx*. 2013 Jan 11.
127. Hamilos DL. Chronic rhinosinusitis patterns of illness. *Clin Allergy Immunol*. 2007;20:1-13.
128. Soler ZM, Sauer DA, Mace JC, Smith TL. Ethmoid histopathology does not predict olfactory outcomes after endoscopic sinus surgery. *Am J Rhinol Allergy*. 2010 Jul-Aug;24(4):281-5.

129. Szczeklik A, Nizankowska E. Clinical features and diagnosis of aspirin induced asthma. *Thorax*. 2000 Oct;55 Suppl 2:S42-4.
130. Samoliński B. Polipy nosa. W: Krzeski A, Janczewski G, red. *Choroby nosa i zatok przynosowych*. Wrocław: Urban & Partner; 2003. s. 187-94.
131. Katotomichelakis M, Riga M, Davris S, Tripsianis G, Simopoulou M, Nikolettos N, Simopoulos K, Danielides V. Allergic rhinitis and aspirin-exacerbated respiratory disease as predictors of the olfactory outcome after endoscopic sinus surgery. *Am J Rhinol Allergy*. 2009 May-Jun;23(3):348-53.
132. Schubert CR, Cruickshanks KJ, Klein BE, Klein R, Nondahl DM. Olfactory impairment in older adults: five-year incidence and risk factors. *The Laryngoscope*. 2011 Apr;121(4):873-8.
133. Janczewski G. Symptomatologia chorób nosa i zatok przynosowych. W: Krzeski A, Janczewski G, red. *Choroby nosa i zatok przynosowych*. Wrocław: Urban&Partner; 2003. s. 118-36.
134. Simola M, Malmberg H. Sense of smell in allergic and nonallergic rhinitis. *Allergy*. 1998 Feb;53(2):190-4.
135. Alobid I, Cardelus S, Benitez P, Guilemany JM, Roca-Ferrer J, Picado C, Bernal-Sprekelsen M, Mullol J. Persistent asthma has an accumulative impact on the loss of smell in patients with nasal polyposis. *Rhinology*. 2011 Dec;49(5):519-24.
136. Apter AJ, Gent JF, Frank ME. Fluctuating olfactory sensitivity and distorted odor perception in allergic rhinitis. *Arch Otolaryngol Head Neck Surg*. 1999 Sep;125(9):1005-10.
137. Cowart BJ, Flynn-Rodden K, McGeady SJ, Lowry LD. Hyposmia in allergic rhinitis. *J Allergy Clin Immunol*. 1993 Mar;91(3):747-51.
138. Seiden AM, Duncan HJ. The diagnosis of a conductive olfactory loss. *The Laryngoscope*. 2001 Jan;111(1):9-14.
139. Olsson P, Berglind N, Bellander T, Stjarne P. Prevalence of self-reported allergic and non-allergic rhinitis symptoms in Stockholm: Relation to age, gender, olfactory sense and smoking. *Acta Otolaryngol*. 2003;123(1):75-80.
140. Frye RE, Schwartz BS, Doty RL. Dose-related effects of cigarette smoking on olfactory function. *Jama*. 1990 Mar 2;263(9):1233-6.
141. Damm M, Eckel HE, Jungehulsing M, Hummel T. Olfactory changes at threshold and suprathreshold levels following septoplasty with partial inferior turbinectomy. *Ann Otol Rhinol Laryngol*. 2003 Jan;112(1):91-7.
142. Jones NS. Zapalenia zatok przynosowych a ból głowy. W: Krzeski A, Gromek I, red. *Zapalenia zatok przynosowych*. Gdańsk: Via Medica; 2008. s. 301-10.
143. Tarabichi M. Characteristics of sinus-related pain. *Otolaryngol Head Neck Surg*. 2000 Jun;122(6):842-7.
144. Daudia AT, Jones NS. Facial migraine in a rhinological setting. *Clin Otolaryngol Allied Sci*. 2002 Dec;27(6):521-5.
145. Moretz WH, Kountakis SE. Subjective headache before and after endoscopic sinus surgery. *Am J Rhinol*. 2006 May-Jun;20(3):305-7.
146. Fahy C, Jones NS. Nasal polyposis and facial pain. *Clin Otolaryngol Allied Sci*. 2001 Dec;26(6):510-3.
147. Jones NS, Cooney TR. Facial pain and sinonasal surgery. *Rhinology*. 2003 Dec;41(4):193-200.
148. Bhattacharyya N. Assessing the additional disease burden of polyps in chronic rhinosinusitis. *Ann Otol Rhinol Laryngol*. 2009 Mar;118(3):185-9.

