



**Politechnika  
Śląska**

**Wydział Automatyki, Elektroniki i Informatyki  
Politechniki Śląskiej**

**ROZPRAWA DOKTORSKA**

**Modele predykcyjne wspomagające diagnostykę  
choroby Alzheimera oraz łagodnych zaburzeń  
funkcji poznawczych z wykorzystaniem  
obrazowania metodą rezonansu magnetycznego**

**Anna Marcisz**

**PROMOTOR  
Prof. dr hab. Joanna Polańska**

**Gliwice 2022**

*Pracę tę dedykuję śp. dr Franciszkowi Binczykowi,  
który zainspirował mnie do zajęcia się tematem  
obrazowania mózgu w chorobie Alzheimera  
oraz był moim Promotorem pomocniczym.*

*Najserdeczniejsze podziękowania składam  
Pani prof. dr hab. Joannie Polańskiej za mentorską opiekę,  
cenne, merytoryczne rady, poświęcony czas  
oraz wsparcie w wielu trudnych momentach.*

*Dziękuję Koleżankom i Kolegom z Katedry Inżynierii  
i Analizy Eksploracyjnej Danych za pomocne wskazówki  
oraz za dzielenie się wiedzą i doświadczeniem.*

*Z serca dziękuję mojemu Mężowi i Dzieciom  
- Stanisławowi, Róży, Helenie, Samuelowi, Glorii i Miriam,  
za cierpliwość i wyrozumiałość oraz wspólnie niesiony trud  
wielu godzin poświęconych na przygotowanie  
i napisanie tej pracy.*

## Spis treści

1. Wstęp.....	7
1.1. Cel i teza rozprawy.....	7
1.2. Zawartość rozprawy.....	9
2. Wprowadzenie .....	12
2.1. Choroba Alzheimera .....	12
2.1. Biomarkery choroby Alzheimera .....	16
2.2. Metody obrazowania.....	18
2.2.1. Rentgenografia .....	18
2.2.2. Ultrasonografia .....	19
2.2.3. Tomografia komputerowa .....	19
2.2.4. Rezonans magnetyczny.....	20
2.2.5. Pozytonowa tomografia emisyjna.....	23
2.3. Metody klasyfikacyjne .....	24
3. Wstępne przetwarzanie danych .....	26
3.1. Bazy danych .....	27
3.2. Struktura analizowanych danych (MPRAGE) .....	28
3.3. Charakterystyka kliniczna pacjentów z baz ADNI oraz EDS.....	29
3.4. Korekta niejednorodności natężenia pola .....	34
3.5. Ekstrakcja mózgu .....	35
3.6. Segmentacja .....	38
4. Analiza danych.....	49
4.1. Zmienne.....	49
4.1.1. Względna objętość mózgu .....	49
4.1.1.1. Względna objętość mózgu - metoda wyznaczania .....	50
4.1.1.2. Względna objętość mózgu - wyniki .....	51
4.1.2. Współczynnik pofałdowania .....	53
4.1.2.1. Współczynnik pofałdowania - metoda wyznaczania .....	53
4.1.2.2. Współczynnik pofałdowania - wyniki .....	59
4.1.3. Względna objętość komór bocznych mózgu .....	61
4.1.3.1. Względna objętość komór bocznych mózgu - metoda wyznaczania.....	61
4.1.3.2. Względna objętość komór bocznych mózgu - wyniki.....	61
4.2. Model matematyczny .....	63
5. Wyniki .....	69
5.1. Model wielomianowej regresji logistycznej z walidacją krzyżową na zbiorze ADNI.....	69

5.2.	Wskaźniki jakości modelu predykcyjnego – walidacja krzyżowa na zbiorze ADNI .....	71
5.3.	Model wielomianowej regresji logistycznej testowany na niezależnym zbiorze EDSD.	78
5.4.	Wskaźniki jakości modelu predykcyjnego testowanego na niezależnym zbiorze EDSD	79
5.5.	Efektywność modelu predykcyjnego – krzywe ROC.....	82
5.6.	Zgodność predykcji ze zmianą diagnozy podczas okresu obserwacji .....	85
6.	Dyskusja .....	92
7.	Podsumowanie .....	98
8.	Wnioski .....	102
9.	Podziękowania .....	104
10.	Spis skrótów .....	105
11.	Spis tabel .....	107
12.	Spis rysunków.....	109
13.	Spis literatury .....	113

## Streszczenie

Choroba Alzheimera (ang. Alzheimer's disease, AD) to postępująca, neurodegeneracyjna choroba mózgu, która powoduje utratę pamięci, zmiany w zachowaniu i problemy z codziennymi czynnościami. Wczesne wykrycie choroby Alzheimera oraz łagodnych zaburzeń funkcji poznawczych jest kluczowe z punktu widzenia podjęcia terapii i wydłużenia życia człowieka oraz złagodzenia objawów. Głównym celem pracy było znalezienie łatwo dostępnych biomarkerów choroby Alzheimera (AD) i łagodnego upośledzenia funkcji poznawczych (MCI) oraz zbudowanie modelu predykcyjnego, który będzie efektywnie klasyfikował pacjentów z chorobą Alzheimera oraz z łagodnym upośledzeniem funkcji poznawczych opierając się na danych z obrazowania metodą rezonansu magnetycznego a także rozpozna z czasowym wyprzedzeniem tych pacjentów, którzy z łagodnych zaburzeń funkcji poznawczych rozwiną chorobę Alzheimera.

Wielomianowa regresja logistyczna została zastosowana do predykcji statusu diagnozy: choroba Alzheimera (AD), łagodne upośledzenie funkcji poznawczych (MCI) oraz status osoby zdrowej (NC, grupa kontrolna). W celu wyboru najlepszego modelu zastosowano Bayesowskie kryterium informacyjne. Przeprowadzono 5-krotną walidację krzyżową na zbiorze ADNI, ponadto model testowano na niezależnym zbiorze EDSD. Spośród analizowanych biomarkerów najlepszym predyktorem pochodzącym z obrazowania MRI okazała się względna objętość mózgu. W pracy udowodnione zostało, że włączenie powyższego biomarkera opartego na MRI do standardowych klinicznych predyktorów AD prowadzi do powstania efektywnego i użytecznego w codziennej praktyce klinicznej modelu usprawniającego proces diagnozy poprzez wstępną detekcję pacjentów cierpiących na chorobę Alzheimera lub łagodne upośledzenie funkcji poznawczych. Ponadto, wykazane zostało, że model rozpoznaje pacjentów zmieniających status choroby z MCI na AD podczas okresu obserwacji jako pacjentów z chorobą Alzheimera już na kilka lat wcześniej niż standardowy, medyczny proces diagnozy, co pozwala na wcześniejsze podjęcie leczenia i opóźnienie postępu choroby. Model bazuje na łatwo dostępnych danych, które można znaleźć w historii choroby pacjentów z podejrzeniem choroby Alzheimera, a klasyczne metody statystyczne nie wymagają dużej mocy obliczeniowej, co sprawia, że model jest gotowy do zastosowania w codziennej praktyce lekarskiej jako test przesiewowy choroby Alzheimera i łagodnych zaburzeń funkcji poznawczych.

## Abstract

Alzheimer's disease (AD) is a progressive, neurodegenerative brain disease that causes memory loss, changes in behaviour, and problems with everyday tasks. Detecting early-stage Alzheimer's disease is still problematic in clinical practice. This work aimed to find T1-weighted MRI-based markers for Alzheimer's disease (AD) and mild cognitive impairment (MCI) and build the mathematical model to improve the screening process. The assumption was to build a screening model that would be accessible and easy to use for physicians in their daily clinical routine.

The multinomial logistic regression was used to detect status: AD, MCI, and normal control (NC) combined with the Bayesian information criterion for model selection. Several T1-weighted MRI-based features were considered as explanatory variables in the prediction model. A 5-fold cross-validation was executed for ADNI dataset. The multinomial logistic regression model was also trained on the whole ADNI dataset and tested on the independent EDSB dataset.

The best T1-weighted MRI-based predictor was the relative brain volume. This work shows that the proposed T1-weighted MRI-based biomarker, combined with standard clinical predictors, gives excellent early-stage AD status predictions. Moreover, this method, as based on MRI, doesn't require invasive and expensive laboratory tests and, as being a classical statistical learning model, doesn't require large calculation power. Additionally, this model detects some patients transitioning from MCI to AD as AD patients a few years earlier before regular medical diagnosis, allowing for earlier treatment and delaying disease progression.

The created method is non-invasive, inexpensive, clinically accessible, and efficiently supports the AD/MCI screening. The model is based on easily available parameters (T1-weighted MRI is standard) and can be calculated in a simple way, so this method is ready to use in medical practice.

# 1. Wstęp

## 1.1. Cel i teza rozprawy

Głównym celem pracy było zbudowanie modeli predykcyjnych, które będą wspomagały diagnostykę choroby Alzheimera oraz łagodnych zaburzeń funkcji poznawczych wykorzystując obrazowanie metodą rezonansu magnetycznego.

Cel główny zawierał w sobie dwa cele pośrednie: było to znalezienie łatwo dostępnych biomarkerów choroby Alzheimera (ang. Alzheimer's disease, AD) i łagodnego upośledzenia funkcji poznawczych (ang. mild cognitive impairment, MCI) oraz zbudowanie modelu klasyfikacyjnego, który będzie efektywnie klasyfikował pacjentów z chorobą Alzheimera oraz z łagodnym upośledzeniem funkcji poznawczych. Dodatkowym wyzwaniem było zbudowanie modelu predykcyjnego, który będzie rozpoznawał chorych z chorobą Alzheimera zanim przejdą oni ze statusu diagnozy „łagodne upośledzenie funkcji poznawczych” do statusu „choroba Alzheimera” w klasycznym procesie diagnozy medycznej.

Postawiona została następująca teza rozprawy:

Model predykcyjny wykorzystujący informację o wynikach testu „Krótka Skala Zaburzeń Poznawczych” (ang. Mini Mental State Examination, MMSE), wieku pacjenta oraz względnej objętości mózgu jest skutecznym testem przesiewowym choroby Alzheimera, który efektywnie i z większym wyprzedzeniem czasowym, w porównaniu do standardowo stosowanego testu przesiewowego opartego wyłącznie na wynikach testu MMSE, rozpoznaje chorych z łagodnymi zaburzeniami funkcji poznawczych, którzy później rozwiną pełnoobjawową chorobę Alzheimera.

Istnieje wiele zaawansowanych i efektywnych metod diagnostycznych (zostały one opisane w dalszej części pracy), jednak często są one kosztowne lub inwazyjne. Brakuje jednak skutecznych metod przesiewowych, które są pierwszym etapem w procesie stawiania właściwej diagnozy. Przedmiotem pracy było zaproponowanie testu przesiewowego pozwalającego wykryć chorobę na wczesnym etapie, który będzie możliwy do zastosowania w codziennej praktyce lekarskiej, dzięki temu, że będzie opierał się na łatwo dostępnych danych oraz nie będzie generował wysokich kosztów.

W pierwszym etapie zaproponowano biomarkery oparte na danych z obrazowania rezonansem magnetycznym (ang. magnetic resonance imaging, MRI). Następnie zbudowano model, który jest propozycją testu przesiewowego mającego usprawnić wczesne wykrywanie pacjentów z chorobą Alzheimera oraz z łagodnym upośledzeniem funkcji poznawczych. Zaproponowany test przesiewowy jest szybki, nie wymaga dużej mocy obliczeniowej, oparty jest na łatwo dostępnych w historii medycznej pacjenta danych oraz nie generuje wysokich kosztów (w porównaniu do niektórych innych stosowanych metod).

Wczesne wykrycie choroby Alzheimera jest bardzo istotne z punktu widzenia pacjenta oraz systemu opieki zdrowotnej. Dla pacjenta jest szansą na wczesne rozpoczęcie leczenia łagodzącego objawy oraz daje możliwość skorzystania z terapii spowalniających rozwój choroby. Z punktu widzenia opieki zdrowotnej wczesne wykrycie choroby pozwala na racjonalizację kosztów, które rosną wraz z postępem choroby. Kolejnym aspektem jest możliwość prowadzenia badań nad lekami, które będą działały nie tylko objawowo, ale będą również hamowały rozwój choroby. Takie badania są możliwe do zrealizowania, kiedy są dostępni potencjalni uczestnicy tych badań, czyli osoby ze zdiagnozowaną chorobą na wczesnym



etapie. Wprowadzenie badań przesiewowych, które będą możliwe do przeprowadzenia przez lekarza w codziennej praktyce jest niezbędne i kluczowe, tym bardziej, że społeczeństwa starzeją się i liczba pacjentów z chorobą Alzheimera wzrasta.

Ubytki w mózgu pacjenta powstają dużo wcześniej przed wystąpieniem pierwszych objawów choroby, zatem zastosowanie technik obrazowania jest dobrym kierunkiem w identyfikacji chorych pacjentów. U pacjentów cierpiących na chorobę Alzheimera mózg kurczy się a wolne przestrzenie zostają wypełnione płynem mózgowo-rdzeniowym (ang. cerebrospinal fluid, CSF) [1] [2]. Ponadto, kurczenie się mózgu powoduje zwiększenie pofałdowania powierzchni mózgu oraz zwiększenie pojemności komór bocznych mózgu. W modelowaniu matematycznym rozważono zatem zmienne opisujące powyższe zjawiska. Zaproponowano własne wskaźniki opisujące kurczenie się oraz pofałdowanie mózgu: względną objętość mózgu bez płynu mózgowo-rdzeniowego oraz wskaźnik pofałdowania. Ponadto rozważono w modelu zmienną opisującą względną objętość komór bocznych mózgu.

## **1.2. Zawartość rozprawy**

Rozdział 1 zawiera omówienie celu pracy, sposobów realizacji tego celu oraz wyjaśnia dlaczego dany problem jest istotny z różnych punktów widzenia.

Rozdział 2 przedstawia czym jest choroba Alzheimera, jakie są objawy, przebieg, leczenie i rokowania. Rozdział ten zawiera omówienie procesu i kryteriów diagnozy choroby. Ponadto omówiono najbardziej charakterystyczne dla choroby Alzheimera oraz łagodnych zaburzeń poznawczych biomarkery. Przedstawiono różne metody obrazowania z informacją, czy dana metoda znajduje zastosowanie

w diagnostyce lub ocenie progresji choroby Alzheimera. Omówiono również pojawiające się najczęściej w literaturze metody klasyfikacyjne dotyczące choroby Alzheimera oraz łagodnych zaburzeń poznawczych.

Rozdział 3 zawiera omówienie metod składających się na wstępne przetwarzanie danych. W rozdziale tym również przedstawiono bazy danych, z których korzystano, oraz typ analizowanych danych. Omówiono metodę korekty niejednorodności pola obrazów MRI oraz metodę zastosowaną do ekstrakcji mózgu. Ponadto opisano dokładnie proces segmentacji obrazu MRI na płyn mózgowo-rdzeniowy i tkankę mózgową.

W rozdziale 4 omówiono zmienne zdefiniowane w celu opisu zmian zachodzących w mózgu osób cierpiących na chorobę Alzheimera. Zaprezentowano wzory oraz algorytmy służące do wyznaczania tych zmiennych. Omówiono zastosowane w analizie metody statystyczne. Opisano również konstrukcję modelu predykcyjnego oraz zaprezentowano wzory służące do obliczania wartości analizowanych parametrów.

W rozdziale 5 zamieszczono wyniki analiz: w tym wyniki otrzymanego modelu oraz wartości wskaźników predykcyjnych uzyskanych na zaleźnym (5-ktorna walidacja krzyżowa) oraz niezależnym zbiorze walidacyjnym.

Rozdział 6 zawiera dyskusję dotyczącą otrzymanych wyników w porównaniu do wyników z przeglądu literatury.

W rozdziale 7 zamieszczono podsumowanie zastosowanym metod oraz uzyskanych wyników.

Rozdział 8 zawiera wnioski oraz odpowiedź na postawioną w pracy tezę.

Na końcu pracy zamieszczono spis skrótów, spisy tabel oraz rysunków, a także bibliografię.

## 2. Wprowadzenie

### 2.1. Choroba Alzheimerera

Choroba Alzheimerera (AD) to postępująca, neurodegeneracyjna choroba mózgu, która powoduje utratę pamięci, zmiany w zachowaniu i problemy z codziennymi czynnościami. Choroba Alzheimerera jest najczęstszą postacią demencji i odpowiada za 60% do 80% przypadków demencji [3] [4]. Zazwyczaj cierpią na nią osoby powyżej 65 roku życia, może jednak wystąpić wcześniej. Według Światowej Organizacji Zdrowia (WHO) prawie 50 milionów ludzi cierpi na demencję, dlatego WHO uznaje demencję za priorytetową kwestię zdrowia publicznego [5]. Szacuje się, że 6,5 miliona Amerykanów w wieku od 65 lat wżwyż cierpi obecnie na chorobę Alzheimerera [6]. Szacuje się, że liczba ta może wzrosnąć do 13,8 milionów do 2060 roku, jeśli nie pojawią się przełomowe odkrycia medyczne mające na celu zapobieganie, spowalnianie lub leczenie choroby Alzheimerera. Choroba Alzheimerera była szóstą najczęstszą przyczyną zgonów w Stanach Zjednoczonych w 2019 r. i siódmą najczęstszą przyczyną zgonów w 2020 i 2021 r [6]. W Polsce w 2005 roku było 5,9 mln osób w wieku powyżej 60. roku życia cierpiących na zespoły otępienne (choroba Alzheimerera jest najczęstszą przyczyną demencji), a według prognoz, do 2030 roku liczba chorych wzrośnie do 9 milionów [7].

Etapem pośrednim pomiędzy normalnym funkcjonowaniem mózgu a demencją jest łagodne upośledzenie funkcji poznawczych (MCI). Chorzy z łagodnym upośledzeniem funkcji poznawczych mają wysoką wartość wskaźnika progresji w kierunku demencji w stosunkowo krótkim czasie, jednak nie u każdego chorego cierpiącego na łagodne upośledzenie funkcji poznawczych rozwinie się choroba Alzheimerera [8]. W 3-letnim okresie obserwacji około 35% pacjentów ze statusem

MCI przechodzi w AD lub otępienie [9]. Roczny współczynnik konwersji wynosi 5%-10% [9].

Chorobę Alzheimera po raz pierwszy opisał Alois Alzheimer w 1906 opisując objawy u pacjenta, którego znał od 1901 roku [10]. W połowie lat osiemdziesiątych ubiegłego wieku powstała bardziej szczegółowa definicja neuropatologiczna choroby Alzheimera, która rozpoznaje również współistniejące neuropatologie, które często przyczyniają się do klinicznej demencji [10].

Objawy choroby zwykle zaczynają się od drobnych problemów z pamięcią i postępują w kierunku upośledzenia funkcji poznawczych, powodujących dysfunkcje w rutynowych, codziennych zadaniach [5]. Poza problemami z pamięcią, które są najczęściej pierwszym objawem, wymienia się problemy z nastrojem, depresję, trudności językowe. Pojawiają się także kolejne objawy: problemy w wykonywaniu zwykłych czynności, problemy związane z funkcjami poznawczymi, dezorientacja, zmiany osobowości i zmiany behawioralne [11]. Początkowo osoba cierpiąca na chorobę Alzheimera dość dobrze pamięta wydarzenia z odległej przeszłości, a nie potrafi sobie przypomnieć co wydarzyło się kilka minut lub kilka godzin temu. Osoba cierpiąca na chorobę Alzheimera zapomina gdzie zostawiła klucze, okulary, nie pamięta wydarzeń bieżących. Z czasem zapomina co wydarzyło się w odległej przeszłości. W głębokim stadium choroby Alzheimera osoba chora nie pamięta kim jest, jak się nazywa, gdzie mieszka oraz nie rozpoznaje członków rodziny. Występujące zaburzenia orientacji początkowo dotyczą problemów z posługiwaniem się mapą i orientacją w nieznanym terenie, a następnie w terenie znanym od lat, aż do problemów z dotarciem do kuchni czy łazienki we własnym domu. U osób cierpiących na chorobę Alzheimera mogą występować również urojenia i omamy oraz

zaburzenia snu. U chorego następują zaburzenia rytmu dobowego, jest aktywny do późnych godzin nocnych, a śpi w ciągu dnia. Choroba Alzheimera postępuje powoli i w ostatnim stadium choroby objawia się utratą zdolności chodzenia, siedzenia, uśmiechania się, a zasób słownictwa zostaje ograniczony do około 6 słów [7].

Diagnoza choroby Alzheimera przebiega wielopłaszczyznowo. Opiera się na dokładnym wywiadzie i zbadaniu historii medycznej pacjenta, rozmowie z opiekunem, obejmuje również testy neuropsychologiczne dotyczące funkcji poznawczych, a także testy laboratoryjne (dodatkowo również testy dotyczące badania płynu mózgowo-rdzeniowego) oraz metody neuroobrazowania [12]. Stosowanymi testami neuropsychologicznymi są Krótka Skala Zaburzeń Poznawczych (ang. Mini Mental State Examination, MMSE) [13] oraz Montrealska Skala Oceny Funkcji Poznawczych (ang. Montreal Cognitive Assessment, MoCA) [14]. Testy te pozwalają wstępnie ocenić skalę zaburzeń i są obecnie stosowane jako testy przesiewowe, a także służą do oceny progresji choroby [15]. Do podstawowych metod neuroobrazowania należą: tomografia komputerowa (ang. computed tomography, CT) oraz rezonans magnetyczny (ang. magnetic resonance imaging, MRI). W diagnostyce wykorzystuje się także metodę pozytonowej tomografii emisyjnej (ang. positron emission tomography, PET) w celu obrazowania złogów  $\beta$ -amyloidu oraz białka tau [7].

W Polsce podstawowym kryteriami w rozpoznawaniu choroby Alzheimera są kryteria ICD-10 (ang. International Statistical Classification of Diseases and Related Health Problems) oraz DSM-V (ang. Diagnostic and Statistical Manual of Mental Disorders, wydanie piąte) [15] [16] [17].

W 2007 roku zaktualizowane zostały kryteria rozpoznawania choroby Alzheimera wg NINDCDS-ARDB (ang. National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association), które w Polsce mają charakter uzupełniający [18].

Średnia długość życia pacjenta z chorobą Alzheimera zmniejsza się od 8 do 10 lat od diagnozy potwierdzającej chorobę [5].

Obecne leczenie choroby Alzheimera polega wyłącznie na leczeniu objawowym. Do tej pory amerykańska Agencja Żywności i Leków (ang. Food and Drug Administration, FDA) zatwierdziła jedynie kilka leków łagodzących objawy choroby Alzheimera. Leki te są stosowane choć nie przynoszą pożądanego efektu terapeutycznego. Problemem jest wprowadzenie leków do mózgu, tak aby wywierały one konkretne działanie. Głównym ograniczeniem przedostawania się leków do mózgu jest bariera krew-mózg, a następnie bariera krew i płyn mózgowo-rdzeniowy. Bariery te mają charakter selektywnie przepuszczalnej błony, które umożliwiają wnikanie do mózgu jedynie określonych cząstek. Istnieje duża potrzeba odkrywania nowych leków oraz nowych technik wprowadzania leków do mózgu. Nadzieją są nanoleki, które są nowatorskim podejściem w leczeniu choroby Alzheimera. W leczeniu wykorzystuje się nanonośniki, które umożliwiają celowane dostarczanie leków [5].

Ostatnie doniesienia naukowe pokazują, że trwają badania nad nowym lekiem o nazwie Lecanemab, który obniża poziom amyloidu w mózgu, przez co spowalnia spadek funkcji poznawczych u pacjentów we wczesnych stadiach choroby Alzheimera [19] [20]. Terapia ta daje nadzieję na zmianę przebiegu choroby

i opóźnienie pojawiania się kolejnych objawów, szczególnie u pacjentów, u których choroba zdiagnozowana zostanie na bardzo wczesnym etapie.

## **2.1. Biomarkery choroby Alzheimerera**

Znalezienie odpowiednich biomarkerów choroby Alzheimerera oraz łagodnego upośledzenia funkcji poznawczych jest przedmiotem badań od wielu lat. Wczesne wykrycie choroby Alzheimerera oraz łagodnych zaburzeń poznawczych jest kluczowe z punktu widzenia podjęcia terapii i wydłużenia życia człowieka oraz złagodzenia objawów.