149. Sahlstrand-Johnson P, Ohlsson B, Von Buchwald C, Jannert M, Ahlner-Elmqvist M. A multi-centre study on quality of life and absenteeism in patients with CRS referred for endoscopic surgery. *Rhinology*. 2011 Oct;49(4):420-8.
150. DelGaudio JM, Wise SK, Wise JC. Association of radiological evidence of frontal sinus disease with the presence of frontal pain. *Am J Rhinol*. 2005 Mar-Apr;19(2):167-73.
151. Mudgil SP, Wise SW, Hopper KD, Kasales CJ, Mauger D, Fornadley JA. Correlation between presumed sinusitis-induced pain and paranasal sinus computed tomographic findings. *Ann Allerg Asthma Im*. 2002 Feb;88(2):223-6.
152. West B, Jones NS. Endoscopy-negative, computed tomography-negative facial pain in a nasal clinic. *The Laryngoscope*. 2001 Apr;111(4 Pt 1):581-6.
153. Tatar M, Plevkova J, Brozmanova M, Pecova R, Kollarik M. Mechanisms of the cough associated with rhinosinusitis. *Pulm Pharmacol Ther*. 2009 Apr;22(2):121-6.
154. Irwin RS, Pratter MR, Holland PS, Corwin RW, Hughes JP. Postnasal drip causes cough and is associated with reversible upper airway obstruction. *Chest*. 1984 Mar;85(3):346-52.
155. Weldon DR. Gastroesophageal reflux disease and sinusitis: their role in patients with chronic cough. *Allergy Asthma Proc*. 2006 Jan-Feb;27(1):36-44.
156. O'Hara J, Jones NS. "Post-nasal drip syndrome": most patients with purulent nasal secretions do not complain of chronic cough. *Rhinology*. 2006 Dec;44(4):270-3.
157. Irwin RS, French CL, Curley FJ, Zawacki JK, Bennett FM. Chronic cough due to gastroesophageal reflux. Clinical, diagnostic, and pathogenetic aspects. *Chest*. 1993 Nov;104(5):1511-7.
158. Stoikes NF, Dutton JM. The effect of endoscopic sinus surgery on symptoms of eustachian tube dysfunction. *Am J Rhinol*. 2005 Mar-Apr;19(2):199-202.
159. Hong CK, Park DC, Kim SW, Cha CI, Cha SH, Yeo SG. Effect of paranasal sinusitis on the development of otitis media with effusion: influence of eustachian tube function and adenoid immunity. *Int J Pediatr Otorhinolaryngol*. 2008 Nov;72(11):1609-18.
160. Leo G, Piacentini E, Incorvaia C, Consonni D. Sinusitis and Eustachian tube dysfunction in children. *Pediatr Allergy Immunol*. 2007 Nov;18 Suppl 18:35-9.
161. Leo G, Piacentini E, Incorvaia C, Consonni D, Cazzavillan A. Prevalence of tympanometric alterations in children with chronic sinusitis. *Int J Pediatr Otorhinolaryngol*. 2008 Mar;72(3):315-9.
162. Pagella F, Colombo A, Gatti O, Giourgos G, Matti E. Rhinosinusitis and otitis media: the link with adenoids. *Int J Immunopathol Pharmacol*. 2010 Jan-Mar;23(1 Suppl):38-40.
163. Finkelstein Y, Ophir D, Talmi YP, Shabtai A, Strauss M, Zohar Y. Adult-onset otitis media with effusion. *Arch Otolaryngol Head Neck Surg*. 1994 May;120(5):517-27.
164. Parietti-Winkler C, Jankowski R. Is there an association between otitis media and nasal polyposis? *Curr Allergy Asthma Rep*. 2011 Dec;11(6):521-5.
165. Iino Y, Tomioka-Matsutani S, Matsubara A, Nakagawa T, Nonaka M. Diagnostic criteria of eosinophilic otitis media, a newly recognized middle ear disease. *Auris Nasus Larynx*. 2011 Aug;38(4):456-61.
166. Parietti-Winkler C, Baumann C, Gallet P, Gauchard G, Jankowski R. Otitis media with effusion as a marker of the inflammatory process associated to nasal polyposis. *Rhinology*. 2009 Dec;47(4):396-9.

167. Gibson WS, Jr., Cochran W. Otagia in infants and children--a manifestation of gastroesophageal reflux. *Int J Pediatr Otorhinolaryngol*. 1994 Jan;28(2-3):213-8.
168. Al-Saab F, Manoukian JJ, Al-Sabah B, Almot S, Nguyen LH, Tewfik TL, Daniel SJ, Schloss MD, Hamid QA. Linking laryngopharyngeal reflux to otitis media with effusion: pepsinogen study of adenoid tissue and middle ear fluid. *J Otolaryngol Head Neck Surg*. 2008 Aug;37(4):565-71.
169. Miura MS, Mascaro M, Rosenfeld RM. Association between otitis media and gastroesophageal reflux: a systematic review. *Otolaryngol Head Neck Surg*. 2012 Mar;146(3):345-52.
170. Kong SK, Chon KM, Goh EK, Lee IW, Lee JW, Wang SG. Histologic changes in the auditory tube mucosa of rats after long-term exposure to cigarette smoke. *Am J Otolaryngol*. 2009 Nov-Dec;30(6):376-82.
171. Hoffman HJ, Daly KA, Bainbridge KE, Casselbrant ML, Homoe P, Kvestad E, Kvaerner KJ, Vernacchio L. Panel 1: epidemiology, natural history, and risk factors. *Otolaryngol Head Neck Surg*. 2013 Apr;148(4 Suppl):E1-E25.
172. Deron P, Clement PA, Derde MP. Septal surgery and tubal function: early and late results. *Rhinology*. 1995 Mar;33(1):7-9.
173. Pelikan Z. Role of nasal allergy in chronic secretory otitis media. *Curr Allergy Asthma Rep*. 2009 Mar;9(2):107-13.
174. Skoner AR, Skoner KR, Skoner DP. Allergic rhinitis, histamine, and otitis media. *Allergy Asthma Proc*. 2009 Sep-Oct;30(5):470-81.
175. Doyle WJ. The link between allergic rhinitis and otitis media. *Curr Opin Allergy Clin Immunol*. 2002 Feb;2(1):21-5.
176. Pratibha PK, Bhat KM, Bhat GS. Oral malodor: a review of the literature. *J Dent Hyg*. 2006 Summer;80(3):8.
177. Marsicano JA, de Moura-Grec PG, Bonato RC, Sales-Peres Mde C, Sales-Peres A, Sales-Peres SH. Gastroesophageal reflux, dental erosion, and halitosis in epidemiological surveys: a systematic review. *Eur J Gastroenterol Hepatol*. 2013 Feb;25(2):135-41.
178. Chester AC, Sindwani R, Smith TL, Bhattacharyya N. Fatigue improvement following endoscopic sinus surgery: a systematic review and meta-analysis. *The Laryngoscope*. 2008 Apr;118(4):730-9.
179. Sautter NB, Mace J, Chester AC, Smith TL. The effects of endoscopic sinus surgery on level of fatigue in patients with chronic rhinosinusitis. *Am J Rhinol*. 2008 Jul-Aug;22(4):420-6.
180. Sedaghat AR, Bhattacharyya N. Chronic rhinosinusitis symptoms and computed tomography staging: improved correlation by incorporating radiographic density. *Int Forum Allergy Rhinol*. 2012 Sep-Oct;2(5):386-91.
181. Albu S, Tomescu E, Mexca Z, Nistor S, Necula S, Cozlean A. Recurrence rates in endonasal surgery for polyposis. *Acta Otorhinolaryngol Belg*. 2004;58(1):79-86.
182. Katotomichelakis M, Gouveris H, Tripsianis G, Simopoulou M, Papathanassiou J, Danielides V. Biometric predictive models for the evaluation of olfactory recovery after endoscopic sinus surgery in patients with nasal polyposis. *Am J Rhinol Allergy*. 2010 Jul-Aug;24(4):276-80.
183. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform*. 2002 Oct-Dec;35(5-6):352-9.
184. Mohamed EI, Bruno E, Linder R, Alessandrini M, Di Girolamo A, Poppl SJ, Puija A, De Lorenzo A. A novel method for diagnosing chronic rhinosinusitis based on an electronic nose. *An Otorrinolaringol Ibero Am*. 2003;30(5):447-57.