Obecnie wiadomo, że u pacjentów we wczesnym stadium choroby Alzheimerera w płynie mózgowo rdzeniowym obserwuje się obniżony poziom  $\beta$ -amyloidu  $A\beta_1 - 42$  i podwyższony poziom białka tau oraz fosforylowanego białka tau (p-tau) [21]. Biomarkery te wykrywane są w badaniach obrazowych PET (ang. positron emission tomography - pozytonowa emisyjna tomografia komputerowa) i płynie mózgowo-rdzeniowym (CSF). Badanie PET charakteryzuje się bardzo wysokimi kosztami i ma ograniczoną dostępność, natomiast badanie płynu mózgowo-rdzeniowego, mimo iż jest relatywnie niedrogi, to jest badaniem inwazyjnym [22].

Złogi białek w mózgu powodują znaczną utratę neuronów i w ten sposób prowadzą do zmniejszenia masy mózgu. Zmiany wolumetryczne powstałe w mózgu są możliwe do zaobserwowania za pomocą rezonansu magnetycznego. Obrazowanie MRI jest powszechnie stosowane w diagnozowaniu choroby Alzheimerera. Standardowo wykorzystuje się obrazy T1-zależne w celu pomiaru stopnia zaniku mózgu, zaniku istoty szarej, a także zmian w takich obszarach mózgu jak: przyśrodkowe obszary skroniowe, hipokamp czy boczne komory mózgowe [2]. Obrazowanie MRI pozwala także na zaobserwowanie zmian



w grubości kory mózgowej, która również jest biomarkerem [2]. Kolejnym biomarkerem neurodegeneracyjnym jest ocena metabolizmu glukozy ( $^{18}\text{F}$  - fluorodeoksyglukoza,  $^{18}\text{F}$ -FDG) za pomocą pozytonowej tomografii emisyjnej po podaniu fluorodeoksyglukozy (FDG-PET) [23].

Na podstawie przeglądu literatury biomarkery stosowane w diagnozie oraz klasyfikacji AD oraz MCI można podzielić na trzy grupy.

Pierwszą grupą biomarkerów są biomarkery związane z płynem mózgowo-rdzeniowym dotyczące biochemicznej oceny poziomu białek  $\beta$ -amyloid  $\text{A}\beta$ 1-42, tau oraz p-tau [24] [25] [26]. Drugą grupą są biomarkery uzyskane z obrazowania MRI oparte na zmianach strukturalnych w budowie mózgu [27] [28] [29] [30]. Trzecią grupą są biomarkery opisujące zmiany metaboliczne mózgu mierzone za pomocą FDG-PET [31] [32].

Ponadto można wyróżnić podgrupę obrazowania MRI stosowaną w diagnozowaniu choroby Alzheimera oraz łagodnych zmian poznawczych: obrazowanie tensora dyfuzji (ang. diffusion tensor imaging, DTI) i funkcjonalne obrazowanie metodą rezonansu magnetycznego (ang. functional MRI, fMRI) [33] [24] [34] [35]. Wiele badań wykorzystuje kombinacje biomarkerów z dwóch lub większej liczby grup biomarkerów w celu wczesnej diagnozy AD i MCI, np. biomarkerów opartych o obrazowanie MRI, biomarekrów związanych z badaniem płynu mózgowo-rdzeniowego lub biomarkerów związanych z PET [33] [36] [24] [25] [37]. Uzyskanie biomarkerów ze wszystkich trzech grup: PET, CSF i MRI (lub DTI lub fMRI) jest bardzo kosztowne, czasochłonne a część metod jest inwazyjna [38] [37].

## **2.2. Metody obrazowania**

W medycynie stosuje się wiele różnych technik obrazowania. Najpopularniejsze z nich to: rentgenografia, tomografia komputerowa, rezonans magnetyczny oraz pozytonowa tomografia emisyjna. W kolejnych podrozdziałach omówiono wymienione techniki obrazowania.

### **2.2.1. Rentgenografia**

Rentgenografia (ang. X-ray) polega na wykorzystaniu promieni X do obrazowania narządów wewnętrznych. Obraz powstaje w taki sposób, że ciało człowieka lub pewien fragment ciała zostają poddane działaniu promieni X. Promienie te przenikają ciało i są w różnym stopniu pochłaniane przez różne narządy, co powoduje powstanie obrazu z jaśniejszymi i ciemniejszymi obszarami. Promienie X nazywane są promieniami rentgenowskimi od nazwiska Wilhelma Roentgena, który dokonał odkrycia tych promieni w 1895 roku. W 1901 roku Roentgen otrzymał nagrodę Nobla za osiągnięcie, które zrewolucjonizowało medycynę, pozwalając wytworzyć obraz wnętrza ciała człowieka w sposób bezinwazyjny [39].

Promieniowanie rentgenowskie ma jednak szkodliwe działanie na organizm człowieka. Kolejną wadą tej techniki obrazowania jest to, że nie można zobaczyć narządów, które są ukryte za innymi narządami stąd rentgenografia nie nadaje się do obrazowania mózgu, który zasłaniany jest przez kości czaszki. Technika rentgenowska najczęściej stosowana jest do obrazowania kości oraz płuc.

### **2.2.2. Ultrasonografia**

Ultrasonografia (USG) wykorzystuje wiązki ultradźwięków, które odbijają się od powierzchni narządów i w ten sposób powodują powstawanie obrazu [39]. Do zalet ultrasonografii należą: mobilność, możliwość obrazowania narządów wewnętrznych oraz ich ruchu, możliwość obrazowania przepływów (np. przepływu krwi), możliwość trójwymiarowej rekonstrukcji ruchomych obiektów (np. płodu), niski koszt oraz bezinwazyjność. Technika ta posiada również wady: niską jakość obrazu, występujące liczne artefakty, trudności w interpretacji obrazu. Ultrasonografię można stosować wspomagająco w diagnozowaniu choroby Alzheimera a także jako narzędzie do obserwacji progresji choroby [40].

### **2.2.3. Tomografia komputerowa**

Tomografia komputerowa (ang. computed tomography, CT) jest techniką obrazowania opartą na rentgenografii ale o wiele bardziej skomplikowaną. W tomografii komputerowej wykonuje się serię zdjęć rentgenowskich pod różnym kątem, dzięki czemu uzyskuje się liczne przekroje wybranych struktur. Następnie przekroje te są nakładane na siebie i łączone w obraz 2D lub 3D za pomocą specjalistycznego oprogramowania, stąd określenie tomografia *komputerowa*. Obraz uzyskany metodą tomografii komputerowej dostarcza zdecydowanie więcej informacji niż klasyczne zdjęcie rentgenowskie. Do zalet tej metody obrazowania należą: bardzo dobra jakość uzyskanych obrazów oraz nieinwazyjność. Wadami metody są: brak różnicowania tkanek miękkich oraz szkodliwość wynikająca ze stosowania promienia jonizującego. Dla porównania, pojedyncze badanie CT klatki piersiowej wiąże się z dawką około 7-10 mSv, podczas gdy zdjęcie rentgenowskie klatki piersiowej to 0,1 mSv [41].

Pojawienie się tomografii komputerowej w początkach lat siedemdziesiątych zwiastowało rewolucję w nieinwazyjnym obrazowaniu mózgu. Początkowo w procesie diagnozy choroby Alzheimera stosowano CT w celu wykluczenia innych przyczyn demencji, dopiero później obrazowanie tą metodą było wykorzystywane w celu potwierdzenia atrofii mózgu [42] [43].

#### **2.2.4. Rezonans magnetyczny**

Rezonans magnetyczny (ang. magnetic resonance image, MRI) jest nieinwazyjną metodą uzyskiwania obrazów medycznych. Obecnie jest jedną z podstawowych technik w diagnostyce obrazowej. Obraz powstaje dzięki wykorzystaniu silnego pola magnetycznego oraz fal radiowych. Obrazowanie za pomocą rezonansu magnetycznego nie wymaga zastosowania szkodliwego promieniowania jonizującego, jak w przypadku zdjęć rentgenowskich czy tomografii komputerowej, co jest kluczową zaletą tej metody. Kolejną zaletą to bardzo dobre obrazowanie zmian w tkankach chorobowych, przez co rezonans magnetyczny może być wykorzystywany do obrazowania tkanek, nawet tych, które zasłonięte są kośćmi. Dzięki tym właściwościom rezonans magnetyczny ma szerokie zastosowanie w obrazowaniu mózgu. Wadą jest koszt badania, który jest znacznie wyższy niż w przypadku zastosowania rentgenografii czy tomografii komputerowej.

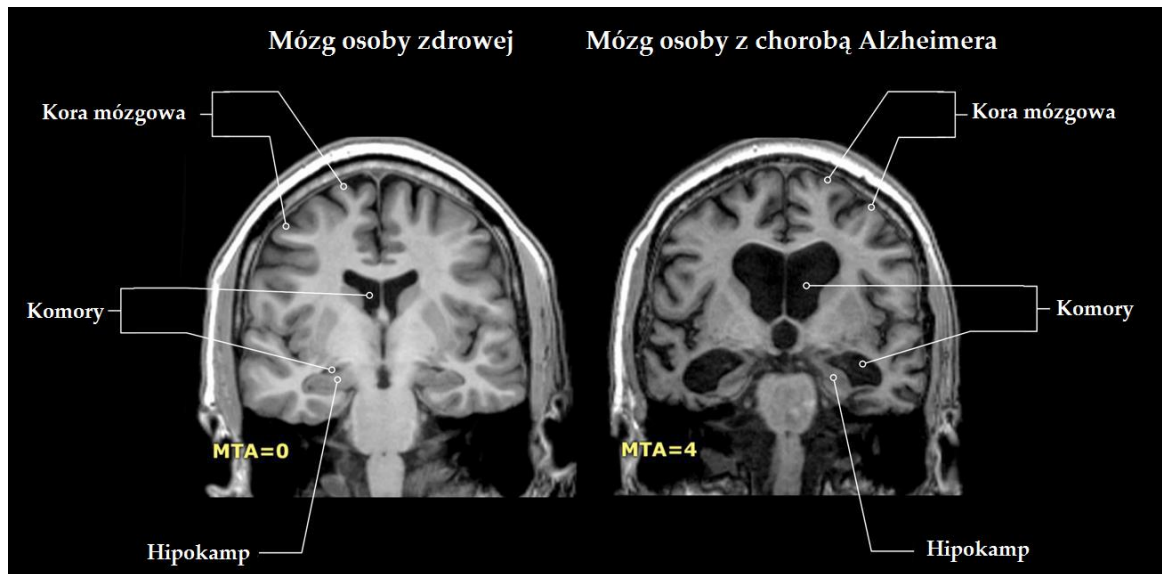
Początkowo obrazowanie metodą rezonansu magnetycznego, podobnie jak tomografia komputerowa, wykorzystywane było jedynie w celu wykluczenia innych przyczyn demencji niż choroba Alzheimera, później jednak stało się ważnym narzędziem diagnostycznym, ze względu na możliwość zaobserwowania atrofii mózgu [44]. Metoda ta pozwala na różnicowanie tkanek, które w innych metodach obrazowania były identyczne [39].

Obrazowanie rezonansem magnetycznym może być przeprowadzone w różnych sekwencjach w zależności od potrzeb diagnostycznych. Ze względu na sekwencje obrazowanie można podzielić na:

- Obrazy T1-zależne – tkanki o wysokiej zawartości tłuszczu (takie jak istota biała) się jasne, a obszary wypełnione wodą (CSF) się ciemne. Obrazy te najlepiej odzwierciedlają strukturę anatomiczną mózgu.
- Obrazy T2-zależne – obszary wypełnione wodą (np. przez płyn mózgowo-rdzeniowy) przedstawiane są w jasnych odcieniach, a tkanki o wysokiej zawartości tłuszczu (takie jak istota biała) przedstawiane są w ciemnych odcieniach. Obrazy stosowane są często do wykazywania patologii, ponieważ część zmian np. uszkodzenia w tkance, mają tendencje do rozwoju obrzęku, co wiąże się ze wzrostem zawartości wody.
- FLAIR (ang. fluid-attenuated inversion recovery) – jest to modyfikacja sekwencji T2-zależnej, w której sygnał z płynu mózgowo-rdzeniowego jest tłumiony, co powoduje, że CSF przedstawiony jest w bardzo ciemnych odcieniach.
- Obrazowanie dyfuzyjne (ang. diffusion-weighted imaging, DWI) – sekwencja ta umożliwia obrazowanie swobodnego ruchu molekuł wody w przestrzeni wewnątrzkomórkowej i zewnątrzkomórkowej. Zaburzenia dyfuzji możliwe są do interpretacji dzięki rzeczywistemu współczynnikowi dyfuzji (ang. apparent diffusion coefficient, ADC) pozwalającemu na obrazowanie rozkładu dyfuzji poprzez mapy ADC [45]. Obrazowanie DWI jest również wykorzystywane w detekcji choroby Alzheimera [46].
- Obrazowanie tensora dyfuzji (ang. diffusion tensor imaging, DTI) – obrazowanie DTI opisuje wartość i kierunek dyfuzji protonów wodoru

molekuł wody. Informacji o kierunku dyfuzji dostarczają mapy FA (ang. fractional anisotropy). Obrazowanie DTI pozwala również na tworzenie map średniej dyfuzyjności (ang. mean diffusivity, MD) oraz map dróg nerwowych istoty białej (traktografia) [45]. Obrazowanie tensora dyfuzji (DTI) to technika nieinwazyjna, która wykrywa zmiany mikrostrukturalne w istocie białej, co może być wykorzystywane w ocenie progresji choroby Alzheimera [47].

- Funkcjonalne obrazowanie metodą rezonansu magnetycznego (ang. functional magnetic resonance imaging, fMRI) – jest to rozszerzenie obrazowania rezonansem magnetycznym (MRI), które pozwala na obserwację zmian w utlenowaniu krwi w aktywowanych obszarach mózgu. Gdy uruchamia się obserwowany obszar mózgu, to zwiększa się zapotrzebowanie na tlen a sygnał odbierany z aktywnych rejonów mózgu przekładany jest na obraz. Funkcjonalne obrazowanie metodą rezonansu magnetycznego może być wsparciem w diagnozowaniu choroby Alzheimera poprzez obserwację zmian aktywności pewnych obszarów mózgu pomiędzy zdrowymi starszymi osobami, a pacjentami z łagodnym upośledzeniem funkcji poznawczych oraz pacjentami z chorobą Alzheimera [48].



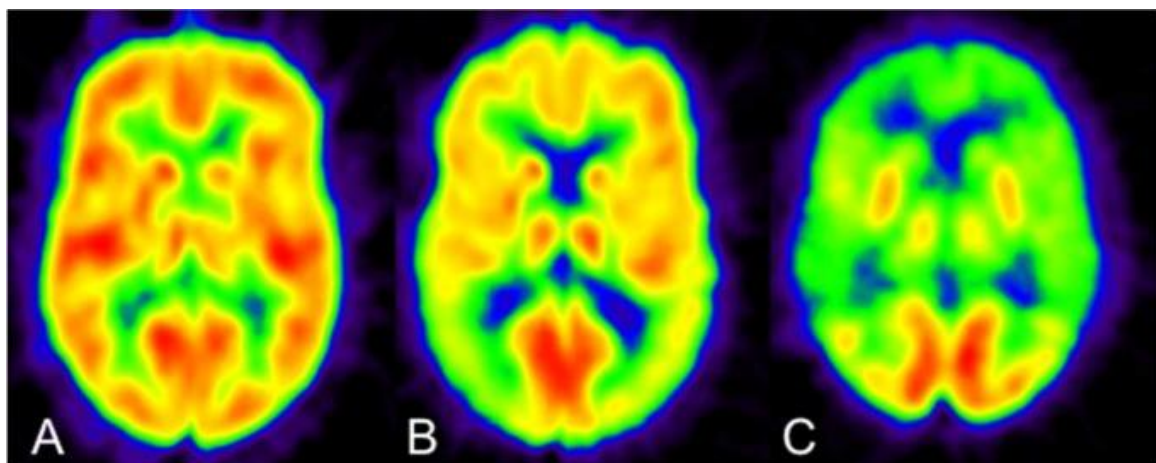
Rysunek 1 Obraz z rezonansu magnetycznego: po lewej stronie mózg osoby zdrowej, po prawej mózg osoby z chorobą Alzheimera. Mózg osoby z chorobą Alzheimera charakteryzuje się większym pofałdowaniem kory mózgowej, powiększonymi komorami oraz zmniejszonym hipokampem. MTA (ang. medial temporal lobe atrophy): MTA = 0 oznacza brak atrofii w przyśrodkowej części płata skroniowego, MTA=4 oznacza poważną utratę objętości hipokampu. Źródło: [49].

### 2.2.5. Pozytonowa tomografia emisyjna

Pozytonowa tomografia emisyjna (ang. positron emission tomography, PET) jest techniką obrazowania, w której obserwuje się promieniowanie powstałe w skutek rozpadu izotopów podanych pacjentowi przed badaniem. Podana pacjentowi substancja promieniotwórcza zawiera izotopy o krótkim czasie połowicznego rozpadu, co wymaga uruchomienia w pobliżu cyklotronu, dlatego, że maksymalny czas transportu jest krótki, ze względu na krótki czas połowicznego rozpadu. Obrazowanie PET pozwala na zbadanie funkcjonalności i aktywności narządów wewnętrznych. Badanie PET jest przydatne w obrazowaniu zaburzeń mózgu i jest stosowane w diagnozowaniu choroby Alzheimera. W przypadku choroby Alzheimera stosuje się pozytonową tomografię emisyjną w celu obrazowania złogów  $\beta$ -amyloidu oraz białka tau [7].

Wadą obrazowania PET jest długi czas badania, znacznie dłuższy niż czas badania tomografem komputerowym czy czas wykonania zdjęcia rentgenowskiego. Badanie jest nieinwazyjne ale wymaga podania środka promieniotwórczego, którego działanie może być potencjalnie niebezpieczne, co jest niewątpliwie minusem badania PET. Ponadto badanie to jest kosztowne, jest znacznie droższe niż zwykłe badanie rezonansem magnetycznym czy tomografem komputerowym oraz ma ograniczoną dostępność.

W diagnozowaniu choroby Alzheimera stosuje się również pozytonową tomografię emisyjną po podaniu fluorodeoksyglukozy (FDG-PET). Obrazowanie FDG-PET pokazuje metabolizm glukozy w tkance o obniżonej aktywności, czyli w obszarach dysfunkcji i degeneracji neuronów [50].



Rysunek 2 Wyniki obrazowania PDG-PET dla: A – zdrowej osoby, B – osoby z wczesnym stadium choroby Alzheimera, C – osoby z zaawansowaną chorobą Alzheimera. Źródło: [32]

### 2.3. Metody klasyfikacyjne

Wiele różnych metod klasyfikacyjnych zostało zaproponowanych w ostatnich latach w odniesieniu do choroby Alzheimera oraz łagodnego upośledzenia funkcji poznawczych. Wiele z tych metod bazuje na algorytmach uczenia maszynowego

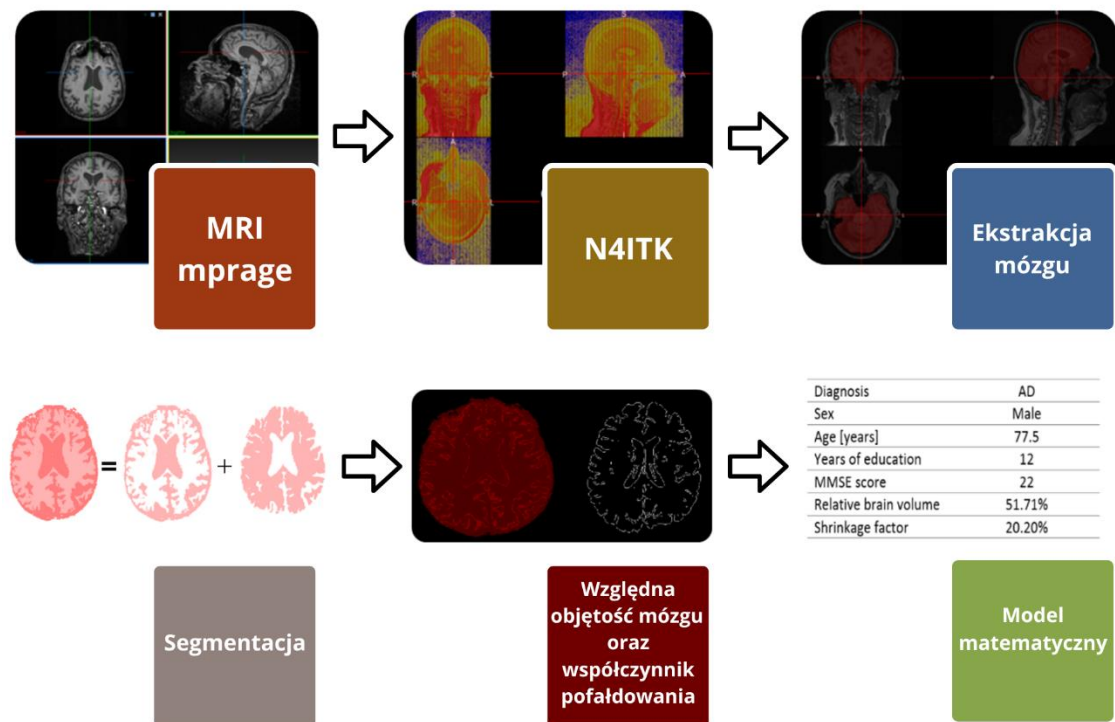


[51], najczęściej na konwolucyjnych sieciach neuronowych (ang. convolutional neural networks, CNN) [27] [28] [25] [33] [52] [30], metodzie wektorów nośnych (ang. support vector machines, SVM) [24] [29] [53], modelach regresyjnych [9] [54] [55] [56] [57] [58] [59], a także na innych metodach klasyfikacyjnych, takich jak metoda grafów [60], las losowy (ang. random forest, RF) [29] [59], drzewa decyzyjne (ang. decision tree, DT) [29], liniowa analiza dyskryminacyjna (ang. linear discriminant analysis, LDA) [61], ortogonalna metoda częściowych najmniejszych kwadratów (ang. orthogonal partial least-squares, OPLS) [37], sieć bayesowska (ang. bayes network) [62], czy algorytm genetyczny (ang. genetic algorithm) [63]. Stosowane są również kombinacje dwóch lub kilku metod, np. modelu regresyjnego i lasu losowego [59]. Część z wymienionych metod to metody bardzo zaawansowane, bardzo kosztowne obliczeniowo (np. CNN), wymagające bardzo dużej próby danych, wrażliwe na protokoły pomiaru i wstępne przetwarzanie, przez co są trudne do zastosowania w codziennej praktyce lekarskiej, pomimo, że dają bardzo dobre wyniki.

### **3. Wstępne przetwarzanie danych**

W poniższym rozdziale zamieszczono charakterystykę baz danych, na których przeprowadzona została analiza a także omówiono typ analizowanych danych (MPRAGE). W kolejnych podrozdziałach przedstawiono i szczegółowo omówiono kroki wstępnego przetwarzania danych.

Pierwszym krokiem jest korekta niejednorodności pola obrazu MRI za pomocą algorytmu N4ITK. Drugi krok to ekstrakcja mózgu. W trzecim kroku następuje segmentacja na płyn mózgowo-rdzeniowy i tkankę mózgową za pomocą dostosowanego do naszych potrzeb algorytmu MiMSeg. Kolejne kroki omówione już w kolejnym rozdziale, dotyczącym analizy danych, to zdefiniowanie zmiennych opisujących zmiany zachodzące w mózgu osoby cierpiącej na łagodne upośledzenie funkcji poznawczych lub chorobę Alzheimera. Wszystkie te kroki prowadzą do zbudowania modelu predykcyjnego. Na poniższym schemacie przedstawiono kolejne kroki wstępnego przetwarzania danych prowadzące do głównej analizy.



Rysunek 3 Schemat przedstawiający kolejne kroki wstępnego przetwarzania danych prowadzące do głównej analizy.

### 3.1. Bazy danych

Analiza została przeprowadzona na dwóch zbiorach danych. Zbiorem uczącym był zbiór danych z bazy ADNI (The Alzheimer's Disease Neuroimaging Initiative), a zbiór testowy pochodził z bazy EDSD (The European DTI Study on Dementia).