185. Zurada JM. Introduction to Artificial Neural Systems. St. Paul, New York, Los Angeles, San Francisco: West Publishing Company; 1992.
186. Duch W, Korbicz, J., Rutkowski, L., Tadeusiewicz, R., red. Sieci neuronowe. 1 ed. Warszawa: Akademicka Oficyna Wydawnicza EXIT; 2000.
187. Tadeusiewicz R. Sieci Neuronowe. . Warszawa: Akademicka Oficyna Wydawnicza; 1993.
188. Tadeusiewicz R. Możliwości i problemy wykorzystania sieci neuronowych w prognozowaniu procesów gospodarczych. Zeszyty Naukowe Akademii Ekonomicznej w Krakowie. 1997;493:5-18.
189. Tadeusiewicz R. Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami. Bolc L, red. Warszawa: Akademia Oficyna Wydawnicza PLJ; 1998.
190. Kappen HJ, Neijt JP. Neural network analysis to predict treatment outcome. Ann Oncol. 1993 January 1, 1993;4(suppl 4):S31-S4.
191. Theeuwens M, Kappen B, Neijt J, editors. Neural network analysis to predict treatment outcome in patients with ovarian cancer. . Proceedings Session 5, Medicine, International Conference on Artificial Neural Networks; 1995; Paris.
192. Ripley RM, Harris AL, Tarassenko L. Neural network models for breast cancer prognosis. Neural Comput Appl. 1998;7(4):367-75.
193. Ravdin P, Clark G. A practical application of neural network analysis for predicting outcome of individual breast cancer patients. Breast Cancer Research and Treatment. 1992;22(3):285-93.
194. Lisboa PJG, Wong H, Harris P, Swindell R. A Bayesian neural network approach for modelling censored data with an application to prognosis after surgery for breast cancer. Artificial intelligence in medicine. 2003;28(1):1-25.
195. Bottaci L, Drew PJ, Hartley JE, Hadfield MB, Farouk R, Lee PWR, Macintyre IMC, Duthie GS, Monson JRT. Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. Lancet. 1997 Aug 16;350(9076):469-72.
196. Ohno-Machado L. Sequential use of neural networks for survival prediction in AIDS. Proc AMIA Annu Fall Symp. 1996:170-4.
197. Ohno-Machado L, Musen MA. Sequential versus standard neural networks for pattern recognition: an example using the domain of coronary heart disease. Comput Biol Med. 1997 Jul;27(4):267-81.
198. Baxt WG. Application of artificial neural networks to clinical medicine. Lancet. 1995;346(8983):1135.
199. Szaleniec J, Skladzien J, Tadeusiewicz R, Oles K, Konior M, Przeklasa R. How can an otolaryngologist benefit from artificial neural networks? Otolaryngol Pol. 2012 Jul;66(4):241-8.
200. Jones AS, Taktak AGF, Helliwell TR, Fenton JE, Birchall MA, Husband DJ, Fisher AC. An artificial neural network improves prediction of observed survival in patients with laryngeal squamous carcinoma. Eur Arch Oto-Rhino-L. 2006 Jun;263(6):541-7.
201. Bryce TJ, Dewhurst MW, Floyd CE, Hars V, Brizel DM. Artificial neural network model of survival in patients treated with irradiation with and without concurrent chemotherapy for advanced carcinoma of the head and neck. Int J Radiat Oncol. 1998 May 1;41(2):339-45.
202. Szaleniec J, Wiatr M, Szaleniec M, Skladzien J, Tomik J, Oles K, Tadeusiewicz R. Artificial neural network modelling of the results of tympanoplasty in chronic suppurative otitis media patients. Comput Biol Med. 2013 Jan 1;43(1):16-22.

203. Szaleniec J, Wiatr M, Szaleniec M, Skladzien J, Tomik J, Strek P, Tadeusiewicz R, Przeklasa R. Przydatność sztucznych sieci neuronowych w prognozowaniu pooperacyjnej poprawy słuchu u chorych z przewlekłym zapaleniem ucha środkowego. *Przegl Lek.* 2009;66(11):924-9.
204. Williams PE, inventor Method and Apparatus for Diagnosing an Allergy of the Upper Respiratory Tract Using a Neural Network.2009.
205. Juhola M, Laurikkala J, Viikki K, Kentala E, Pyykko I. Classification of patients on the basis of otoneurological data by using Kohonen networks. *Acta Otolaryngol Suppl.* 2001;545:50-2.