ADNI to wieloośrodkowe badanie obserwacyjne, którego celem jest zebranie serii danych z obrazowania rezonansem magnetycznych, pozytonowej tomografii emisyjnej oraz danych dotyczących biomarkerów biologicznych a także danych klinicznych i neuropsychologicznych pozwalających na ocenę progresji łagodnych zaburzeń poznawczych i wczesnej choroby Alzheimera. Badanie ADNI rozpoczęło się w 2004 roku i obecnie obejmuje 63 ośrodki w Stanach Zjednoczonych i Kanadzie. Początkowo pięcioletnie badanie ADNI-1 zostało

w 2009 roku przedłużone o 2 lata dzięki grantowi „Grand Opportunities” (ADNI-GO), a w 2011 i 2016 r. przez kolejne odnowienia grantu ADNI-1 (ADNI-2, ADNI-3). W analizie wykorzystano podzbiór z bazy ADNI-1: 638 pacjentów, którzy ukończyli pierwszy rok obserwacji i posiadają wyniki z obrazowania (1.5 Tesla) w punkcie początkowym oraz po 6 i 12 miesiącach obserwacji. Dane podzielono na trzy podgrupy ze względu na status chorobowy: pacjenci cierpiący na chorobę Alzheimera (ang. Alzheimer’s disease, AD), pacjenci z łagodnym upośledzeniem funkcji poznawczych (MCI) oraz grupa kontrolna, która stanowią zdrowe, starsze osoby (ang. normal control, NC) [64].

Drugi zbiór danych, który jest niezależnym zbiorem testowym w analizie pochodzi z bazy ESDS. Badanie ESDS („The European DTI Study on Dementia”) rozpoczęło się w 2010 roku. Koordynatorem badania jest Niemieckie Centrum Chorób Neurodegeneracyjnych (DZNE). Początkowo dane z neuroobrazowania, MRI i DTI, były retrospektywnie zbierane od zdrowych osób oraz od pacjentów z chorobą Alzheimera w 10 europejskich ośrodkach. W 2013 roku rozpoczęto zbieranie danych również od pacjentów z łagodnym upośledzeniem funkcji poznawczych (MCI). Baza stale się rozrasta powiększając również o dane dotyczące biomarkerów oraz o dane kliniczne. Baza zawiera obrazy typu FLAIR, T1- zależne MPRAGE (1.5 oraz 3.0 Tesla) oraz DTI (nie dla każdego pacjenta dostępne są wszystkie trzy typu obrazów) [65].

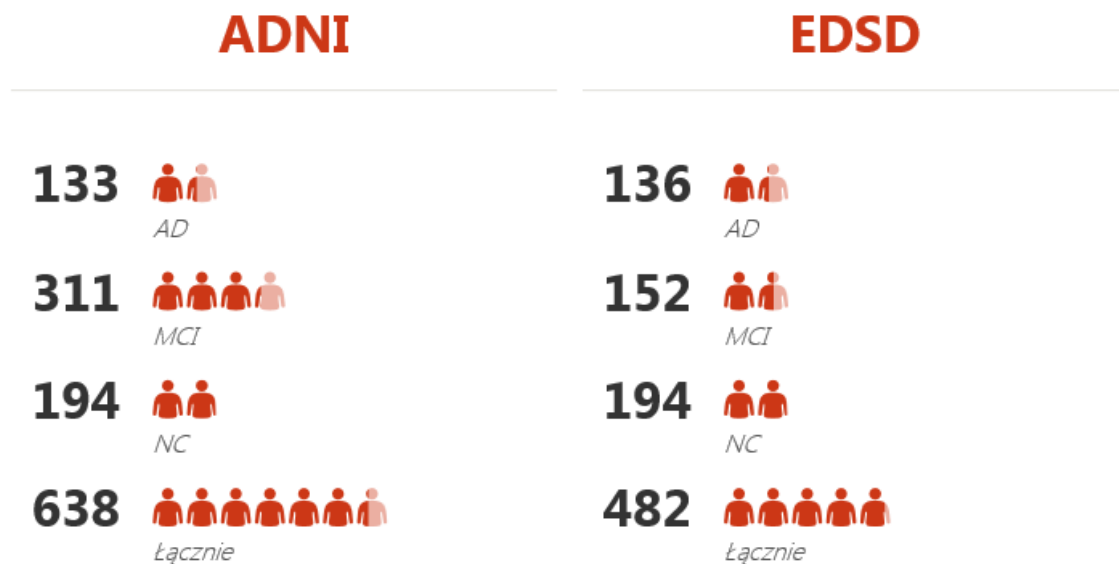
### **3.2. Struktura analizowanych danych (MPRAGE)**

Analizę rozpoczęto od danych MRI typu FLAIR (ang. fluid attenuated inversion recovery). Są to obrazy, w których obszary z małą ilością wody przedstawione są w ciemniejszych barwach, a obszary z dużą ilością wody w jaśniejszych. Jednak

po wstępnej analizie i porównaniu obrazów typu FLAIR i T1-zależny MPRAGE (ang. magnetization-prepared rapid gradient-echo) podjęto decyzję o przeprowadzeniu całej analizy na danych typu T1-zależny MPRAGE. MPRAGE to jedna z najpopularniejszych sekwencji do strukturalnego obrazowania mózgu. Sekwencja rejestruje wysoki kontrast tkanek i zapewnia wysoką rozdzielczość przestrzenną z pokryciem całego mózgu w krótkim czasie skanowania oraz pozwala na późniejszą rekonstrukcję obrazu w trzech planach: poprzecznym, strzałkowym i czołowym. Te cechy obrazowania sekwencją MPRAGE sprawiają, że bardzo dobrze spełnia ona swoją rolę w diagnozowaniu choroby Alzheimera oraz łagodnego upośledzenia funkcji poznawczych [66].

### **3.3. Charakterystyka kliniczna pacjentów z baz ADNI oraz EDSD**

Analizę przeprowadzono na danych dotyczących 638 pacjentów (133 pacjentów z chorobą Alzheimera, 311 z łagodnym upośledzeniem funkcji poznawczych oraz 194 zdrowe, starsze osoby) pochodzących z bazy ADNI, dokładnie z ADNI-1 zawężonej do pacjentów, którzy ukończyli pierwszy rok obserwacji i posiadają wyniki z obrazowania (1.5 Tesla) w punkcie początkowym oraz po 6 i 12 miesiącach obserwacji. Jako niezależny zbiór testowy analizowano dane 482 pacjentów pochodzące z bazy EDSD (136 pacjentów z chorobą Alzheimera, 152 pacjentów z łagodnym upośledzeniem funkcji poznawczych oraz 194 zdrowe, starsze osoby). Do analizy włączono pacjentów oznaczonych w bazie jako „niewykluczony”. Baza EDSD zawiera wyniki obrazowania dla części pacjentów dla 1,5 Tesla, a dla części 3 Tesla. Poza danymi z obrazowania wyciągnięto z baz wspólne dostępne dane, takie jak: płeć, wiek w momencie diagnozy, liczba lat edukacji, wynik Krótkiej Skali Zaburzeń Poznawczych (MMSE, ang. Mini-Mental State Examination) [13].



Rysunek 4 Liczebność analizowanych zbiorów danych z baz ADNI oraz EDSD z podziałem na status diagnozy.

Krótką Skalą Zaburzeń Poznawczych (MMSE) to proste narzędzie przesiewowe w postaci kwestionariusza. Maksymalna liczba punktów, które można uzyskać to 30. Przyjmuje się według kryteriów DSM-V (ang. Diagnostic and Statistical Manual of Mental Disorders – klasyfikacja zaburzeń psychicznych Amerykańskiego Towarzystwa Psychiatrycznego, wydanie piąte) oraz ICD-10, że wynik niższy niż 24 punkty w skali MMSE sugeruje obecność zespołu otępiennego. Ze względu na to, iż wiek oraz poziom wykształcenia wpływają na wynik testu proponuje się, aby pacjentów z wynikiem poniżej 27 punktów skierować na dalsze badania mające na celu wykluczenie lub potwierdzenie występowania zespołu otępiennego.

W poniższej tabeli została przedstawiona charakterystyka kliniczna pacjentów z bazy ADNI oraz z bazy EDSD, którzy zostali włączeni do analizy. Pacjentów podzielono na trzy grupy ze względu na status chorobowy: zdrowe osoby (NC,

grupa kontrolna), pacjenci z łagodnym upośledzeniem funkcji poznawczych (MCI), pacjenci z chorobą Alzheimera (AD). Dla zmiennych ciągłych (wiek, liczba lat edukacji, punkty MMSE) zostały obliczone wartości średniej oraz odchylenia standardowego a dla zmiennych skategoryzowanych (płeć) wynik przedstawiono w postaci odsetka. Porównania pomiędzy grupami zostały przeprowadzone przy użyciu nieparametrycznego testu Kruskala-Wallis dla zmiennych ciągłych (w analizie post-hoc zastosowany został test wariancji rang Conovera) oraz testu  $\chi^2$  do porównania proporcji dla danych skategoryzowanych. Dla każdego testowania statystycznego przedstawiono wartość  $p$ . Ponadto dla zmiennych ciągłych obliczono wartości wielkości efektu  $\eta^2$  (ang. eta squared) wraz z 95% przedziałem ufności [67] [68].

Tabela 1 Charakterystyka kliniczna analizowanych zbiorów z baz ADNI i EDSD.

Charakterystyka	NC	MCI	AD	Wartość $p$	Wielkość efektu $\eta^2$ [95%CI]
ADNI, n	194	311	133	-	-
Wiek, średnia (SD) [lata]	75,9 (5,08)	74,9 (7,06)	74,7 (7,59)	0,4643 *	0,0053 [0; 0,0199]
Edukacja, średnia (SD) [lata]	16,0 (2,79)	15,7 (3,00)	14,7 (3,11)	0,0003 * 0,0004 † < 0,0001 ‡	0,0253 [0,0057; 0,0523]
MMSE, średnia (SD)	29,1 (1,03)	27,0 (1,78)	23,5 (1,91)	< 0,0001 * < 0,0001 † ‡ §	0,6023 [0,5581; 0,6387]
Kobiety [%]	47,9	35,4	48,1	0,0053 **	-
EDSD, n	194	152	136	-	-
Wiek, średnia (SD), [lata]	68,7 (5,90)	71,2 (6,76)	72,4 (8,28)	< 0,0001 * 0,0001 † < 0,0001 §	0,0497 [0,0170; 0,0900]

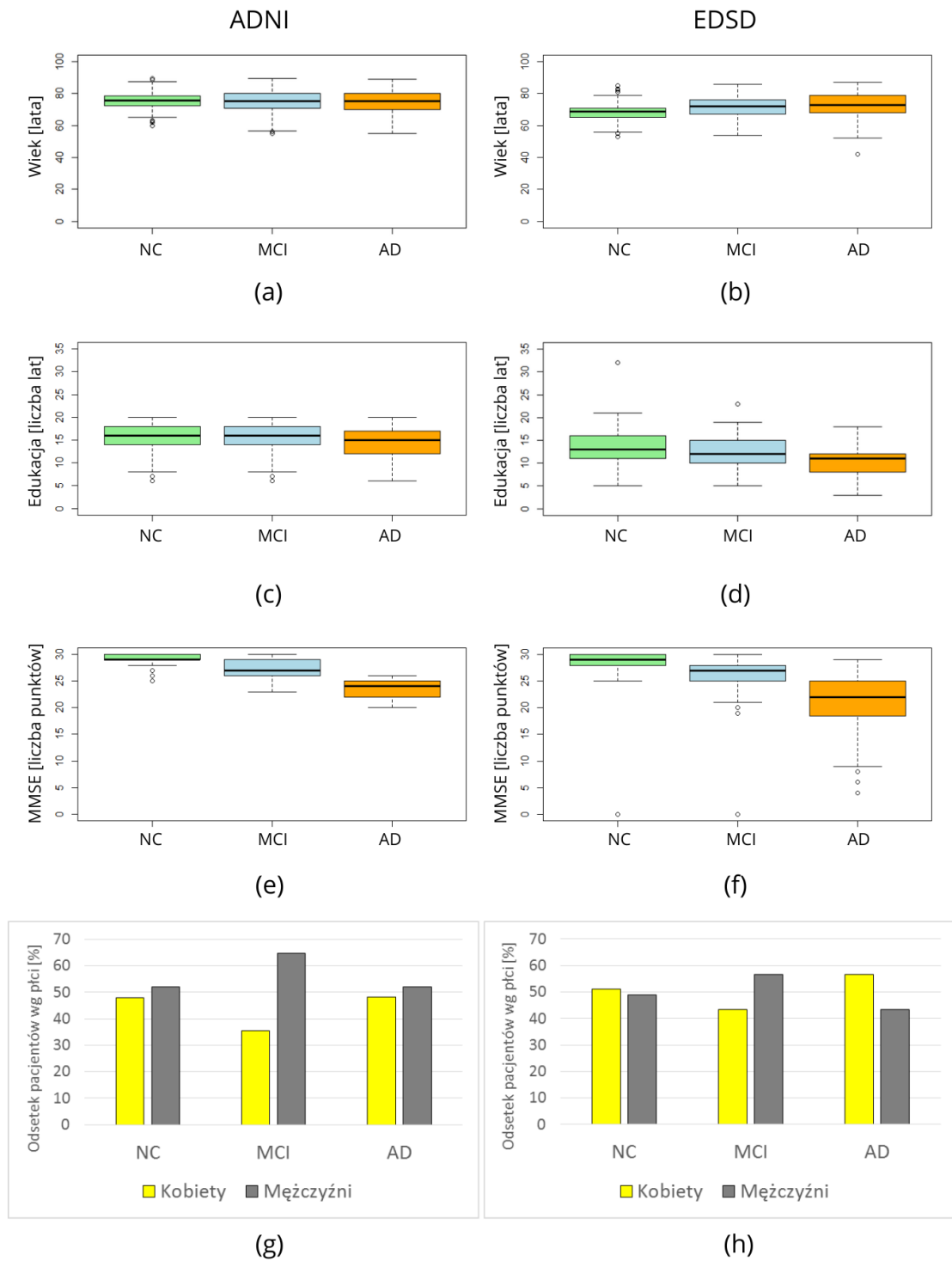
Edukacja, średnia (SD), [lata]	13,1 (3,67) n=173	12,4 (3,35) n=132	10,3 (3,33) n=134	< 0,0001 * < 0,0001 ††	0,1036 [0,0538; 0,1572]
MMSE, średnia (SD)	27,4 (6,49)	26,3 (3,14)	20,8 (5,36)	< 0,0001 * < 0,0001 †‡§	0,2198 [0,1569; 0,2793]
Kobiety [%]	51,0	43,4	56,6	0,07854 **	-

\*Test Kruskala-Wallisa; † Test Conovera: AD vs MCI; ‡ Test Conovera: AD vs NC; § Test Conovera: MCI vs NC; \*\* Test chi<sup>2</sup>

Wyniki charakterystyki klinicznej dla ADNI pokazują, że różnice pomiędzy co najmniej dwoma medianami są statystycznie istotne dla następujących zmiennych: liczba lat edukacji oraz punkty MMSE. Dla EDSD, różnice pomiędzy dwoma medianami są istotne statystycznie dla wszystkich zmiennych ciągłych: wiek, liczba lat edukacji, punkty MMSE. Wartość wielkości efektu dla zmiennej wiek jest bardzo mała dla danych z obu baz: ADNI i EDSD. Wartość wielkości efektu dla edukacji jest mała dla zbioru ADNI i średnia dla zbioru EDSD, natomiast wartość wielkości efektu dla punktów MMSE jest duża dla zbioru ADNI i bardzo duża dla zbioru EDSD. Wynik testu chi<sup>2</sup> dla proporcji informuje nas, że hipoteza „proporcja kobiet jest taka sama w grupie NC, MCI i AD” powinna być odrzucona dla zbioru ADNI, ale nie dla zbioru EDSD.

Na poniższym rysunku przedstawiono w postaci wykresów ramka-wąsy dla danych ciągłych (wiek, liczba lat edukacji, punktacja MMSE) oraz w postaci wykresów kolumnowych dla danych skategoryzowanych (płeć) dane z powyższej tabeli zawierającej charakterystykę kliniczną analizowanych zbiorów ADNI oraz EDSD.



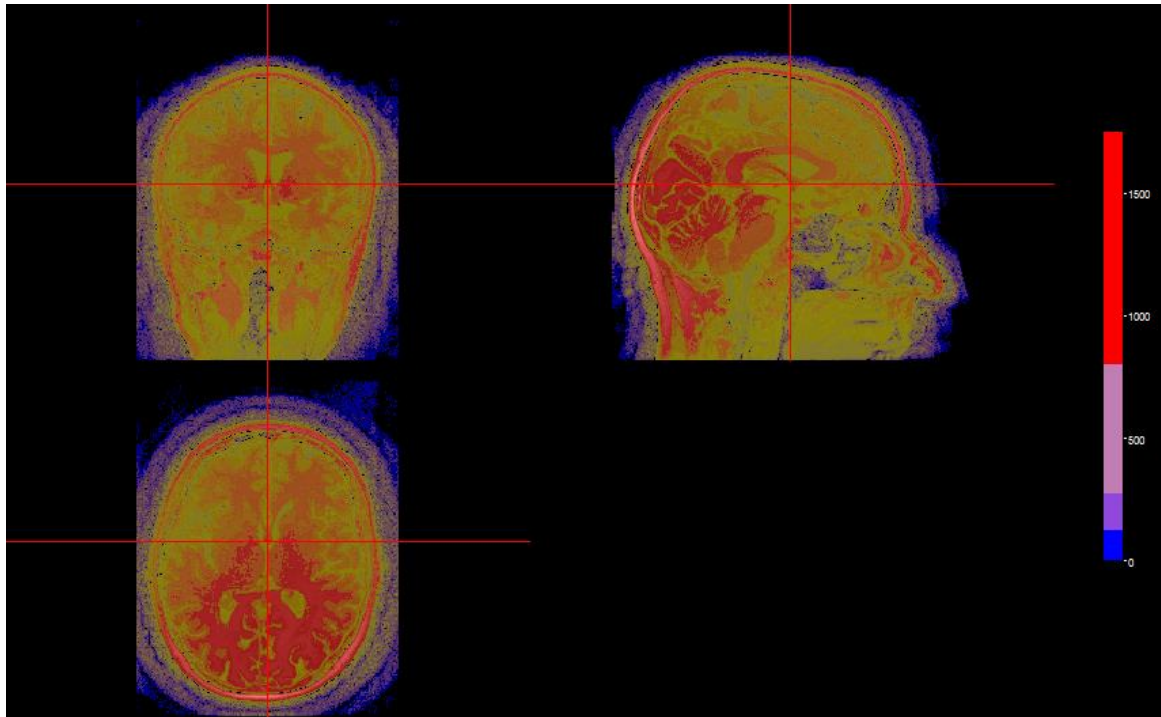


Rysunek 5 (a) Wiek pacjentów ze zbioru ADNI w momencie diagnozy dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (b) Wiek pacjentów ze zbioru EDSD w momencie diagnozy dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (c) Liczba lat edukacji pacjentów ze zbioru ADNI dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (d) Liczba lat edukacji pacjentów ze zbioru EDSD dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (e) Liczba punktów MMSE pacjentów ze zbioru ADNI dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (f) Liczba punktów MMSE pacjentów ze zbioru EDSD dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (g) Odsetek kobiet i mężczyzn z analizowanego zbioru ADNI dla grup: NC, MCI, AD. Wykres kolumnowy. (h) Odsetek kobiet i mężczyzn z analizowanego zbioru EDSD dla grup: NC, MCI, AD. Wykres kolumnowy.

### **3.4. Korekta niejednorodności natężenia pola**

Korekta niejednorodności natężenia pola to pierwszy krok wstępnego przetwarzania danych. Baza ADNI zawiera obrazy ze znormalizowanym natężeniem pola polaryzacji [64]. W przypadku bazy EDSO dostępne były dane natywne, do których zastosowano korektę niejednorodności natężenia pola za pomocą algorytmu N4ITK [69]. Niejednorodność natężenia pola charakteryzuje się niższą częstotliwością natężenia w niektórych obszarach obrazu MRI, co utrudnia np. segmentację obrazu.

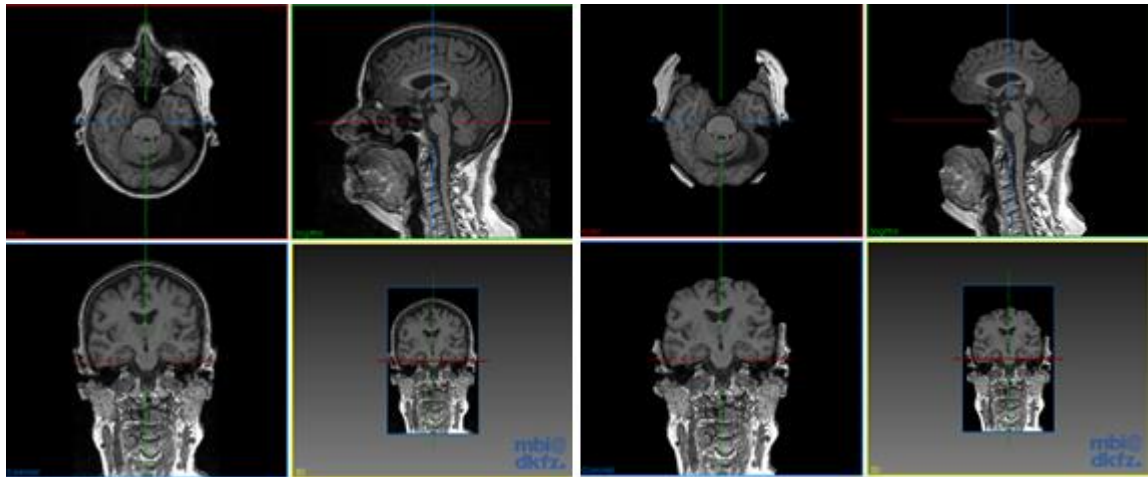
Do korekty niejednorodności pola zastosowano algorytm N4 korekty błędu pola (N4 bias field correction) [69]. Algorytm N4 jest wariantem popularnego nieparametrycznego algorytmu niejednorodnej normalizacji natężenia (N3) do korekcji pola polaryzacji [70]. Na poniższym rysunku przedstawiono w postaci mapy ciepła wynik porównania obrazów MRI mózgu przed korektą niejednorodności natężenia pola i po korekcie. Obszary zaznaczone na czerwono charakteryzują się największą różnicą (przed - po korekcie) natężenia pola a obszary oznaczone kolorem niebieskim najniższą.



Rysunek 6 Mapa cieplna przedstawiająca wynik porównania obrazów MRI mózgu przed korektą niejednorodności natężenia pola i po korekcie metodą N4ITK.

### 3.5. Ekstrakcja mózgu

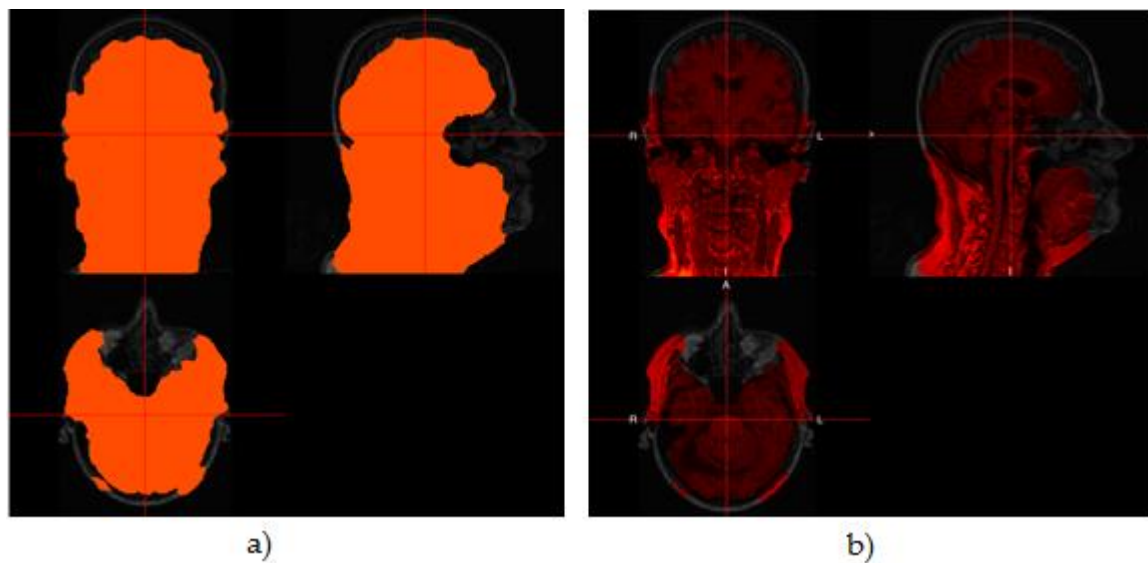
Kolejnym krokiem wstępnego przetwarzania danych jest ekstrakcja mózgu, czyli oddzielenie mózgu od czaszki. Początkowo jako narzędzie do ekstrakcji wybrano FSL-BET, który dawał dobre rezultaty dla danych typu flair, jednak nie przyniósł pożądaných rezultatów dla danych typu MPRAGE [71]. W wyniku procedury ekstrakcji za pomocą narzędzia FSL-BET dla danych typu MPRAGE otrzymano mózg z pozostałościami kości czaszki. Na poniższym rysunku po lewej stronie zaprezentowano oryginalny obraz MRI (a) a po prawej wynik ekstrakcji (b).



Rysunek 7 a) Obraz MRI przed ekstrakcją mózgu; b) Wynik ekstrakcji mózgu za pomocą FSL-BET.

Na rysunku (b) wyraźnie widoczne są pozostałości czaszki po ekstrakcji za pomocą FSL-BET.

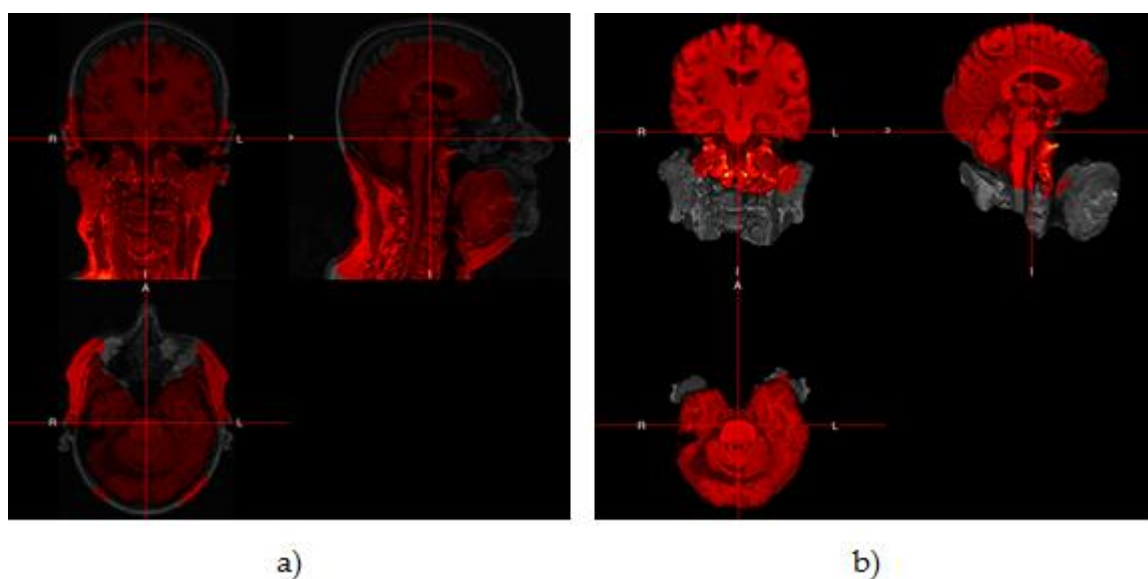
Na poniższych rysunkach przedstawiono wynik tej samej ekstrakcji ale w inny sposób, poprzez zacięniowanie fragmentu po ekstrakcji na tle oryginalnego obrazu MRI przed ekstrakcją.



Rysunek 8 Wynik ekstrakcji mózgu za pomocą narzędzia FSL-BET. a) Zacięniowane na pomarańczowo fragmenty prezentują wynik ekstrakcji mózgu na tle oryginalnego obrazu przed ekstrakcją. b) Zacięniowana na czerwono transparentne fragmenty prezentują wynik ekstrakcji mózgu na tle oryginalnego obrazu przed ekstrakcją,

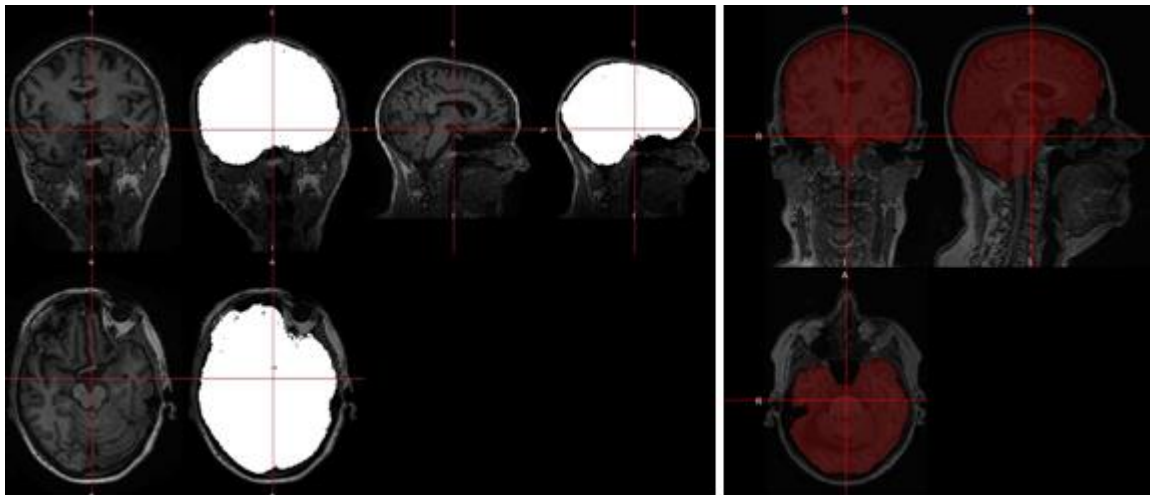
Można zauważyć, że niektóre rejony mają spore pozostałości czaszki, a w innych wycięto zbyt dużo (poza czaszką również fragmenty mózgu).

Początkowo próbowano zastosować procedurę ekstrakcji mózgu bez zastosowania korekty niejednorodności natężenia pola polaryzacji. W rezultacie otrzymano powyższe obrazy. Po zastosowaniu korekty niejednorodności natężenia pola wynik ekstrakcji uległ poprawie, jednak nadal nie był satysfakcjonujący. W kolejnym kroku przeprowadzono dodatkowo estymację środka ciężkości, co znacząco poprawiło rezultaty ekstrakcji mózgu. Nadal jednak obraz będący wynikiem ekstrakcji zawierał zbyt duży obszar. Poniższy rysunek prezentuje rezultat ekstrakcji mózgu bez korekty niejednorodności natężenia pola (a), oraz po korekcje niejednorodności (obszary w szarym kolorze na rysunku (b)) a także dodatkowo z zastosowaniem estymacji środka ciężkości (obszary w kolorze czerwonym na rysunku (b)).



Rysunek 9 Rezultaty ekstrakcji mózgu: a) bez korekty niejednorodności, b) po korekcje niejednorodności (kolor szary) oraz dodatkowo z zastosowaniem estymacji środka ciężkości (kolor czerwony).

Zastosowanie korekty niejednorodności natężenia pola oraz dodanie estymacji środka ciężkości poprawiło wyniki ekstrakcji mózgu, jednak nadal nie były one zadowalające. Poszukiwanie innej procedury ekstrakcji mózgu doprowadziło do znalezienia narzędzia SPM 12 działającego w środowisku MATLAB (<https://www.fil.ion.ucl.ac.uk/spm/>) [72]. Otrzymane rezultaty były bardzo dobre i ostatecznie to narzędzie wybrano do ekstrakcji mózgow z obrazów pochodzących z baz ADNI oraz ESD. Rezultaty ekstrakcji mózgu w wyniku zastosowania narzędzia SPM 12 (po uprzednim przeprowadzeniu korekty niejednorodności natężenia pola) zamieszczono na poniższych rysunkach.



Rysunek 10 Wyniki ekstrakcji mózgu za pomocą narzędzia SPM 12 (rezultat ekstrakcji zaznaczono na kolor biały oraz transparenty czerwony).

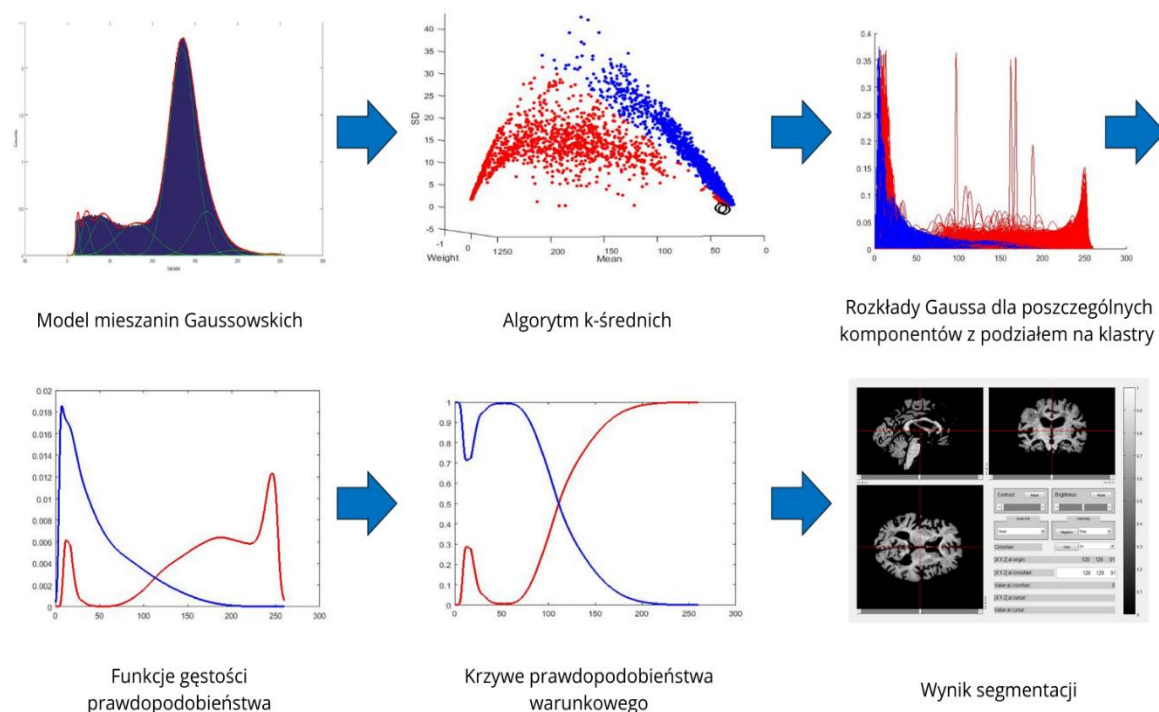
### 3.6. Segmentacja

Segmentacja mózgu na płyn mózgowo-rdzeniowy (ang. cerebrospinal fluid, CSF) oraz resztę mózgu (łącznie istotę szarą i białą) została przeprowadzona oddzielnie dla każdego pacjenta. W celu przeprowadzenia segmentacji dostosowano algorytm MiMSeg [73]. W oryginalnej wersji algorytm ten służył do detekcji tkanek nowotworowych na podstawie obrazów DWI. Algorytm ten dostosowano do naszej

sytuacji, w której analizujemy obrazy MR T1-zależne i chcemy oddzielić płyn mózgowo-rdzeniowy od mózgu.

Algorytm ten opiera się na modelu mieszanin gaussowskich i segmentuje płyn mózgowo-rdzeniowy od mózgu wyznaczając punkt odcięcia na skali szarości.

Poniższy rysunek przedstawia kolejne kroki segmentacji.



Rysunek 11 Schemat przedstawiający kolejne kroki segmentacji mózgu za pomocą dostosowanego algorytmu MiMSeg.

Cały proces segmentacji przeprowadzono oddzielnie dla każdego pacjenta. W pierwszym kroku dla każdej warstwy obrazu MRI analizowanego pacjenta wykonano, za pomocą mieszanin Gaussowskich, dekompozycję dystrybucji intensywności sygnału odpowiadających tym warstwom.

Dekompozycja polega na przedstawieniu widma pola rezonansu jądrowego za pomocą modelu mieszanin gaussowskich, gdzie „pik” jest reprezentowany przez pojedynczy komponent modelu. Model konstruowany jest przy założeniu,

że możliwe jest znalezienie takiej liczby komponentów  $n_s$ , aby ich suma była przybliżeniem oryginalnego sygnału. Funkcja gęstości prawdopodobieństwa dla  $s$ -tej warstwy obrazu  $X$  wyraża się przez sumę  $n_s$  ważonych (z wagą  $w_{i,s}$  dla  $i$ -tego komponentu) funkcji gęstości prawdopodobieństwa rozkładu normalnego o średniej  $\mu_{i,s}$  i odchyleniu standardowym  $\sigma_{i,s}$  odpowiadających poszczególnym komponentom gaussowskim:

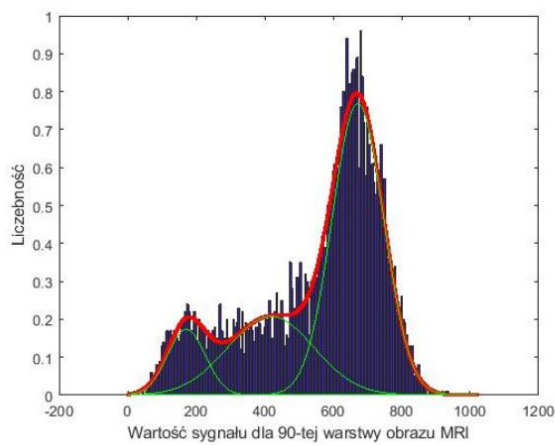
$$f_s(X) = \sum_{i=1}^{n_s} w_{i,s} \cdot \varphi(X; \mu_{i,s}, \sigma_{i,s}^2)$$

Na określenie odpowiedniej liczby komponentów modelu pozwala bayesowskie kryterium informacyjne (BIC), które wyraża się wzorem:

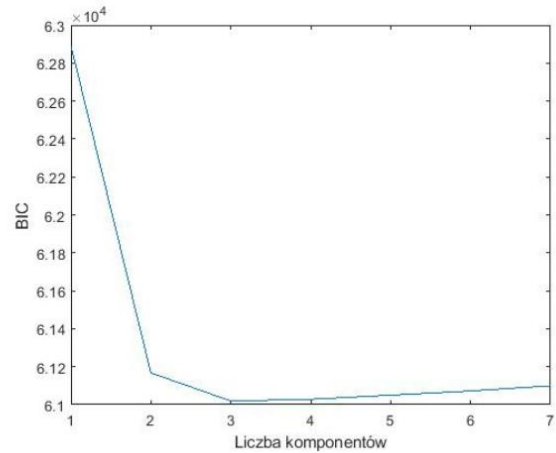
$$BIC_s = -2 \ln(L_{s,n_s}) + (3n_s - 1) \ln(k_s),$$

gdzie  $L_{s,n_s}$  oznacza logarytm funkcji wiarygodności dla  $s$ -tej warstwy obrazu  $X$  i  $n_s$ -tej składowej modelu mieszanin gaussowskich, a  $k_s$  oznacza liczbę pikseli w analizowanej  $s$ -tej warstwie obrazu  $X$  [73]. Do estymacji parametrów modelu mieszanin gaussowskich zastosowano algorytm maksymalizacji wartości oczekiwanej (ang. Expectation-Maximization algorithm, EM) [74]. Na poniższym rysunku przedstawiono wyniki dekompozycji za pomocą modelu mieszanin gaussowskich dla trzech warstw jednego obrazu MRI mózgu ze zbioru ADNI: 90-tej, 100-tej i 110-tej. Na zielono wrysowano w histogramy poszczególne komponenty. Ponadto zamieszczono wykresy przedstawiające zależność pomiędzy liczbą komponentów a bayesowskim kryterium informacyjnym dla analizowanych warstw. Można zauważyć, że dla każdej z trzech warstw mamy inną liczbę komponentów: dla warstwy 90-tej mamy 3 komponenty, dla warstwy 100-tej – 4 komponenty a dla warstwy 110-tej mamy 5 komponentów.

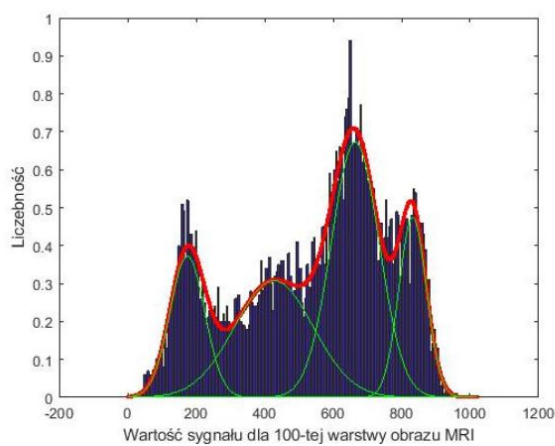




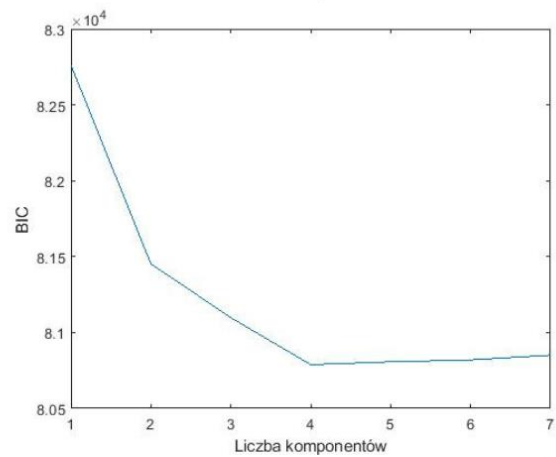
a)



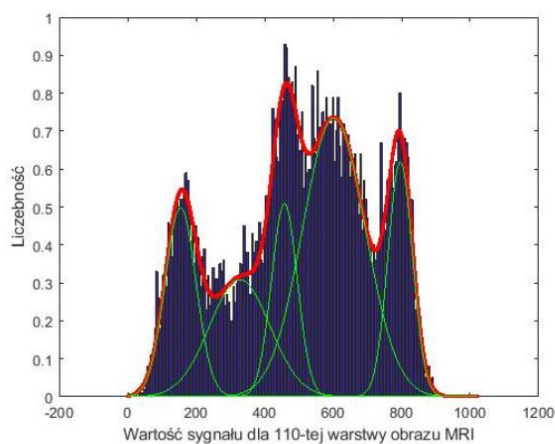
b)



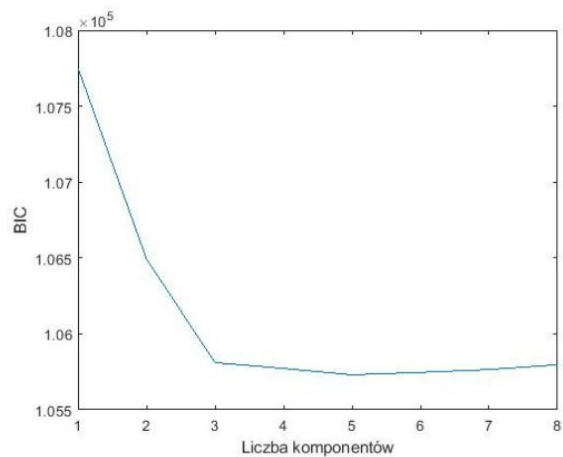
c)



d)



e)



f)

Rysunek 12 Wynik dekompozycji za pomocą modelu mieszanin gaussowskich: a) model mieszanin gaussowskich z trzema komponentami dla 90-tej warstwy obrazu MRI wpisany w histogram, b) zależność pomiędzy BIC a liczbą komponentów dla 90-tej warstwy obrazu MRI, c) model mieszanin gaussowskich z czterema komponentami dla 100-tej warstwy obrazu MRI, d) zależność pomiędzy BIC a liczbą komponentów dla 100-tej warstwy obrazu MRI, e) model mieszanin gaussowskich z pięcioma komponentami dla 110-tej warstwy obrazu MRI, f) zależność pomiędzy BIC a liczbą komponentów dla 110-tej warstwy obrazu MRI.

Dla każdej warstwy pojedynczego obrazu MRI w wyniku dekompozycji otrzymano zbiór wartości parametrów składowych modelu mieszanin gaussowskich: średniej, odchylenia standardowego oraz wagi. Wartości tych parametrów zebrano ze wszystkich warstw i umieszczono w jednej macierzy, gdzie w kolumnach umieszczono wartości parametrów, a wiersze odpowiadały poszczególnym komponentom.

Idea dekompozycji rozkładu sygnału polega na tym, iż poszczególne komponenty, lub grupy komponentów odpowiadają poszczególnym tkankom mózgu. Najczęściej spotykanym podziałem jest podział mózgu na płyn mózgowo-rdzeniowy, istotę szarą oraz istotę białą. Specyfika obrazów  $T_1$  – zależnych, w których płyn mózgowo-rdzeniowy jest ciemniejszy niż tkanka mózgowa pozwala na oddzielenie płynu CSF od tkanki za pomocą metod klasteryzacji.

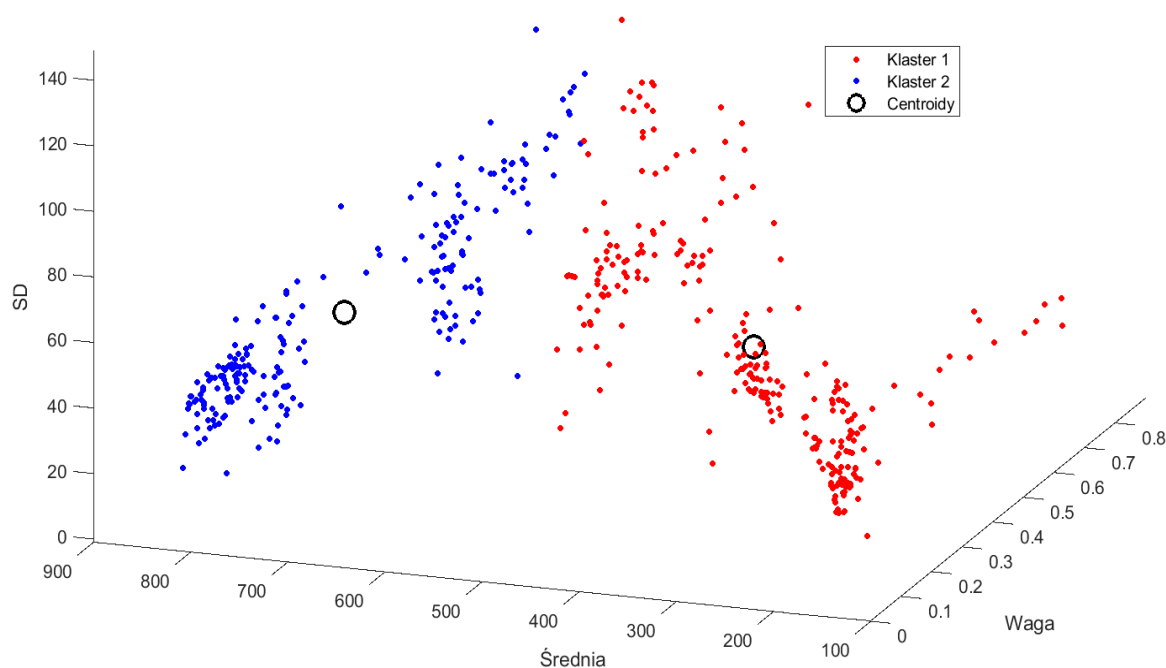
W celu znalezienia punktu odcięcia, który będzie segmentował mózg na płyn mózgowo-rdzeniowy (CSF) oraz tkankę mózgu zastosowano metodę klasteryzacji  $k$ -średnich do klasyfikacji podobnych wartości parametrów składowych GMM, zaproponowaną w algorytmie MiMSeg do znalezienia punktu odcięcia dla tkanki nowotworowej i zdrowej tkanki mózgu [73].

Algorytm  $k$ -średnich jest nienadzorowanym algorytmem uczenia, który pozwala na klasyfikację danych, w tym przypadku punktów w trójwymiarowej przestrzeni wyznaczonej przez wartości średniej, odchylenia standardowego oraz wagi, na ustaloną a priori liczbę klastrów ( $k$ ) [75]. Kolejne kroki algorytm przedstawiają się następująco:

- w pierwszym kroku wybranych zostaje  $k$  punktów spośród analizowanych danych, które będą określały początkowe centroidy (środki) dla klastrów;

- następnie następuje podział punktów na klastry za pomocą funkcji określającej odległość w przyjętej metryce między centroidem a poszczególnymi punktami;
- obliczona zostaje nowa wartość centroidu jako średnia wartości należących do danego klastra;
- procedurę powtarza się aż do momentu, gdy odległość (w ustalonej metryce) nie zmienia się znacząco w porównaniu do poprzedniej iteracji.