206. el-Solh AA, Mador MJ, Ten-Brock E, Shucard DW, Abul-Khoudoud M, Grant BJ. Validity of neural network in sleep apnea. *Sleep.* 1999 Feb 1;22(1):105-11.
207. Wochlik I, Bulka J, Izworski A, Paslowski A. Preliminary processing and detection of characteristic features of Brainstem Auditory Evoked Potentials (BAEP). *Proceedings of the 23rd Annual International Conference of the Ieee Engineering in Medicine and Biology Society, Vols 1-4.* 2001;23:1765-8.
208. Buller G, Lutman ME. Automatic classification of transiently evoked otoacoustic emissions using an artificial neural network. *Brit J Audiol.* 1998 Aug;32(4):235-47.
209. Ziavra N, Kastanioudakis I, Trikalinos TA, Skevas A, Ioannidis JPA. Diagnosis of sensorineural hearing loss with neural networks versus logistic regression modeling of distortion product otoacoustic emissions. *Audiol Neuro-Otol.* 2004;9(2):81-7.
210. Ritchings RT, McGillion M, Moore CJ. Pathological voice quality assessment using artificial neural networks. *Med Eng Phys.* 2002 Sep-Oct;24(7-8):561-4.
211. Schonweiler R, Hess M, Wubbelt P, Ptok M. Novel approach to acoustical voice analysis using artificial neural networks. *J Assoc Res Otolaryngol.* 2000 Dec;1(4):270-82.
212. Linder R, Albers AE, Hess M, Poppl SJ, Schonweiler R. Artificial neural network-based classification to screen for dysphonia using psychoacoustic scaling of acoustic voice features. *J Voice.* 2008 Mar;22(2):155-63.
213. Szaleniec J. Zastosowanie sieci neuronowych w różnicowaniu mowy patologicznej od fizjologicznej. *Przegl Lek.* 2005;2005/62/Supplement.
214. Tadeusiewicz R, Jaworek J, Kańtoch E, Miller J, Pięciak T, Przybyło J. Wprowadzenie do modelowania systemów biologicznych oraz ich symulacji w środowisku MATLAB. Lublin: UMCS; 2012.
215. Lula P, Tadeusiewicz R, red. Wprowadzenie do sieci neuronowych. Kraków: Statsoft; 2001.
216. Tadeusiewicz R. Sieci neuronowe - przewodnik problemowy. *Elektrotechnika.* 1991;10(2):125 – 67.
217. Livingstone DJ, Manallack DT. Statistics using neural networks: chance effects. *Journal of medicinal chemistry.* 1993 Apr 30;36(9):1295-7.
218. Andrea TA, Kalayeh H. Applications of neural networks in quantitative structure-activity relationships of dihydrofolate reductase inhibitors. *J Med Chem.* 1991;34(9):2824-36.
219. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics.* 2002;35(5-6):352-9.
220. Tetko IV, Livingstone DJ, Luik AI. Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of chemical information and computer sciences.* 1995;35(5):826-33.
221. Schwarzer G, Vach W, Schumacher M. On the misuses of artificial neural networks for prognostic and diagnostic classification in oncology. *Statistics in Medicine.* 2000;19(4):541-61.

222. Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? *Bmj*. 1995 1995-12-09 08:00:00;311(7019):1539-41.
223. Shannon CE. A Mathematical Theory of Communication. *The Bell System Technical Journal*. 1948;27(July, October):379–423, 623–56.
224. Stanisław A. Przystępny kurs statystyki w oparciu o STATISTICA PL na przykładach z medycyny. Kraków: StatSoft; 2001.
225. Tadeusiewicz R. Sieci neuronowe w badaniach mózgu. W: M. D-W, red. Zaawansowane metody badania mózgu. Kraków: Instytut Farmakologii PAN; 2011. s. 7 – 25.
226. Szaleniec J, Szaleniec M, Stręk P, Boroń A, Jabłońska K, Gawlik J, Składzień J. Outcome prediction in endoscopic surgery for chronic rhinosinusitis – a multidimensional model. *Advances in Medical Sciences*. 2014;59/1. (praca przyjęta do druku)
227. Ohno-Machado L. Modeling Medical Prognosis: Survival Analysis Techniques. *Journal of Biomedical Informatics*. 2001;34(6):428-39.