W algorytmie k-średnich zastosowanym do klasyfikacji wartości parametrów składowych GMM zastosowano metrykę euklidesową. Na poniższym rysunku przedstawiono wynik klasyfikacji na dwa klastry. Kolorem czerwonym oznaczone zostały punkty należące do pierwszego klastra, a kolorem niebieskim – do drugiego. Jeden klaster odpowiada płynowi mózgowo-rdzeniowemu CSF na obrazie MRI analizowanego pacjenta, a drugi odpowiada tkance mózgu.

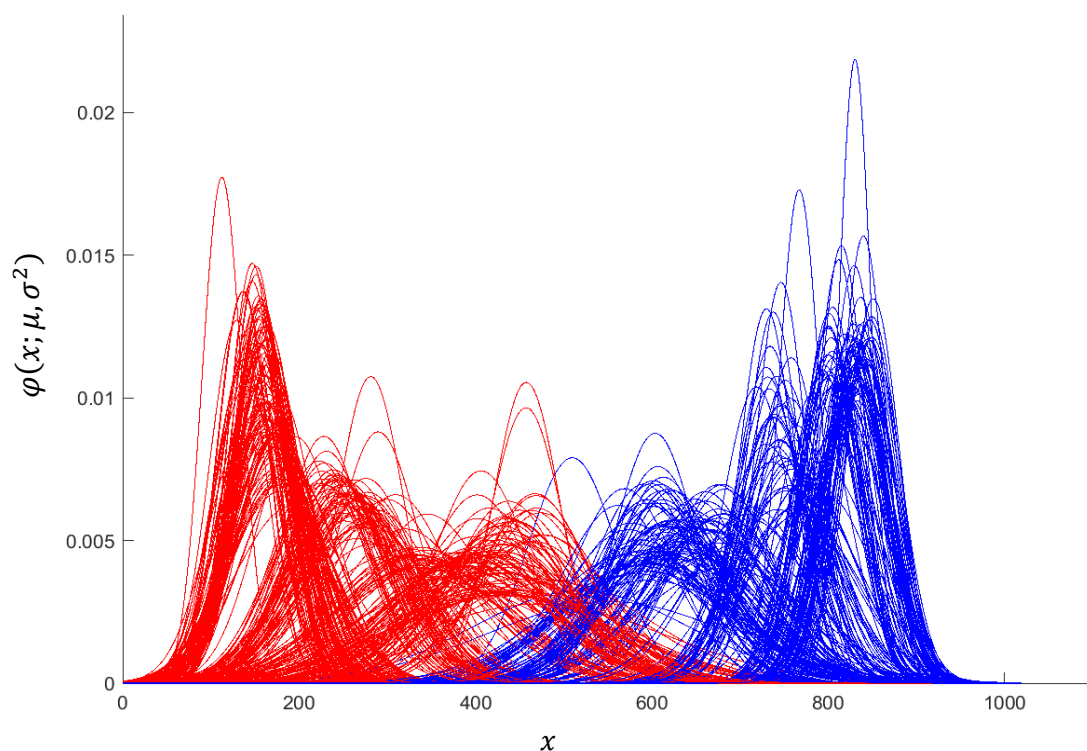


Rysunek 13 Rezultat klasyfikacji wartości parametrów składowych GMM za pomocą algorytmu k-średnich.

Punkty zakwalifikowane do poszczególnych klastrów charakteryzują odpowiednie komponenty modelu mieszanin gaussowskich. Funkcja gęstości prawdopodobieństwa rozkładu normalnego dla  $j$ -tego komponentu należącego do klastra  $c$  ( $c = 1, 2$ ) opisana jest wzorem:

$$\varphi(X; \mu_{j,c}, \sigma_{j,c}^2) = \frac{1}{\sigma_{j,c}\sqrt{2\pi}} \cdot e^{-\frac{(X-\mu_{j,c})^2}{2\sigma_{j,c}^2}}$$

gdzie  $\mu_{j,c}$  i  $\sigma_{j,c}$  są odpowiednio wartościami średniej i odchylenia standardowego charakteryzującymi  $j$ -ty komponent klastra  $c$ . Na poniższym rysunku przedstawiono funkcje gęstości prawdopodobieństwa komponentów GMM należących dla klastra 1 (kolor czerwony) oraz klastra 2 (kolor niebieski).

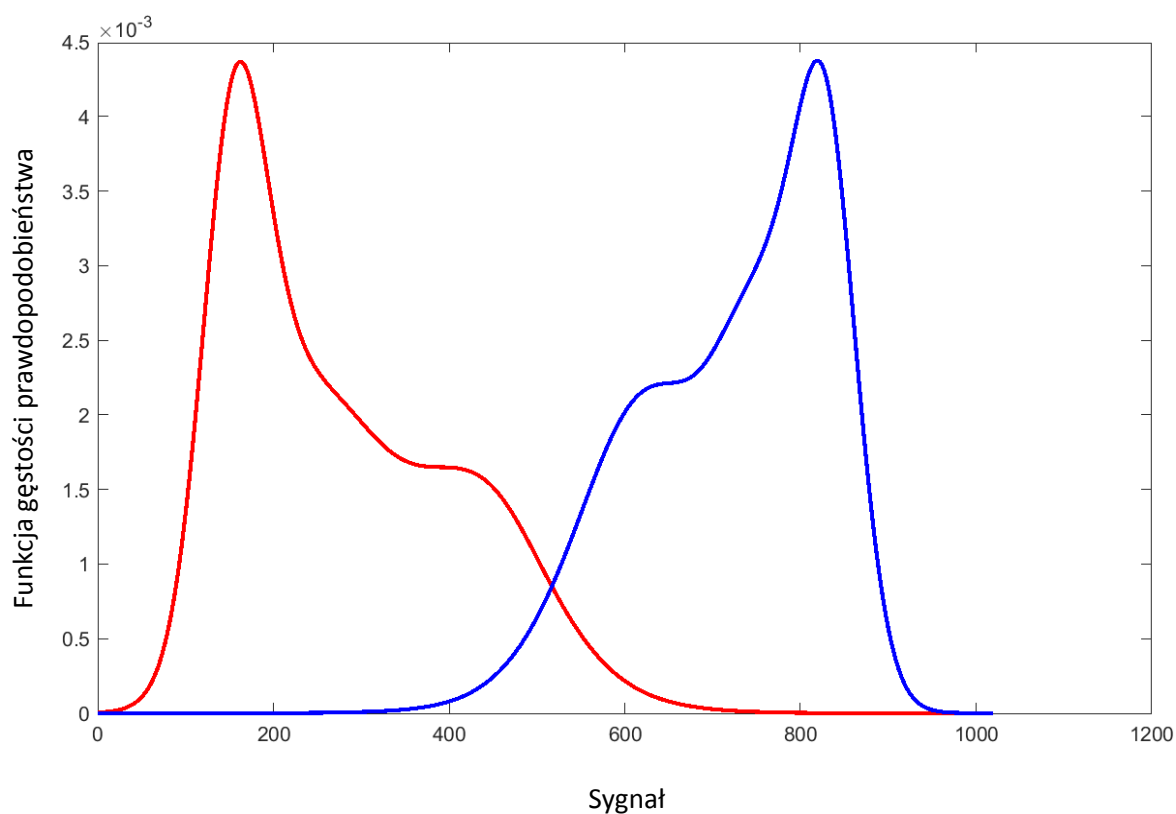


Rysunek 14 Funkcje gęstości prawdopodobieństwa komponentów GMM z podziałem na dwa klastry.

Suma (z odpowiednimi wagami) funkcji gęstości prawdopodobieństwa komponentów modelu GMM należących do klastra  $c$  ( $c = 1, 2$ ) określa funkcję gęstości prawdopodobieństwa  $c$ -tego klastra obrazu  $X$ , co wyraża się wzorem:

$$\tilde{f}_c(X) = \sum_{h=1}^{n_c} w_{h,c} \cdot \varphi(X; \mu_{h,c}, \sigma_{h,c}^2) = \sum_{h=1}^{n_c} w_{h,c} \cdot \frac{1}{\sigma_{h,c}\sqrt{2}} \cdot e^{-\frac{(X-\mu_{h,c})^2}{2\sigma_{h,c}^2}}$$

gdzie  $w_{h,c}$  jest wagą  $h$ -tego komponentu GMM w klastrze  $c$ ,  $\varphi(X; \mu_{h,c}, \sigma_{h,c}^2)$  jest funkcją gęstości prawdopodobieństwa rozkładu normalnego dla  $h$ -tego komponentu GMM w klastrze  $c$ , a  $\mu_{h,c}$  oraz  $\sigma_{h,c}$  są wartościami średniej i odchylenia standardowego opisującymi  $h$ -ty komponent klastra  $c$  ( $c = 1, 2$ ). Na poniższym wykresie przedstawiono funkcje gęstości prawdopodobieństwa pierwszego i drugiego klastra obrazu  $X$ .



Rysunek 15 Funkcje gęstości prawdopodobieństwa dla pierwszego i drugiego klastra obrazu  $X$ .

Funkcja gęstości prawdopodobieństwa dla sygnału obrazu  $X$  (łącznie dla obu klastrów) jest sumą funkcji gęstości prawdopodobieństwa klastra pierwszego i drugiego:

$$f(X) = \sum_{t=1}^2 \tilde{f}_t(X) = \sum_{t=1}^2 \sum_{h=1}^{n_t} w_{h,t} \cdot \varphi(X; \mu_{h,t}, \sigma_{h,t}^2) = \sum_{t=1}^2 \sum_{h=1}^{n_t} w_{h,t} \cdot \frac{1}{\sigma_{h,t} \sqrt{2}} \cdot e^{-\frac{(X-\mu_{h,t})^2}{2\sigma_{h,t}^2}}$$

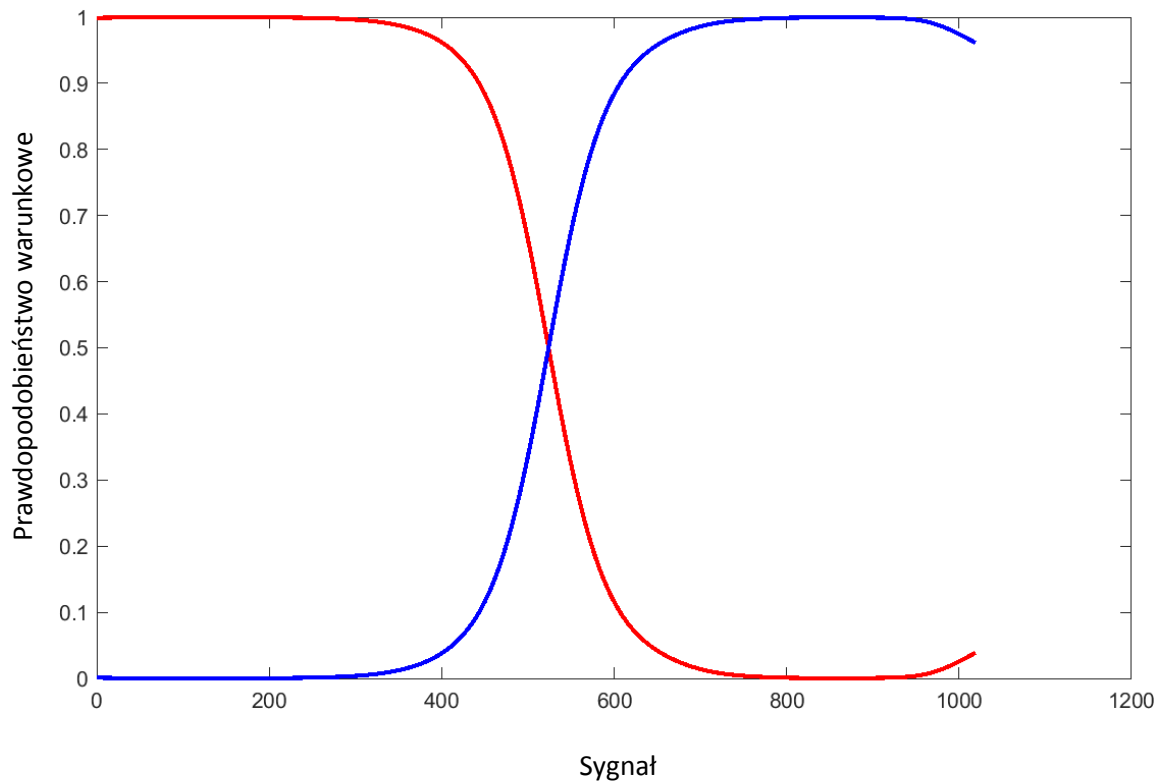
gdzie  $\tilde{f}_t(X)$  jest funkcją gęstości prawdopodobieństwa dla klastra  $t$ ,  $w_{h,t}$  jest wagą  $h$ -tego komponentu GMM w klastrze  $t$ , a  $\varphi(X; \mu_{h,t}, \sigma_{h,t}^2)$  jest funkcją gęstości prawdopodobieństwa rozkładu normalnego dla  $h$ -tego komponentu GMM w klastrze  $t$  o średniej  $\mu_{h,t}$  i odchyleniu standardowym  $\sigma_{h,t}$ .

Mając powyższe funkcje gęstości można estymować punkt odcięcia dla segmentacji płynu CSF i tkanki mózgu u analizowanego pacjenta korzystając z kryterium maksymalnej wartości prawdopodobieństwa warunkowego [73].

Prawdopodobieństwo warunkowe przynależności do klastra  $t$  dla wartości sygnału obrazu  $X$  równej  $x$  wyraża się wzorem:

$$P(t|X = x) = \frac{\tilde{f}_t(x)}{\tilde{f}_1(x) + \tilde{f}_2(x)} = \frac{\tilde{f}_t(x)}{f(x)}$$

Poniższy wykres przedstawia krzywe prawdopodobieństwa warunkowego z podziałem na klastry.



Rysunek 16 Krzywe prawdopodobieństwa warunkowego z podziałem na klastry.

Przypisanie wartości  $x$  do odpowiedniego klastra następuje poprzez wyznaczenie wartości maksymalnej prawdopodobieństw warunkowych. Wartość  $x$  zostaje przypisana do tego klastra, dla którego prawdopodobieństwo warunkowe  $P(t|X = x)$  przyjmuje największą wartość dla  $t = 1,2$ . Zatem punkt odcięcia  $p_{od}$  definiuje się jako punkt, dla którego prawdopodobieństwa warunkowe dla obu klastrów są sobie równe:

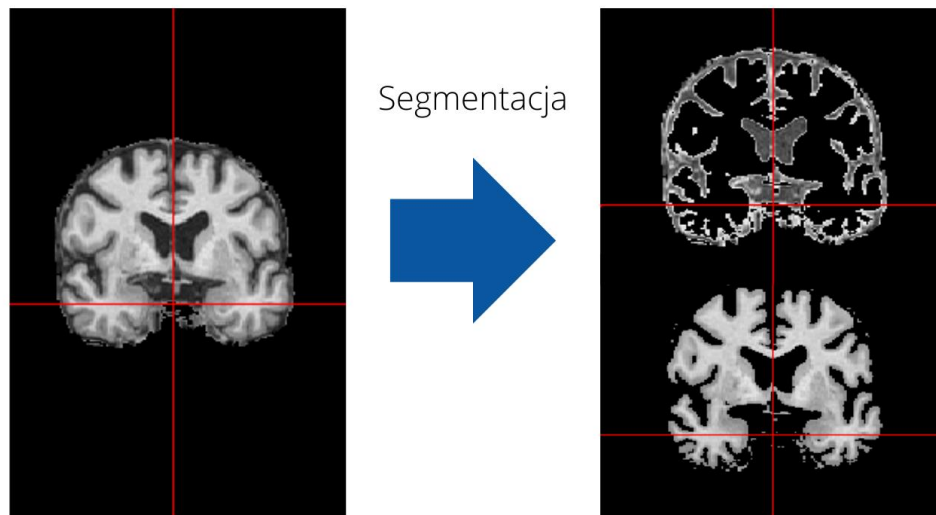
$$\tilde{f}_1(p_{od}) = \tilde{f}_2(p_{od})$$

oraz zachodzą nierówności:

$$\tilde{f}_2(x|x > p_{od}) < \tilde{f}_1(x|x > p_{od})$$

$$\tilde{f}_2(x|x < p_{od}) > \tilde{f}_1(x|x < p_{od})$$

Dla analizowanego pacjenta wartość punktu odcięcia wynosi  $p_{od} = 523,9544$ . Na poniższym rysunku przedstawiono wynik segmentacji obrazu MRI na płyn mózgowo-rdzeniowy oraz tkankę mózgu za pomocą dostosowanego algorytmu MiMSeg opisanego szczegółowo w tym rozdziale.



Rysunek 17 Rezultat segmentacji obrazu MRI na płyn mózgowo-rdzeniowy oraz tkankę mózgu za pomocą dostosowanego algorytmu MiMSeg.

Powyższą procedurę segmentacji przeprowadzono oddzielnie dla każdego pacjenta ze zbiorów ADNI oraz EDSD wyznaczając indywidualny dla każdego analizowanego obrazu MRI punkt odcięcia.



## **4. Analiza danych**

### **4.1. Zmienne**

W pracy zaproponowano trzy biomarkery oparte na danych z obrazowania rezonansem magnetycznym, zatem trzy zmienne zostały zdefiniowane dla danych pochodzących z obrazów MRI. Zmienne te w sposób numeryczny opisują zmiany w strukturze mózgu. Pierwsza zmienna to względna objętość mózgu (bez płynu mózgowo-rdzeniowego), druga to współczynnik pofałdowania, a trzecia to względna objętość komór bocznych mózgu. W kolejnych podrozdziałach omówione zostały dokładniej te trzy zmienne.

#### **4.1.1. Względna objętość mózgu**

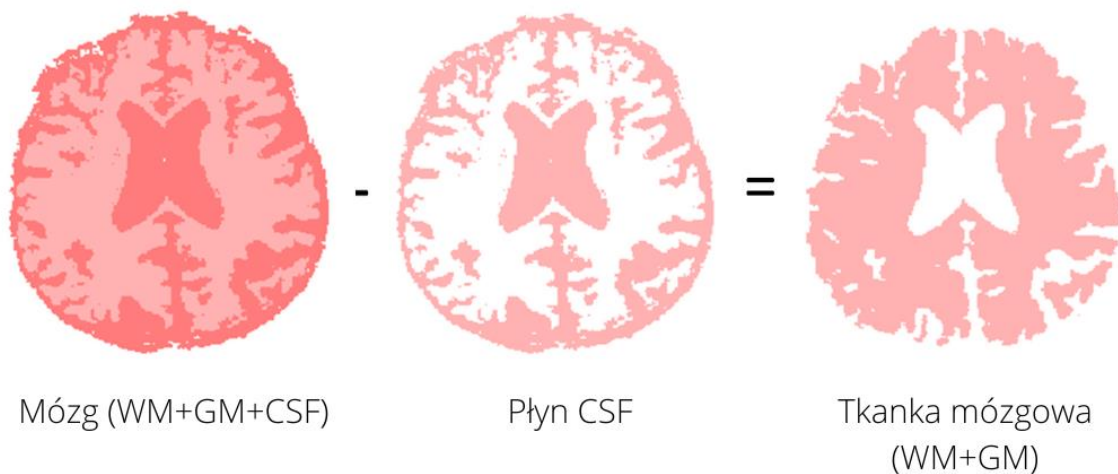
U pacjentów z chorobą Alzheimera mózg kurczy się, a wolne przestrzenie zostają wypełnione płynem mózgowo-rdzeniowym [1] [2]. Zaproponowana zmienna względna objętość mózgu (bez płynu mózgowo-rdzeniowego) ma opisywać tę sytuację. Spodziewamy się, że skoro mózg kurczy się, to objętość mózgu bez płynu mózgowo-rdzeniowego u chorych z chorobą Alzheimera będzie mniejsza niż u zdrowych, starszych osób, a także mniejsza niż u pacjentów z łagodnym upośledzeniem funkcji poznawczych. Ponieważ objętość mózgu zależy również od czynników osobniczych i może różnić się znacznie nawet pomiędzy dwiema zdrowymi osobami, to zaproponowano zmienną, która jest względną objętością mózgu bez płynu mózgowo-rdzeniowego, czyli oceniamy objętość mózgu po oddzieleniu płynu CSF w odniesieniu do objętości mózgu wraz z płynem.

#### 4.1.1.1. Względna objętość mózgu - metoda wyznaczania

Względna objętość mózgu (ang. relative brain volume, RBV) została zdefiniowana jako objętość mózgu bez płynu mózgowo-rdzeniowego CSF ( $V_{\text{CSF}}$ ) podzielona przez objętość całego mózgu ( $V$ ) (wraz z płynem CSF) i wyrażona w procentach. Objętość całego mózgu ( $V$ ) została uzyskana poprzez zliczenie wokseli (woksel to najmniejszy element objętości; jest to trójwymiarowy odpowiednik piksela) należących do mózgu, a objętość mózgu bez płynu mózgowo-rdzeniowego CSF ( $V_{\text{CSF}}$ ) została uzyskana poprzez zliczenie wokseli obrazu mózgu po odjęciu płynu mózgowo-rdzeniowego.

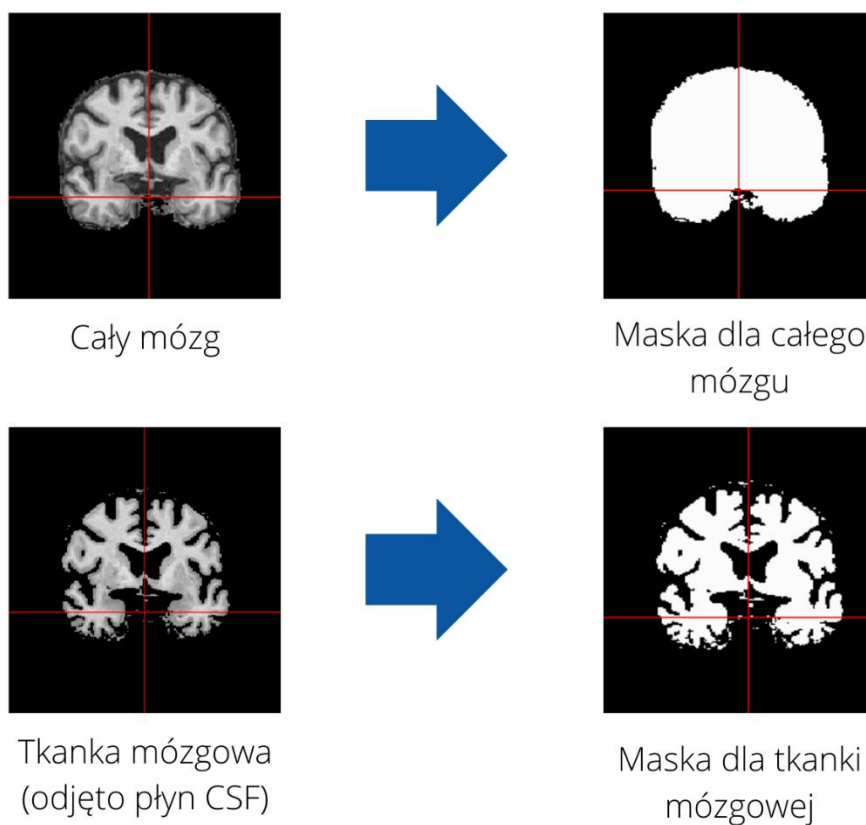
$$\text{RBV} = V_{\text{CSF}} / V \cdot 100\%$$

Na poniższym rysunku przedstawiono, w sposób schematyczny, wynik segmentacji (opisanej w rozdziale 3.5) oraz oddzielenia płynu mózgowo-rdzeniowego od mózgu otrzymując obraz tkanki mózgowej (czyli istoty białej (ang. white matter, WM) oraz szarej (ang. grey matter, GM) łącznie).



Rysunek 18 Schemat uzyskania obrazu tkanki mózgu poprzez segmentację i oddzielenie płynu mózgowo-rdzeniowego od mózgu.

Dla obrazu tkanki mózgu utworzona zostaje maska, czyli obraz binarny i zliczone zostają woksele należące do tej maski. W analogiczny sposób zlicza się woksele należące do całego mózgu poprzez utworzenie maski dla całego mózgu. Na rysunku poniżej przedstawiono oryginalny obraz MRI całego mózgu i tkanki mózgu (po oddzieleniu płynu CSF) oraz utworzone dla nich maski binarne.



Rysunek 19 Rezultat tworzenia maski dla całego mózgu oraz dla tkanki mózgowej uzyskanej w wyniku segmentacji po oddzieleniu płynu mózgowo-rdzeniowego.

#### 4.1.1.2. Względna objętość mózgu - wyniki

U osób z chorobą Alzheimera mózg kurczy się a wolne przestrzenie wypełnia płyn mózgowo-rdzeniowy. Wyniki zamieszczone w poniższej tabeli potwierdzają ten fakt.

Tabela 2 Statystyka opisowa zmiennej względna objętość mózgu (RBV) dla zbioru danych ADNI.

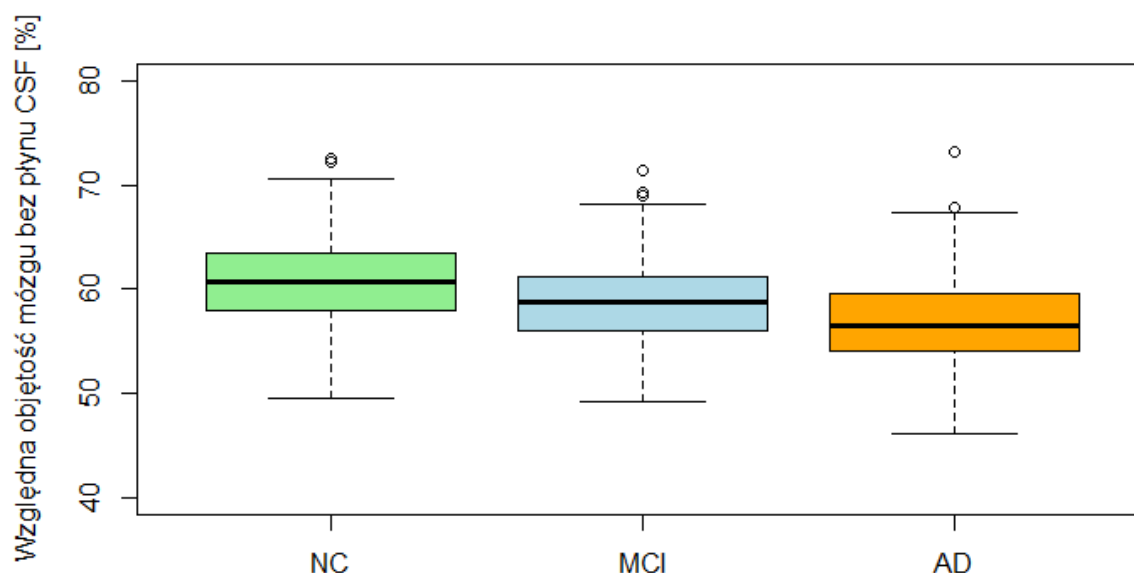
RBV [%]	Średnia	SD	IQR	Min	Q1	Mediana	Q3	Max	N
NC	60,78	4,006	5,53	49,59	57,91	60,73	63,44	72,62	194
MCI	58,79	3,896	5,06	49,18	56,07	58,72	61,13	71,38	311
AD	57,08	4,283	5,44	46,20	54,14	56,42	59,59	73,24	133

Przeprowadzono testowanie normalności rozkładu w grupach NC, MCI i AD, i w każdej z grup nie było podstaw do odrzucenia hipotezy o normalności rozkładu zmiennej RBV, zatem do porównania grup zastosowano test parametryczny ANOVA, którego wynik wskazuje, że istnieje istotna statystycznie różnica między grupami dla zmiennej RBV ( $p < 0,0001$ ). Następnie zastosowano test HSD Tukeya dla porównań parami.

Tabela 3 Wyniki testu post-hoc HSD Tukeya dla porównań parami.

Porównanie	Różnica	Dolna granica 95% przedziału ufności	Górna granica 95% przedziału ufności	Wartość p
MCI-AD	1,72	0,74	2,69	0,0001241
NC-AD	3,70	2,64	4,76	<0,0000001
NC-MCI	1,98	1,12	2,85	0,0000003

Wyniki testu Tukeya pokazują (powyższa tabela), że różnice dla zmiennej RBV pomiędzy każdą parą grup są istotne statystycznie. Zatem względna objętość mózgu jest mniejsza u pacjentów z chorobą Alzheimera w porównaniu z osobami zdrowymi a nawet chorymi z łagodnym upośledzeniem funkcji poznawczych. Również względna objętość mózgu u pacjentów z łagodnym upośledzeniem funkcji poznawczych jest mniejsza niż u osób zdrowych. Najmniejszą średnią względną objętość mózgu mamy dla osób z chorobą Alzheimera (57,08%, SD=4,283), a największą dla osób zdrowych (60,78%, SD=4,006). Graficzne przedstawienie wyników zamieszczono na poniższym wykresie.



Rysunek 20 Względna objętość mózgu bez płynu mózgowo-rdzeniowego dla grup: NC, MCI i AD. Wykres typu ramka-wąsy.

#### 4.1.2. Współczynnik pofałdowania

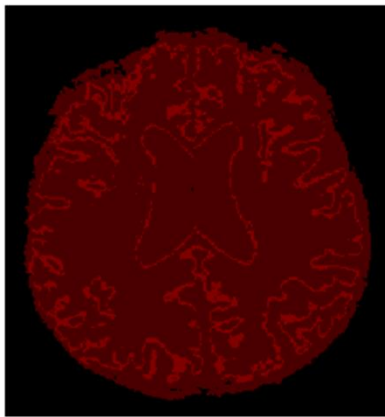
Wiadomo, że mózg u osób cierpiących na chorobę Alzheimera ulega atrofii, czyli kurczy się, a to powoduje zwiększenie pofałdowania powierzchni mózgu. Druga zaproponowana zmienna ma opisywać właśnie pofałdowanie mózgu. Spodziewamy się, że skoro mózg staje się bardziej pofałdowany to powierzchnia mózgu musi się zwiększać. Ponieważ podobnie jak objętość mózgu, tak również powierzchnia zależą od czynników osobniczych, to zaproponowano parametr względny czyli powierzchnię mózgu bez płynu mózgowo-rdzeniowego w odniesieniu do objętości mózgu bez płynu CSF.

##### 4.1.2.1. Współczynnik pofałdowania - metoda wyznaczania

Zmienna współczynnik pofałdowania (ang. shrinkage factor, SF) została zdefiniowana jako iloraz powierzchni mózgu bez płynu mózgowo-rdzeniowego CSF ( $S_{CSF}$ ) i objętości mózgu bez płynu ( $V_{CSF}$ ) oraz wyrażona w procentach:

$$SF = S_{\text{-CSF}} / V_{\text{-CSF}} \cdot 100\%$$

Powierzchnia mózgu bez płynu mózgowo-rdzeniowego to liczba wokseli na powierzchni mózgu, która wyznaczona została w taki sposób, że dla każdej warstwy mózgu z usuniętym płynem mózgowo-rdzeniowym wyznaczono kontur i zliczono woksele tworzące kontury wszystkich warstw. Objętość mózgu bez płynu CSF, czyli tkanki mózgowej obliczono zliczając dla każdej warstwy danego obrazu MRI woksele należące do tkanki mózgowej.



Kontur tkanki mózgowej  
- linia jasnoczerwona



Mózg po oddzieleniu płynu CSF  
- czerwony obszar

Rysunek 21 Warstwa obrazu MRI przedstawiająca kontur tkanki mózgowej (linia jasnoczerwono po lewej stronie) oraz tkankę mózgową (kolor czerwony po prawej stronie).

Do wyznaczenia konturu stworzono algorytm, którego poszczególne kroki przedstawiono poniżej.

Algorytm wyznaczania konturu tkanki mózgowej (czyli tkanki pozostałej po usunięciu z mózgu płynu mózgowo-rdzeniowego) dla analizowanej warstwy obrazu:

1. W pierwszym kroku utworzono maskę binarną tkanki mózgowej, gdzie „1” w macierzy reprezentującej warstwę oznacza, że dany obszar należy do tkanki, a „0” oznacza tło obrazu.



Rysunek 22 Maska binarna tkanki mózgu dla analizowanej warstwy.

2. W kolejnym kroku przechodzimy komórka po komórce przez macierz reprezentującą analizowaną warstwę z maską. Jeśli natrafimy na tkankę, czyli komórkę zawierającą „1”, to sprawdzamy czy komórki wokół zawierają „1” czy „0”. Jeśli komórki otaczające badaną komórkę nie wszystkie zawierają „1”, to oznacza, że to jest kontur i wtedy takiej komórce przypisujemy w nowej macierzy „1”.

Warstwa W	l-1	l	l+1
j-1			
j			
j+1			

Rysunek 23 Fragment macierzy reprezentującej analizowaną warstwę. Na granatowo zaznaczono komórkę reprezentującą element co, do którego podejmowana jest decyzja, czy należy do konturu tkanki mózgowej, a na niebiesko zaznaczono komórki sąsiadujące, których suma wartości daje odpowiedź na pytanie, czy środkowy element należy do konturu czy nie.

Jeśli  $W(j, l) = 1$  i  $W(j - 1, l) + W(j, l - 1) + W(j, l + 1) + W(j + 1, l) < 4$ , to  $\tilde{W}(j, l) = 1$ , gdzie  $W$  jest macierzą reprezentującą analizowaną warstwę z maską, a  $\tilde{W}$  jest nową macierzą, w której elementy konturu będą oznaczone jako „1”, a wszystko, co nie jest konturem jako „0”.

3. Jeśli warunek z kroku 2 nie był spełniony, to komórkę oznaczano jako „0”.

Poniższy rysunek przedstawia kontur powstały w wyniku przeprowadzenia opisywanej procedury.



Rysunek 24 Kontur tkanki mózgu dla analizowanej warstwy obrazu MRI.

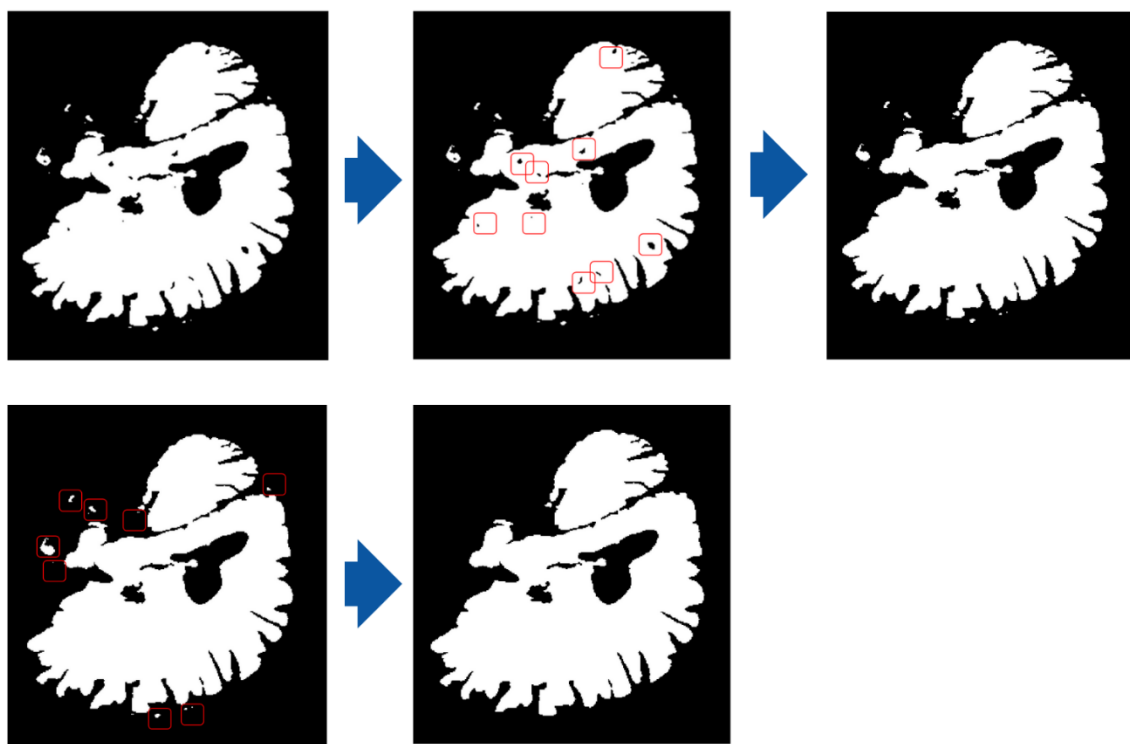
Powyższy algorytm wyznaczania konturu dla pojedynczej warstwy przeprowadzono dla każdej warstwy analizowanego obrazu MRI uzyskując po zsumowaniu wyników ze wszystkich warstw powierzchnię dla całej tkanki w obrazie MRI.

Ponieważ część obrazów po wyznaczeniu konturu zawierała wiele drobnych konturów (nazywanych dalej „czarnymi dziurami”, jeśli znajdowały się wewnątrz



obszaru tkanki lub „białymi plamami”, jeśli znajdowały się poza obszarem tkanki), to przygotowano algorytm usuwający „czarne dziury” oraz „białe plamy”.

Poniższy rysunek przedstawia kolejne kroki usuwania drobnych konturów: detekcję „czarnych dziur”, usuwanie „czarnych dziur”, detekcję „białych plam”, usuwanie „białych plam”.

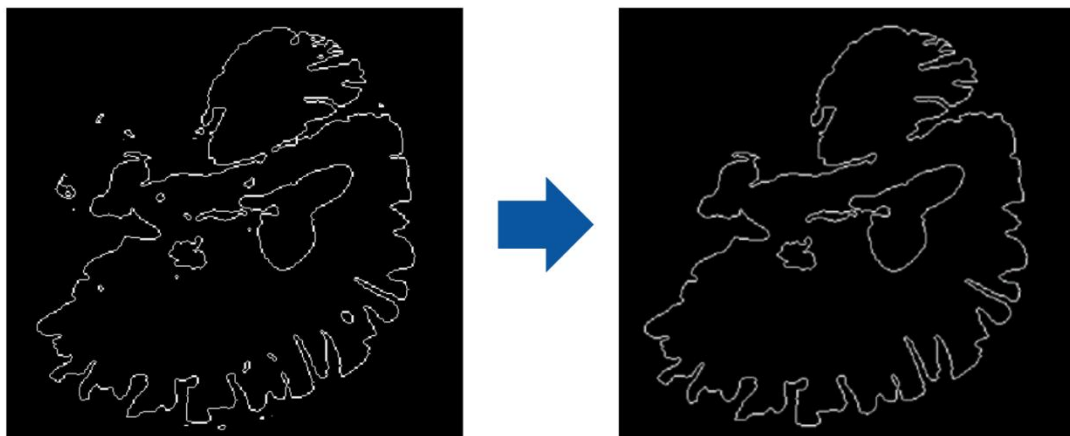


Rysunek 25 Kolejne kroki procedury usuwania drobnych konturów z obrazu MRI.

Algorytm usuwania „czarnych dziur” działa w następujący sposób: na analizowanej warstwie z maską tkanki mózgowej wyszukiwano „czarne dziury” w ten sposób, że sprawdzano czy komórki oznaczone jako „0” (kolor czarny) są otoczone przez tkankę (kolor biały) i czy „dziury” mieszczą się w ramce o wielkości co najwyżej 10x10 wokseli. Jeśli powyższe warunki były spełnione to w macierzy zmieniano „0” na „1”, czyli zaliczano dany obszar do tkanki. W analogiczny sposób znajdowano i usuwano „białe plamy”, które znajdowały się na czarnym tle. Procedury

te powtarzano dla każdej warstwy obrazu MRI. Powyższy rozmiar ramki 10x10 wokseli przyjęto po szeregu analiz wielkości „dziur” oraz „plam”. Przeprowadzono również analizę wrażliwości wielkości ramki w odniesieniu do rozmiaru obrazu MRI. Ponieważ obrazy nieznacznie różnią się rozdzielczością, to przyjęcie stałej wielkości ramki zamiast proporcjonalnej do wielkości obrazu nie wpływa w istotny sposób na wynik analizy.

Po usunięciu „czarnych dziur” i „białych plam” wyznaczano ponownie kontur dla każdej warstwy oddzielnie. Poniższy rysunek przedstawia rezultat wyznaczenia konturu po zastosowaniu algorytmu do usuwania „dziur” oraz „plam”. Po lewej znajduje się obraz z drobnymi konturami, a po prawej obraz po zastosowaniu procedury do usuwania tych konturów. Zarówno drobne kontury na zewnątrz jak i wewnątrz tkanki mózgu zostały usunięte.



Rysunek 26 Rezultat procedury usuwania drobnych konturów oraz procedury wyznaczenia konturu tkanki mózgu.

Po usunięciu dziur zliczono woksele należące do konturu i zsumowano liczbę wokseli z konturów dla wszystkich warstw obrazu MRI otrzymując w ten sposób liczbę wokseli dla powierzchni mózgu bez płynu mózgowo-rdzeniowego.

Następnie obliczano wartość wskaźnika pofałdowania dzieląc liczbę wokseli dla powierzchni mózgu bez płynu mózgowo-rdzeniowego przez liczbę wokseli opisującą objętość mózgu bez płynu mózgowo-rdzeniowego. Wynik przedstawiano w postaci procentowej.

#### 4.1.2.2. Współczynnik pofałdowania - wyniki

W wyniku rozwoju choroby Alzheimera mózgu ulega mocniejszemu pofałdowaniu. Znajduje to odzwierciedlenie w wynikach zamieszczonych w poniższej tabeli.

Tabela 4 Statystyka opisowa dla współczynnika pofałdowania SF dla danych ADNI.

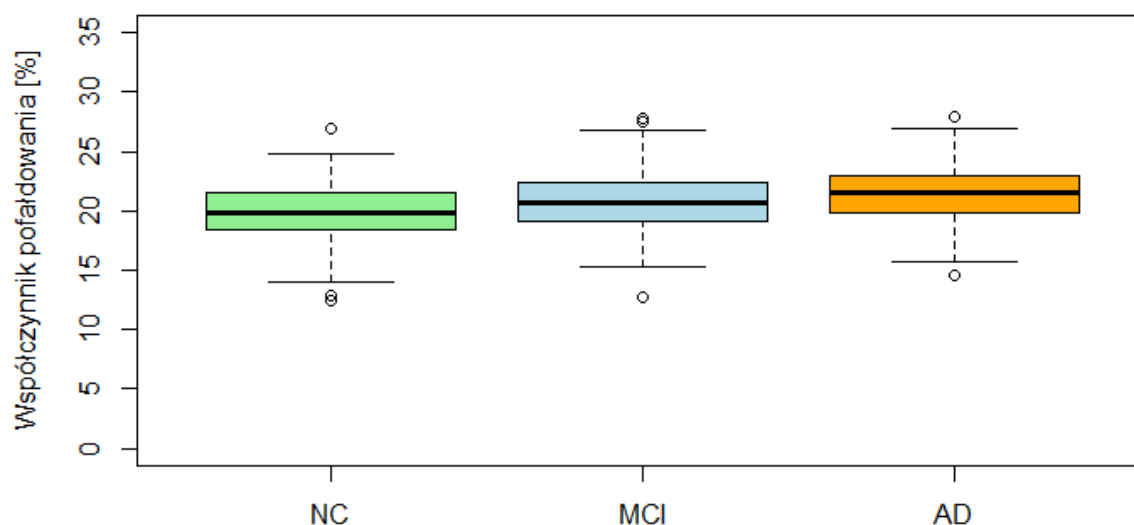
SF [%]	Średnia	SD	IQR	Min	Q1	Mediana	Q3	Max	N
NC	19,88	2,282	3,13	12,39	18,40	19,87	21,53	26,97	194
MCI	20,78	2,367	3,26	12,80	19,08	20,68	22,35	27,78	311
AD	21,43	2,505	3,19	14,55	19,76	21,50	22,95	27,86	133

Przeprowadzono testowanie normalności rozkładu w grupach NC, MCI i AD, i w każdej z grup nie było podstaw do odrzucenia hipotezy o normalności rozkładu zmiennej SF, zatem do porównania grup zastosowano test parametryczny ANOVA, którego wynik wskazuje, że istnieje istotna statystycznie różnica między grupami dla zmiennej SF ( $p < 0,0001$ ). Poniższa tabela zawiera wyniki dla porównań parami.

Tabela 5 Wyniki testu post-hoc HSD Tukeya dla porównań parami.

Porównanie	Różnica	Dolna granica 95% przedziału ufności	Górna granica 95% przedziału ufności	Wartość p
MCI-AD	-0,66	-1,24	-0,08	0,0204612
NC-AD	-1,55	-2,18	-0,93	<0,0000001
NC-MCI	-0,89	-1,40	-0,38	0,0001268

Wyniki testu Tukeya pokazują (Tabela 5), że różnice dla zmiennej SF pomiędzy każdą parą grup: MCI-AD, NC-AD oraz NC-MCI są istotne statystycznie. Zatem wartość współczynnika pofałdowania jest istotnie wyższa u pacjentów z chorobą Alzheimera w porównaniu z osobami zdrowymi, jest również istotnie wyższa u chorych z łagodnym upośledzeniem funkcji poznawczych w porównaniu do osób zdrowych, a także u pacjentów z chorobą Alzheimera w porównaniu z chorymi z łagodnym upośledzeniem funkcji poznawczych. Największą średnią wartość współczynnika pofałdowania mają pacjenci z chorobą Alzheimera (21,43%, SD=2,505), a najmniejszą osoby zdrowe (19,88%, SD=2,282). Graficzne przedstawienie wyników zamieszczono na poniższym wykresie.



Rysunek 27 Współczynnik pofałdowania dla grup: NC, MCI i AD. Wykres typu ramka-wąsy.

### **4.1.3. Względna objętość komór bocznych mózgu**

Kurczenie się mózgu u pacjentów z chorobą Alzheimera powoduje powiększanie się komór bocznych mózgu. Jako zmienną opisującą tę sytuację wybrano względną objętość komór bocznych mózgu. Oczekujemy, że względna objętość komór bocznych będzie większa u pacjentów z chorobą Alzheimera niż u zdrowych osób, a nawet osób z łagodnym upośledzeniem funkcji poznawczych.

#### **4.1.3.1. Względna objętość komór bocznych mózgu - metoda wyznaczania**

Trzecią zmienną jest względna objętość komór bocznych mózgu  $V_{LV}$ . Do obliczenia objętości komór bocznych zastosowano algorytm Automatic Lateral Ventricle delineatioN (ALVIN). ALVIN to w pełni zautomatyzowany algorytm segmentacji komór bocznych z obrazów MRI (ALVIN działa w ramach SPM8 jako nakładka na MATLAB) [76]. Algorytm ALVIN działa w następujący sposób: z obrazu MRI segmentowany jest płyn mózgowo-rdzeniowy oraz obraz jest normalizowany, następnie za pomocą maski uzyskanej z obrazów dla 275 pacjentów w wieku od 18 do 94 lat wyznaczane są komory boczne mózgu i obliczana jest ich objętość. Uzyskana za pomocą algorytmu ALVIN objętość komór bocznych dzielona jest przez objętość całego mózgu, a następnie mnożona przez 100% i w ten sposób otrzymuje się względną objętość komór bocznych, która wyrażona jest w procentach.

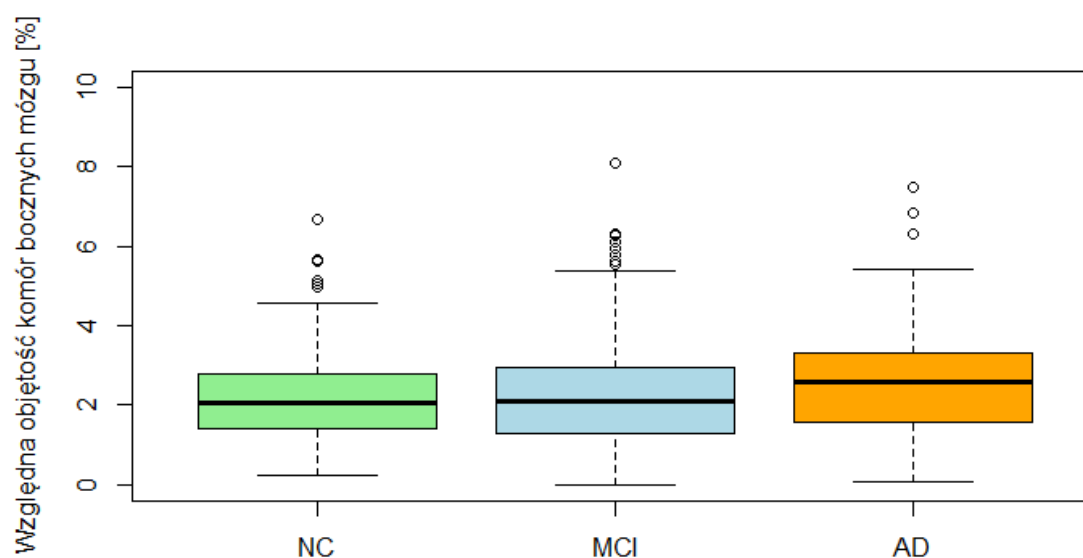
#### **4.1.3.2. Względna objętość komór bocznych mózgu - wyniki**

Rozwój choroby Alzheimera powoduje powiększenie komór bocznych mózgu. Wyniki statystyki opisowej dla zmiennej: względna objętość komór bocznych mózgu zamieszczono w poniższej tabeli.

Tabela 6 Statystyka opisowa dla zmiennej: względna objętość komór bocznych mózgu dla danych ze zbioru ADNI.

V <sub>LV</sub> [%]	Średnia	SD	IQR	Min	Q1	Mediana	Q3	Max	N
NC	2,20	1,139	1,38	0,23	1,42	2,07	2,79	6,66	194
MCI	2,25	1,352	1,69	0,00	1,27	2,09	2,97	8,10	311
AD	2,56	1,391	1,74	0,07	1,56	2,59	3,31	7,48	133

Przeprowadzono testowanie normalności rozkładu w grupach NC, MCI i AD, i w każdej z grup odrzucono hipotezy o normalności rozkładu zmiennej V<sub>LV</sub>, zatem do porównania grup zastosowano test nieparametryczny Kruskala-Wallisa, którego wynik wskazuje, że różnice między grupami dla zmiennej V<sub>LV</sub> są istotne statystycznie ( $p = 0,0217$ ). W analizie post-hoc zastosowany został test wariacji rang Conovera. Dla porównań parami otrzymano istotne statystycznie różnice pomiędzy grupami AD i NC ( $p=0,0067$ ) oraz AD i MCI ( $p=0,0054$ ). Graficzne przedstawienie wyników z podziałem na grupy zamieszczono na poniższym wykresie.



Rysunek 28 Względna objętość komór bocznych mózgu dla grup: NC, MCI i AD. Wykres typu ramka-wąsy.

## 4.2. Model matematyczny

Wielomianowa regresja logistyczna (ang. multinomial logistic regression) została zastosowana do predykcji statusu diagnozy. Ten typ regresji został wybrany ze względu na to, iż mamy trzy statusy diagnozy: choroba Alzheimera (ang. Alzheimer's disease, AD), łagodne upośledzenie funkcji poznawczych (ang. mild cognitive impairment, MCI) oraz status osoby zdrowej (ang. normal control, NC) (grupa referencyjna). Budując model rozważono następujące zmienne zależne: wiek, płeć, liczbę lat edukacji, punkty w skali MMSE, względną objętość mózgu bez płynu mózgowo-rdzeniowego, współczynnik pofałdowania oraz względną objętość komórek bocznych mózgu. Zmienną zależną był status diagnozy.

Ogólny wzór modelu wielomianowej regresji logistycznej ma postać:

$$\ln\left(\frac{p_{ij}}{p_{iJ}}\right) = \ln\left(\frac{p_{ij}}{1 - \sum_{j=1}^{J-1} p_{ij}}\right) = \sum_{k=0}^K x_{ik} \beta_{kj}$$

$$\text{dla } i = 1, \dots, N, \quad j = 1, \dots, J - 1,$$

gdzie  $X$  jest macierzą zmiennych niezależnych o  $N$  wierszach ( $N$  – liczba obserwacji) i  $K + 1$  kolumnach ( $K$  – liczba zmiennych niezależnych), w której pierwszy element każdego wiersza jest stałą,  $x_{i0} = 1$ .  $\beta$  jest macierzą o  $K + 1$  wierszach i  $J - 1$  kolumnach ( $J$  – liczba kategorii zmiennej zależnej), gdzie  $\beta_{kj}$  jest szacowanym współczynnikiem dla  $k$  – tej zmiennej niezależnej i  $j$  – tej wartości zmiennej zależnej.  $P$  jest macierzą o  $N$  wierszach i  $J - 1$  kolumnach, gdzie  $p_{ij}$  jest prawdopodobieństwem zaobserwowania  $j$  – tej wartości zmiennej zależnej dla dowolnej  $i$  – tej obserwacji [77].

Współczynniki wielomianowej regresji logistycznej były estymowane poprzez maksymalizację funkcji wiarygodności.

Logarytm funkcji wiarygodności dla modelu wielomianowej regresji logistycznej ma postać [77]:

$$l(\beta) = \sum_{i=1}^N \sum_{j=1}^{J-1} \left( y_{ij} \sum_{k=0}^K x_{ik} \beta_{kj} \right) - K \ln \left( 1 + \sum_{j=1}^{J-1} e^{\sum_{k=0}^K x_{ik} \beta_{kj}} \right)$$

gdzie  $y_{ij}$  jest wartością  $j$ -tej zmiennej zależnej dla  $i$ -tej obserwacji a pozostałe występujące symbole opisane są pod wcześniejszym wzorem.

Dla wartości współczynników modelu obliczono wartości dopasowanego parametru ilorazu szans wraz z 95% przedziałem ufności wyznaczonym według metody zaproponowanej przez Woolf'a [78].

Iloraz szans wyraża się wzorem:

$$OR = e^{\beta}$$

gdzie  $\beta$  jest wektorem wartości współczynników wielomianowego modelu regresji logistycznej, a OR (ang. odds ratio) jest wektorem wartości ilorazu szans.

Przedział ufności dla ilorazu szans wyraża się wzorem:

$$\left( e^{\beta - z_{1-\frac{\alpha}{2}} \cdot SE}, e^{\beta + z_{1-\frac{\alpha}{2}} \cdot SE} \right)$$

gdzie  $\alpha$  jest poziomem istotności (w naszym przypadku  $\alpha = 0,05$ ),  $z_{1-\frac{\alpha}{2}}$  jest dystrybuantą rozkładu normalnego dla  $1 - \frac{\alpha}{2}$ ,  $SE$  jest błędem standardowym dla współczynnika  $\beta$ .



Przeanalizowano również modele z dwukierunkowymi interakcjami. Została zastosowana 5-krotna walidacja krzyżowa dla zbioru ADNI [79]. Polegała ona na losowym podziale zbioru na pięć równolicznych podgrup. Analizę przeprowadzano pięciokrotnie budując model na zbiorze złożonym z czterech podgrup (80% pacjentów analizowanego zbioru ADNI) i testując na pozostałej piątej podgrupie (20% pacjentów analizowanego zbioru ADNI). Ponadto zbudowano model na całym analizowanym zbiorze ADNI oraz przeprowadzono walidację na niezależnym zbiorze danych ESDS. Bayesowskie kryterium informacyjne (BIC) zostało zastosowane do wyboru najlepszego modelu [80]. Kryterium to wyraża się wzorem:

$$BIC = -2 \ln(\hat{L}) + K \ln(N)$$

gdzie  $\hat{L}$  jest maksymalną wartością funkcji wiarygodności,  $K$  – liczbą zmiennych niezależnych, a  $N$  – liczbą obserwacji.

Zdefiniowano model podstawowy z jedną zmienną niezależną: MMSE, ponieważ wynik testu MMSE jest obecnie stosowanym testem przesiewowym [15]. Otrzymane w analizie modele porównywano z modelem podstawowym.

Porównywanie dwóch zagnieżdżonych modeli następowało przy zastosowaniu metody ANOVA, ponadto współczynnik Bayesa ( $\exp(\Delta BIC)$ ) był obliczany dla porównywanych modeli.

Krzywe ROC (ang. the receiver operating characteristic curve) wraz z wartościami parametru AUC (ang. the area under the curve) zostały wyznaczone dla obu zbiorów danych [81]. Krzywa ROC budowana jest na podstawie wartości zmiennej zależnej oraz przewidywanego prawdopodobieństwa tej zmiennej. Krzywa ROC pozwala na ocenę efektywności klasyfikacji zbudowanego modelu, a AUC jest

polem pod krzywą ROC. Wartość AUC pozwala na ocenę jakości klasyfikacji.  $AUC = 1$  oznacza klasyfikator idealny, a  $AUC = 0,5$  – klasyfikator losowy. Im wartość AUC bliższa 1, tym klasyfikator jest lepszy.

Dla każdej predykcji na zbiorze testowym obliczono wartości wskaźników jakości modelu predykcyjnego: czułości, swoistości, wartości predykcyjnej dodatniej, wartości predykcyjnej ujemnej, chorobowości oraz zbalansowanej dokładności. Poniższa tabela przedstawia macierz błędów (tabelę pomyłek), czyli wszystkie możliwe sytuacje dotyczące wyników testu w odniesieniu do rzeczywistości. Wynik testu mieści się w jednej z czterech kategorii: prawdziwie pozytywny (ang. true positive, TP), fałszywie pozytywny (ang. false positive, FP), fałszywie negatywny (ang. false negative, FN) i prawdziwie negatywny (ang. true negative, TN).

Tabela 7 Macierz błędów dla modelu predykcyjnego.

Macierz błędów		Rzeczywista diagnoza	
		Chory (+)	Zdrowy (-)
Wynik testu	Chory (+)	TP	FP
	Zdrowy (-)	FN	TN

Czułość to odsetek wyników prawdziwie pozytywnych wśród osób rzeczywiście chorych. Wskaźnik czułości (ang. sensitivity, true positive rate, TPR) opisuje się następującym wzorem:

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

Swoistość to odsetek wyników prawdziwie negatywnych wśród osób rzeczywiście zdrowych. Wzór opisujący czułość (ang. specificity, true negative rate, TNR) przedstawia się następująco:

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

Wartość predykcyjna dodatnia to odsetek wyników prawdziwie dodatnich wśród wszystkich wyników dodatnich testu. Wartość predykcyjna dodatnia (ang. positive predictive value, PPV) wyraża się wzorem:

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

Wartość predykcyjna ujemna to odsetek wyników prawdziwie negatywnych wśród wszystkich wyników ujemnych testu. Wartość predykcyjna ujemna (ang. negative predictive value, NPV) wyraża się wzorem:

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

Wskaźnik „chorobowość” w naszej sytuacji opisuje odsetek osób chorych wśród wszystkich poddanych badaniu i wyraża się wzorem:

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

Zbalansowana dokładność (ang. balanced accuracy, BACC) to średnia czułości i swoistości, czyli średnia odsetków poprawnie zakwalifikowanych pacjentów w każdej kategorii oddzielnie, podczas, gdy dokładność (ang. accuracy, ACC) to odsetek poprawnie zakwalifikowanych pacjentów wśród wszystkich badanych pacjentów. Zbalansowana dokładność wyrażona za pomocą czułości (TPR) i swoistości (TNR) przedstawia się następująco:

$$BACC = \frac{TPR + TNR}{2}$$

Dla 5-krotnej walidacji krzyżowej na zbiorze ADNI uzyskano pięć wartości dla każdego z wyżej wymienionych wskaźników, a następnie obliczono średnią oraz 95% przedział ufności.

W ostatnim etapie analizy porównano otrzymane z modelu wielomianowej regresji logistycznej wyniki predykcyjne z informacjami o zmianie diagnozy podczas okresu obserwacji. Zbudowano również model regresji logistycznej, by zbadać, które zmienne najlepiej opisują przejście ze statusu chorobowego MCI do AD. Zmienną zależną w modelu była zmiana statusu chorobowego z łagodnego upośledzenia funkcji poznawczych na chorobę Alzheimera podczas okresu obserwacji (zmiana z MCI na AD – 1; poziom referencyjny: stabilne MCI (brak zmiany diagnozy) – 0). W modelu rozważono następujące zmienne niezależne: wiek, płeć, liczbę lat edukacji, liczbę punktów MMSE, względną objętość mózgu bez płynu mózgowo-rdzeniowego, współczynnik pofałdowania oraz względną objętość komór bocznych mózgu. Porównano model podstawowy (tylko z jedną zmienną niezależną: liczba punktów MMSE) z modelem rozszerzonym (z dodanymi zmiennymi RBV oraz wiek) za pomocą metody ANOVA.

## 5. Wyniki

### 5.1. Model wielomianowej regresji logistycznej z walidacją krzyżową na zbiorze ADNI

Wielomianowa regresja logistyczna została zastosowana do predykcji statusu diagnozy. Dla każdej z pięciu walidacji krzyżowych otrzymano model o tej samej strukturze (z tymi samymi zmiennymi niezależnymi) kierując się bayesowskim kryterium informacyjnym (BIC). Najlepszy model opisujący status chorobowy zawiera następujące zmienne niezależne: względną objętość mózgu bez płynu CSF (RBV), punkty MMSE oraz wiek, gdzie zmienna wiek pełni funkcję korygującą naturalny proces starzenia się mózgu. Dodanie dwukierunkowych interakcji nie poprawiło modelu. Poniżej znajduje się schematyczny wzór otrzymanego modelu wielomianowej regresji logistycznej:

$$\text{Status chorobowy} \sim \text{RBV} + \text{MMSE} + \text{Wiek}$$

Porównanie modelu podstawowego, czyli bez zmiennych: względna objętość mózgu bez płynu (RBV) oraz wiek (w modelu wzięto pod uwagę jedynie zmienną „punkty MMSE”) z modelem zawierającym dodatkowo zmienne: względna objętość mózgu bez płynu (RBV) oraz wiek pokazało, że różnice pomiędzy modelami są istotne statystycznie ( $p < 0.00001$ ; BIC = 785,54 dla modelu bez dodatkowych zmiennych versus BIC = 809,68 dla modelu z dodanymi zmiennymi: RBV oraz wiek).

Wartość współczynnika Bayesa dla porównywanych modeli wynosi 132646731,7, co wskazuje na bardzo silny dowód za wyborem modelu z dodatkowymi zmiennymi (RBV oraz wiek) [82].

Poniższa tabela przedstawia średnie wartości współczynników (wraz z 95% przedziałem ufności) uzyskanych w 5-krotnej walidacji krzyżowej. Zdrowe, starsze osoby (NC) stanowią grupę kontrolną.

Tabela 8 Średnia wartość współczynników wielomianowego modelu regresji logistycznej dla 5-krotnej walidacji krzyżowej na zbiorze ADNI.

<b>Współczynniki</b>	<b>Stała regresji</b>	<b>RBV</b>	<b>Wiek</b>	<b>MMSE</b>
AD	85,05 [80,24; 89,86]	-0,3010 [-0,3653; -0,2368]	-0,1560 [-0,1760; -0,1360]	-2,0980 [-2,176; -2,019]
MCI	46,32 [43,55; 49,09]	-0,1774 [-0,2156; -0,1387]	-0,1023 [-0,1096; -0,0949]	-0,9755 [-1,0253; -0,9256]

Dla każdej średniej wartości współczynnika została obliczona wartość dopasowanego ilorazu szans. Wyniki w postaci skorygowanego ilorazu szans przedstawiono w poniższej tabeli.

Tabela 9 Skorygowany iloraz szans wraz z 95% przedziałem ufności dla uśrednionego modelu wielomianowej regresji logistycznej dla 5-krotnej walidacji krzyżowej na zbiorze ADNI.

<b>Skorygowany iloraz szans OR (95% CI)</b>	<b>RBV</b>	<b>Wiek</b>	<b>MMSE</b>
AD vs NC	0,74 [0,69; 0,79]	0,86 [0,84; 0,87]	0,12 [0,11; 0,13]
MCI vs NC	0,84 [0,81; 0,87]	0,90 [0,90; 0,91]	0,38 [0,36; 0,40]
AD vs MCI	0,88 [0,86; 0,91]	0,95 [0,94; 0,96]	0,33 [0,32; 0,33]

Z każdym spadkiem w zmiennej RBV o jeden punkt procentowy szansa wystąpienia choroby Alzheimera wzrasta 1,35 krotnie (95% CI [1,27; 1,44]) a szansa wstąpienia łagodnego upośledzenia funkcji poznawczych (MCI) wzrasta 1,19 krotnie (95% CI [1,15; 1,24]). Wśród pacjentów z łagodnym upośledzeniem

funkcji poznawczych, dla każdego spadku w RBV o jeden punkty procentowy szansa choroby Alzheimera wzrasta 1,13 krotnie (95% CI [1,10; 1,16]). Spadek o 1 punkt w punktacji MMSE zwiększa szansę wystąpienia choroby Alzheimera 8,15 krotnie (95% [7,53; 8,81]). Szansa wystąpienia łagodnego upośledzenia funkcji poznawczych wzrasta 2,65 krotnie (95%CI [2,52; 2,79]) dla każdej redukcji o 1 punkt w punktacji MMSE. Dla każdego spadku o 1 punkt w punktacji MMSE, szansa choroby Alzheimera wzrasta 3,07 krotnie (95% CI [2,99; 3,16]) wśród pacjentów z łagodnym upośledzeniem funkcji poznawczych.

## 5.2. Wskaźniki jakości modelu predykcyjnego – walidacja krzyżowa na zbiorze ADNI

W celu oceny jakości klasyfikacji przeanalizowana została macierz błędów. Poniższa macierz błędów zawiera wyniki predykcji w odniesieniu do początkowej diagnozy dla każdej z grup z 5-krotnej walidacji krzyżowej na zbiorze ADNI dla modelu rozszerzonego z dodatkowymi zmiennymi (model rozszerzony: status chorobowy ~ RBV + MMSE + Wiek) oraz dla modelu podstawowego bez dodatkowych zmiennych (model podstawowy: status chorobowy ~ MMSE).

Tabela 10 Macierz błędów dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla dwóch porównywanych modeli: rozszerzonego oraz podstawowego (z dodatkowymi zmiennymi oraz bez)

Model: status chorobowy ~ RBV + MMSE + Wiek				Model: status chorobowy ~ MMSE			
Grupa 1	Diagnoza początkowa			Grupa 1	Diagnoza początkowa		
Predykcja	NC	MCI	AD	Predykcja	NC	MCI	AD
NC	26	8	0	NC	29	16	0
MCI	13	50	8	MCI	10	41	9
AD	0	4	19	AD	0	5	18

<b>Grupa 2</b>		<b>Diagnoza początkowa</b>			<b>Grupa 2</b>		<b>Diagnoza początkowa</b>		
Predykcja		NC	MCI	AD	Predykcja		NC	MCI	AD
NC		31	13	0	NC		34	16	0
MCI		8	39	10	MCI		5	38	9
AD		0	10	17	AD		0	8	18
<b>Grupa 3</b>		<b>Diagnoza początkowa</b>			<b>Grupa 3</b>		<b>Diagnoza początkowa</b>		
Predykcja		NC	MCI	AD	Predykcja		NC	MCI	AD
NC		25	11	0	NC		28	18	0
MCI		14	47	14	MCI		11	41	15
AD		0	4	13	AD		0	3	12
<b>Grupa 4</b>		<b>Diagnoza początkowa</b>			<b>Grupa 4</b>		<b>Diagnoza początkowa</b>		
Predykcja		NC	MCI	AD	Predykcja		NC	MCI	AD
NC		29	10	0	NC		33	12	0
MCI		10	45	5	MCI		6	44	6
AD		0	7	21	AD		0	6	20
<b>Grupa 5</b>		<b>Diagnoza początkowa</b>			<b>Grupa 5</b>		<b>Diagnoza początkowa</b>		
Predykcja		NC	MCI	AD	Predykcja		NC	MCI	AD
NC		25	12	0	NC		26	17	0
MCI		13	46	11	MCI		12	40	9
AD		0	5	15	AD		0	6	17

Na podstawie powyższych macierzy błędów dla poszczególnych grup walidacji krzyżowej obliczono wartości wskaźników predykcji: czułości, swoistości, wartości dodatniej predykcyjnej, wartości ujemnej predykcyjnej, chorobowości oraz zbalansowanej dokładności. Ponieważ mamy trzy grupy: AD, MCI i NC, to wartości wskaźników predykcji obliczono dla następujących porównań: AD vs reszta (czyli MCI i NC), MCI vs reszta (czyli AD i NC) oraz NC vs reszta (czyli AD i MCI). Z tych trzech porównań istotne są dla nas dwa: AD vs reszta i NC vs reszta, i te dwa



porównania będą pojawiały się w dalszych analizach. Ponadto obliczono wartości wskaźników predykcji dla porównań parami: AD vs NC, MCI vs NC oraz AD vs MCI. Poniższa tabela zawiera wartości wskaźników predykcji dla wyszczególnionych wyżej porównań.

Tabela 11 Wskaźniki jakości modelu predykcyjnego rozszerzonego (z dodatkowymi zmiennymi) dla każdej grupy z 5-krotnej walidacji krzyżowej na zbiorze ADNI.

ADNI (Model rozszerzony: status chorobowy ~ RBV + MMSE + Wiek)

<b>Grupa 1</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NC</b>	<b>MCI vs NC</b>	<b>AD vs MCI</b>
Czułość [%]	70,37	80,65	66,67	100	86,21	70,37
Swoistość [%]	96,04	68,18	91,01	100	66,67	92,59
Wartość predykcyjna dodatnia [%]	82,61	70,42	76,47	100	79,37	82,61
Wartość predykcyjna ujemna [%]	92,38	78,95	86,17	100	76,47	86,21
Chorobowość [%]	21,09	48,44	30,47	42,22	59,79	33,33
Zbalansowana dokładność [%]	83,20	74,41	78,84	100	76,44	81,48
<b>Grupa 2</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NC</b>	<b>MCI vs NC</b>	<b>AD vs MCI</b>
Czułość [%]	62,96	62,90	79,49	100	75,00	62,96
Swoistość [%]	90,10	72,73	85,39	100	79,49	79,59
Wartość predykcyjna dodatnia [%]	62,96	68,42	70,45	100	82,98	62,96
Wartość predykcyjna ujemna [%]	90,10	67,61	90,48	100	70,45	79,59
Chorobowość [%]	21,09	48,44	30,47	35,42	57,14	35,53
Zbalansowana dokładność [%]	76,53	67,82	82,44	100	77,24	71,28
<b>Grupa 3</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	48,15	75,81	64,1	100	81,03	48,15

Swoistość [%]	96,04	57,58	87,64	100	64,10	92,16
Wartość predykcyjna dodatnia [%]	76,47	62,67	69,44	100	77,05	76,47
Wartość predykcyjna ujemna [%]	87,39	71,70	84,78	100	69,44	77,05
Chorobowość [%]	21,09	48,44	30,47	35,42	59,79	34,62
Zbalansowana dokładność [%]	72,09	66,69	75,87	100	72,57	70,15
<b>Grupa 4</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	80,77	72,58	74,36	100	81,82	80,77
Swoistość [%]	93,07	76,92	88,64	100	74,36	86,54
Wartość predykcyjna dodatnia [%]	75,00	75,00	74,36	100	81,82	75,00
Wartość predykcyjna ujemna [%]	94,95	74,63	88,64	100	74,36	90,00
Chorobowość [%]	20,47	48,82	30,71	42,00	58,51	33,33
Zbalansowana dokładność [%]	86,92	74,75	81,50	100	78,09	83,65
<b>Grupa 5</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	57,69	73,02	65,79	100	79,31	57,69
Swoistość [%]	95,05	62,5	86,52	100	65,79	90,20
Wartość predykcyjna dodatnia [%]	75,00	65,71	67,57	100	77,97	75,00
Wartość predykcyjna ujemna [%]	89,72	70,18	85,56	100	67,57	80,70
Chorobowość [%]	20,47	49,61	29,92	37,50	60,42	33,77
Zbalansowana dokładność [%]	76,37	67,76	76,15	100	72,55	73,94

Podobne obliczenia wartości wskaźników jakości predykcji wykonano dla modelu podstawowego z jedną zmienną niezależną: MMSE, bez dodatkowych zmiennych, dla każdej grupy w 5-krotnej walidacji krzyżowej na zbiorze ADNI. Poniższa tabela zawiera wyniki tych obliczeń.

Tabela 12 Wskaźniki jakości modelu predykcyjnego podstawowego (tylko ze zmienną niezależną MMSE, bez dodatkowych zmiennych) dla każdej grupy z 5-krotnej walidacji krzyżowej na zbiorze ADNI.

ADNI (Model podstawowy: status chorobowy ~ MMSE)

<b>Grupa 1</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	66,67	66,13	74,36	100	71,93	66,67
Swoistość [%]	95,05	71,21	82,02	100	74,36	89,13
Wartość predykcyjna dodatnia [%]	78,26	68,33	64,44	100	80,39	78,26
Wartość predykcyjna ujemna [%]	91,43	69,12	87,95	100	64,44	82,00
Chorobowość [%]	21,09	48,44	30,47	38,30	59,38	36,99
Zbalansowana dokładność [%]	80,86	68,67	78,19	100	73,14	77,90
<b>Grupa 2</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	66,67	62,90	87,18	100	70,37	66,67
Swoistość [%]	92,08	78,79	83,15	100	87,18	82,61
Wartość predykcyjna dodatnia [%]	69,23	73,58	69,39	100	88,37	69,23
Wartość predykcyjna ujemna [%]	91,18	69,33	93,67	100	68,00	80,85
Chorobowość [%]	21,09	48,44	30,47	34,62	58,06	36,99
Zbalansowana dokładność [%]	79,37	70,85	85,16	100	78,77	74,64
<b>Grupa 3</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	44,44	66,13	71,79	100	69,49	44,44

Swoistość [%]	97,03	60,61	79,78	100	71,79	93,18
Wartość predykcyjna dodatnia [%]	80,00	61,19	60,87	100	78,85	80,00
Wartość predykcyjna ujemna [%]	86,73	65,57	86,59	100	60,87	73,21
Chorobowość [%]	21,09	48,44	30,47	30,00	60,20	38,03
Zbalansowana dokładność [%]	70,74	63,37	75,79	100	70,64	68,81
<b>Grupa 4</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	76,92	70,97	84,62	100	78,57	76,92
Swoistość [%]	94,06	81,54	86,36	100	84,62	88,00
Wartość predykcyjna dodatnia [%]	76,92	78,57	73,33	100	88,00	76,92
Wartość predykcyjna ujemna [%]	94,06	74,65	92,68	100	73,33	88,00
Chorobowość [%]	20,47	48,82	30,71	37,74	58,95	34,21
Zbalansowana dokładność [%]	85,49	76,25	85,49	100	81,59	82,46
<b>Grupa 5</b>	<b>AD vs reszta</b>	<b>MCI vs reszta</b>	<b>NC vs reszta</b>	<b>AD vs NL</b>	<b>MCI vs NL</b>	<b>AD vs MCI</b>
Czułość [%]	65,38	63,49	68,42	100	70,18	65,38
Swoistość [%]	94,06	67,19	80,90	100	68,42	86,96
Wartość predykcyjna dodatnia [%]	73,91	65,57	60,47	100	76,92	73,91
Wartość predykcyjna ujemna [%]	91,35	65,15	85,71	100	60,47	81,63
Chorobowość [%]	20,47	49,61	29,92	39,53	60,00	36,11
Zbalansowana dokładność [%]	79,72	65,34	74,66	100	69,30	76,17

Dla 5-krotnej walidacji krzyżowej na zbiorze ADNI obliczono średnie wartości parametrów predykcji wraz z 95% przedziałem ufności. Wyniki przedstawiono w poniższej tabeli.

Tabela 13 Wskaźniki jakości (wraz z 95% przedziałem ufności) uśrednionego rozszerzonego modelu predykcyjnego z 5-krotną walidacją krzyżową na zbiorze ADNI.

Wskaźnik	AD vs reszta	NC vs reszta	AD vs NC	MCI vs NC	AD vs MCI
ADNI (Model rozszerzony: status chorobowy ~ RBV + MMSE + Wiek)					
Czułość [%]	63,99 [48,61; 79,37]	70,08 [61,92; 78,24]	100	80,67 [75,63; 85,72]	63,99 [48,61; 79,37]
Swoistość [%]	94,06 [90,93; 97,19]	87,84 [85,17; 90,51]	100	70,08 [61,92; 78,24]	88,22 [81,53; 94,9]
Wartość predykcyjna dodatnia [%]	74,41 [65,56; 83,26]	71,66 [67,12; 76,20]	100	79,84 [76,71; 82,96]	74,41 [65,56; 83,26]
Wartość predykcyjna ujemna [%]	90,91 [87,34; 94,47]	87,13 [84,19; 90,07]	100	71,66 [67,12; 76,20]	82,71 [76,16; 89,26]
Chorobowość [%]	20,84 [20,42; 21,26]	30,41 [30,05; 30,77]	38,51 [34,3; 42,73]	59,13 [57,50; 60,76]	34,11 [32,94; 35,29]
Zbalansowana dokładność [%]	79,02 [71,65; 86,4]	78,96 [75,23; 82,69]	100	75,38 [72,10; 78,65]	76,10 [68,52; 83,69]
AUC [%]	94,18 [92,09; 96,28]	90,01 [87,03; 92,99]	99,65 [99,18; 100,00]	79,30 [74,35; 84,24]	90,78 [87,45; 94,11]
Punkt odcięcia	0,16 [0,05; 0,26]	0,30 [0,20; 0,40]	8,38 [3,88; 12,87]	47,24 [35,28; 59,21]	30,13 [10,32; 49,95]

Wartości pola powierzchni pod krzywą ROC (AUC) są bardzo wysokie dla klas AD vs reszta oraz NC vs reszta i dla 5-krotnej walidacji krzyżowej wynoszą odpowiednio 94,18% (95% CI [92,09; 96,28]) oraz 90,01% (95% CI [87,03; 92,99]). Wartość zbalansowanej dokładności dla trzech klas: AD vs reszta, NC vs reszta, MCI vs reszta wynosi 76,10%. Dla klasy AD vs reszta osiągnięto wyższą wartość swoistości niż dla klasy NC vs reszta i wynosi ona 94,06% (95%CI [90,93; 97,19]), przy wartości czułości równej 63,99% (95%CI [48,61; 79,37]) dla tej klasy.

Wartość predykcyjna ujemna (NPV) wynosi 90,91% (95%CI [87,34; 94,47]) dla klasy AD vs reszta, przy wartości predykcyjnej dodatniej (Positive Predictive

Value - PPV) równej 74,41% (95%CI [65,56; 83,26]). Wartości swoistości, czułości, NPV i PPV dla klasy NC vs reszta wynoszą odpowiednio 87,84% (95%CI [85,17; 90,51]), 70,08% (95%CI [61,92; 78,24]), 87,13% (95%CI [84,19; 90,07]) oraz 71,66% (95%CI [67,12; 76,20]). Analiza grup parami daje bardzo duże wartości AUC dla klasyfikacji AD vs NC: 99,65% (95%CI [99,18; 100,00]). Wartości swoistości, czułości, NPV i PPV dla klasy AD vs NC wynoszą 100%.

### 5.3. Model wielomianowej regresji logistycznej testowany na niezależnym zbiorze EDSD

Model wielomianowej regresji logistycznej był również uczony na całym zbiorze ADNI i testowany na niezależnym zbiorze EDSD. Wartości współczynników modelu zostały zaprezentowane w poniższej tabeli.

Tabela 14 Wartości współczynników i błędów standardowych wielomianowego modelu regresji logistycznej zbudowanego na całym zbiorze ADNI.

Współczynniki	Stała regresji	RBV	Wiek	MMSE
AD ± SE	84,66 ± 1,515	-0,3003 ± 0,03962	-0,1547 ± 0,02376	-2,0875 ± 0,11191
MCI ± SE	46,17 ± 2,646	-0,1764 ± 0,02924	-0,1019 ± 0,01883	-0,9726 ± 0,07768

Jak poprzednio, dla każdej zmiennej predykcyjnej obliczono wartość skorygowanego ilorazu szans. Wyniki w postaci skorygowanego ilorazu szans przedstawiono w poniższej tabeli.

Tabela 15 Skorygowany iloraz szans dla zmiennych predykcyjnych w wielomianowym modelu regresji logistycznej dla całego zbioru ADNI.

Skorygowany iloraz szans (95% CI)	RBV	Wiek	MMSE
AD vs NC	0,74 [0,69; 0,80]	0,86 [0,82; 0,90]	0,12 [0,10; 0,15]
MCI vs NC	0,84 [0,79; 0,89]	0,90 [0,87; 0,94]	0,38 [0,32; 0,44]
AD vs MCI	0,88 [0,87; 0,90]	0,95 [0,94; 0,96]	0,33 [0,31; 0,35]

Można zauważyć, że z każdym spadkiem w RBV o 1 punkt procentowy szansa wystąpienia choroby Alzheimera wzrasta 1,35 krotnie (95% CI [1,25; 1,46]) a szansa wystąpienia łagodnego upośledzenia funkcji poznawczych wzrasta 1,19 krotnie (95% CI [1,13; 1,26]), co daje wartości bardzo zbliżone do uzyskanych w uśrednionym modelu dla 5-krotnej walidacji krzyżowej. Wśród pacjentów z łagodnym upośledzeniem funkcji poznawczych, dla każdej redukcji w zmiennej RBV o 1 punkt procentowy, szansa wystąpienia choroby Alzheimera wzrasta 1,13 krotnie (95% CI [1,10; 1,16]). Każdy spadek o 1 punkt w punktacji MMSE zwiększa 8,06 krotnie (95%CI [6,48; 10,04]) szansę wystąpienia choroby Alzheimera i 2,64 krotnie (95% [2,27; 3,08]) szansę wystąpienia łagodnego upośledzenia funkcji poznawczych (MCI). Wśród pacjentów z łagodnym upośledzeniem funkcji poznawczych spadek o 1 punkt w punktacji MMSE powoduje 3,05 krotny (95% [2,85; 2,36]) wzrost wystąpienia choroby Alzheimera.

#### **5.4. Wskaźniki jakości modelu predykcyjnego testowanego na niezależnym zbiorze EDSD**

Model uczony na całym analizowanym zbiorze ADNI był testowany na niezależnym zbiorze danych EDSD i poniższa tabela przedstawia wyniki dla parametrów predykcji uzyskanych z tego modelu. W tabeli tej zamieszczono wyniki zarówno dla ostatecznego modelu ze zmiennymi niezależnymi: RBV, wiek oraz punkty MMSE, jak i dla porównywanego modelu bez zmiennych RBV i wiek, a jedynie ze zmienną niezależną MMSE (model podstawowy).

Tabela 16 Wskaźniki jakości modelu predykcyjnego dla niezależnego zbioru testowego EDSD.

Wskaźnik	AD vs reszta	NC vs reszta	AD vs NC	MCI vs NC	AD vs MCI
<b>EDSD (Model rozszerzony: Status chorobowy ~ RBV + MMSE + Wiek)</b>					
Czułość [%]	69,85	73,20	96,94	75,78	71,43
Swoistość [%]	90,17	88,19	93,42	77,17	80,17
Wartość predykcyjna dodatnia [%]	73,64	80,68	90,48	69,78	79,83
Wartość predykcyjna ujemna [%]	88,39	83,01	97,93	82,08	71,85
Chorobowość [%]	28,22	40,25	39,20	41,03	52,36
Zbalansowana dokładność [%]	80,01	80,70	95,18	76,48	75,80
AUC [%]	89,95	85,36	93,25	71,08	85,74
<b>EDSD (Model podstawowy: Status chorobowy ~ MMSE)</b>					
Czułość [%]	71,32	69,59	97,98	77,78	72,39
Swoistość [%]	89,60	89,58	93,10	73,37	79,03
Wartość predykcyjna dodatnia [%]	72,93	81,82	90,65	66,67	78,86
Wartość predykcyjna ujemna [%]	88,83	81,39	98,54	82,82	72,59
Chorobowość [%]	28,22	40,25	40,57	40,65	51,94
Zbalansowana dokładność [%]	80,46	79,59	95,54	75,57	75,71
AUC [%]	90,19	85,87	93,35	67,04	86,14

Wyniki walidacji dla niezależnego zbioru danych EDSD pokazują, że wartości pola powierzchni pod krzywą ROC (AUC) dla klas: AD vs reszta i NC vs reszta wynoszą odpowiednio 89,95% i 85,36%. Wartość zbalansowanej dokładności dla trzech klas: AD vs reszta, NC vs reszta, MCI vs reszta jest równa 76,83%. Dla klasy AD vs reszta wartość czułości wyniosła 90,17%, swoistości – 69,85%, natomiast



wartość predykcyjna ujemna (NPV) jest równa 88,39% przy wartości predykcyjnej dodatniej wynoszącej 73,64%. Wartości czułości, swoistości, NPV i PPV dla klasy NC vs reszta są odpowiednio równe: 88,19%, 73,20%, 83,01% i 80,68%. Analiza grup parami potwierdziła bardzo wysoką wartość pola pod krzywą ROC (AUC) dla klasyfikacji AD vs NC, wartość ta wyniosła 93,25%. Swoistość dla klasy AD vs NC osiągnęła wartość 93,42%, przy czułości wynoszącej 96,94%, i jest to najwyższa wartość swoistości dla grup analizowanych parami. Wartość predykcyjna ujemna (NPV) jest równa 97,93% dla klasyfikacji AD vs NC przy wartości predykcyjnej dodatniej równej 90,48%.

Wyniki parametrów predykcyjnych uzyskane dla modelu podstawowego bez dodatkowych zmiennych (RBV oraz wiek) a wyłącznie ze zmienną niezależną MMSE potwierdzają, że dodanie do modelu zmiennych: RBV oraz wiek (jako zmienną pełniącą funkcję korygującą) poprawia jakość predykcyjną modelu. Wartość pola pod krzywą ROC (AUC) dla klasyfikacji MCI vs NC wzrasta z 67,04% (dla modelu z jedną zmienną niezależną: MMSE) do 71,08% (dla modelu z dodanymi zmiennymi RBV oraz wiek).

Ponadto obliczono wartości predykcyjne dla dwóch podgrup danych zbioru EDSD: 1,5 T i 3 T, wyniki zamieszczono w poniższej tabeli.

Tabela 17 Wskaźniki jakości modelu predykcyjnego dla podgrup 1,5T oraz 3T dla niezależnego zbioru testowego EDSO.

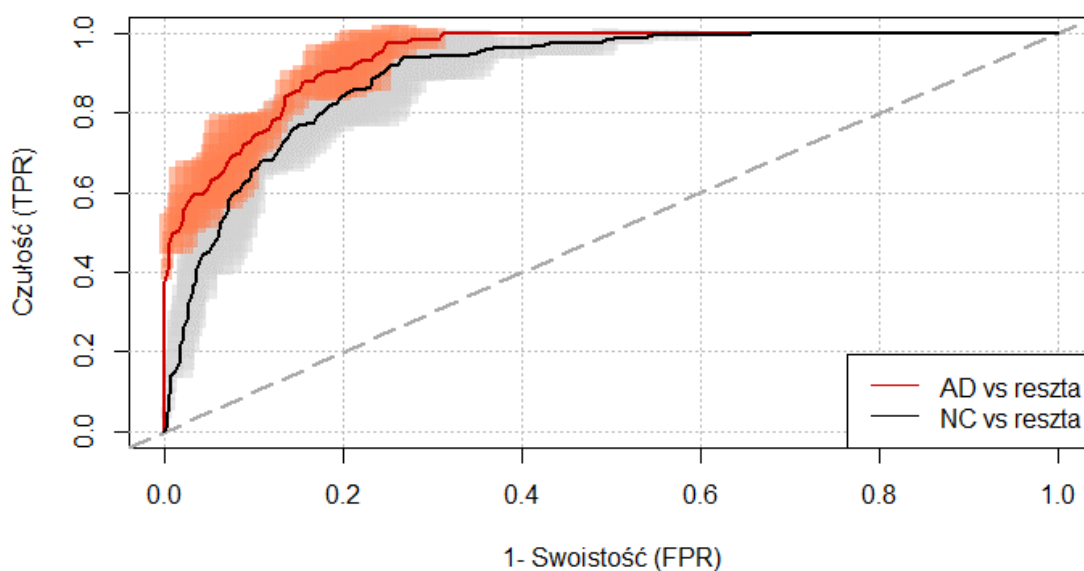
Wskaźnik	EDSD		EDSD – 1,5T		EDSD – 3T	
	AD vs reszta	NC vs reszta	AD vs reszta	NC vs reszta	AD vs reszta	NC vs reszta
Czułość [%]	69,85	73,20	64,71	70,00	72,94	74,31
Swoistość [%]	90,17	88,19	88,57	98,59	90,58	84,79
Wartość predykcyjna dodatnia [%]	73,64	80,68	80,49	97,22	70,45	76,43
Wartość predykcyjna ujemna [%]	88,39	83,01	77,50	82,35	91,58	83,26
Chorobowość [%]	28,22	40,25	42,15	41,32	23,55	39,89
Zbalansowana dokładność [%]	80,01	80,70	76,64	84,30	81,76	79,55
AUC [%]	89,95	85,36	88,10	89,15	90,40	84,46

### 5.5. Efektywność modelu predykcyjnego – krzywe ROC

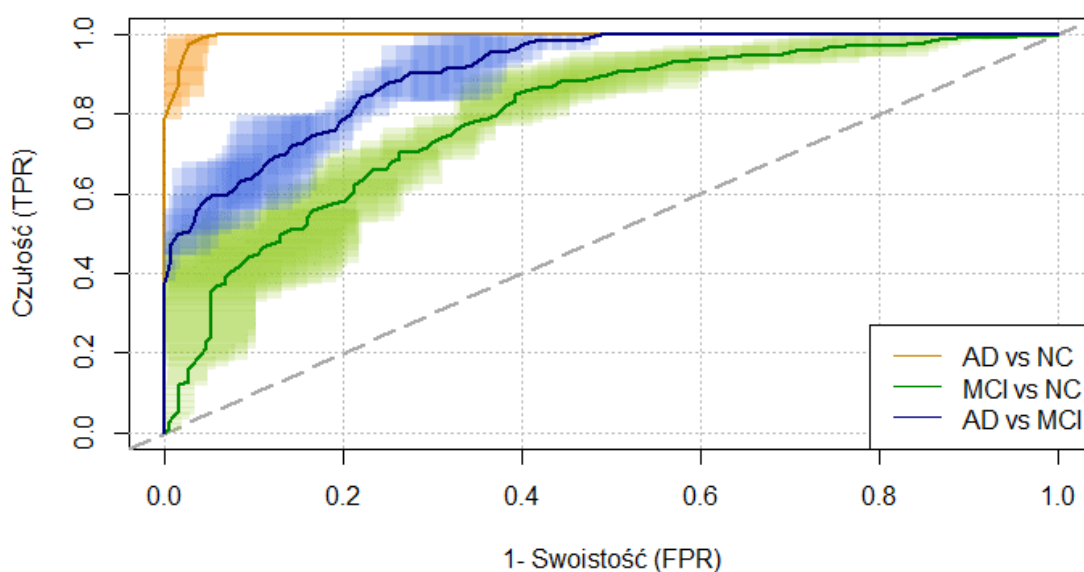
Efektywność rozszerzonego modelu predykcyjnego dla zbiorów danych ADNI i EDSO przedstawiono za pomocą krzywych ROC i zamieszczono na poniższym rysunku. Krzywe ROC zaprezentowano na oddzielnych wykresach z podziałem na zbiory danych (ADNI i EDSO) oraz sposób klasyfikacji (AD vs reszta i NC vs reszta oraz AD vs NC, MCI vs NC i AD vs MCI). Dla zbioru danych ADNI mamy uśredniony model predykcyjny z 5-krotną walidacją krzyżową, co pozwoliło na wyznaczenie 95% przedziału ufności dla krzywych ROC.

Na zbiorze ADNI dla porównań AD vs reszta oraz NC vs reszta wartości pola pod krzywą ROC są wysokie i wynoszą odpowiednio 94,18% (95% CI [92,09%; 96,28%]) i 90,01% (95% CI [87,03; 92,99]). Dla porównań parami bardzo wysoką

wartość AUC otrzymano dla AD vs NC i wynosi ona 99,65% (95% CI [99,18;100,00]). Wysoką wartość uzyskano również dla AD vs MCI: 90,78% (95%CI [87,45; 94,11]), a dla porównania MCI vs NC uzyskana wartość AUC wynosiła 79,30% (94%CI [74,35; 84,24]).

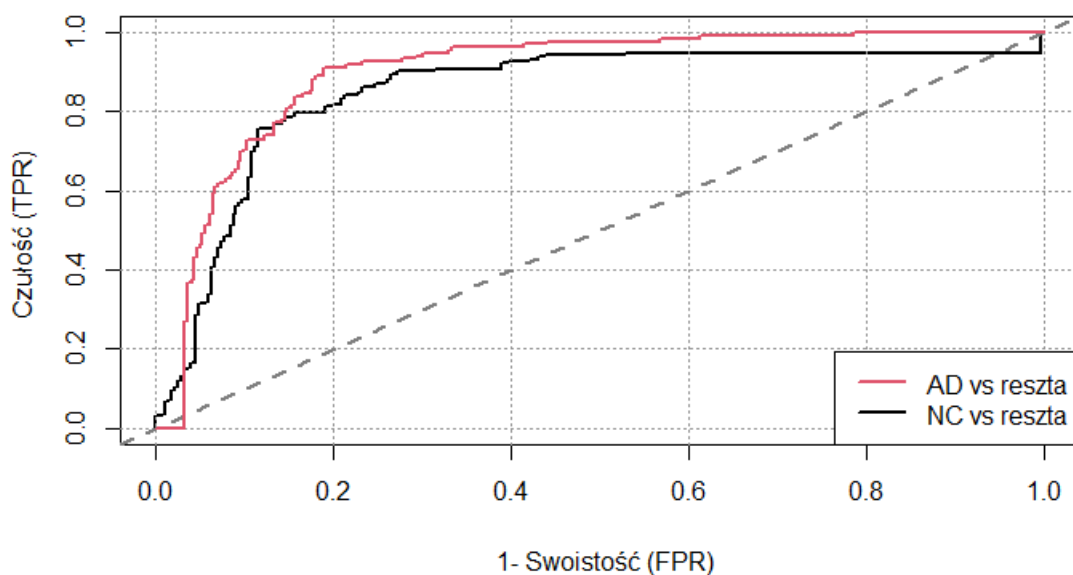


Rysunek 29 Krzywe ROC dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla: AD vs reszta oraz NC vs reszta. Na wykresie zamieszczono również 95% przedziały ufności dla krzywych ROC.

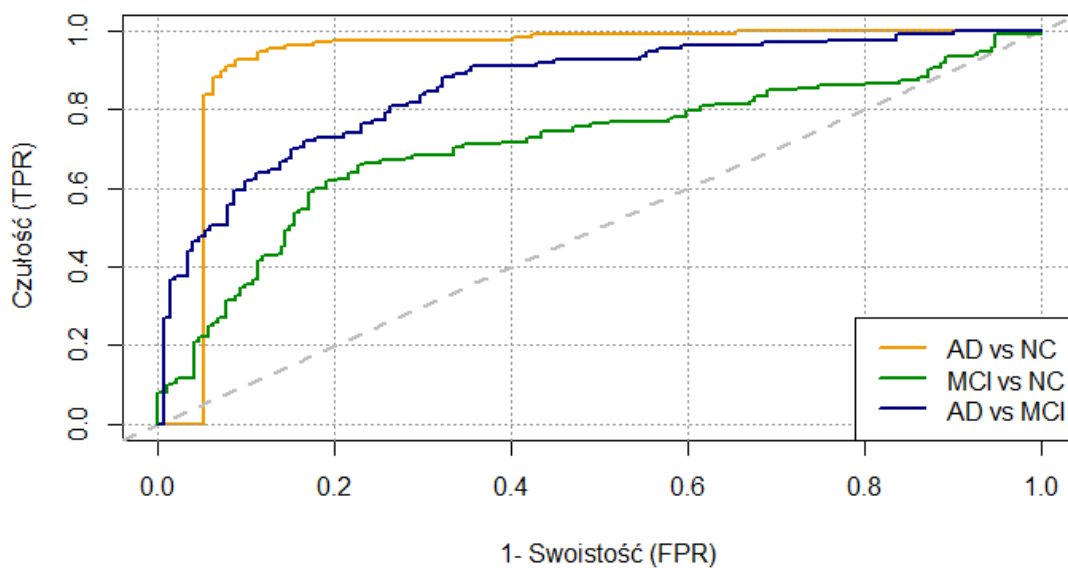


Rysunek 30 Krzywe ROC dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla: AD vs NC, MCI vs NC oraz AD vs MCI. Na wykresie zamieszczono również 95% przedziały ufności dla krzywych ROC.

Dla niezależnego zbioru EDSD wartości pola pod krzywą ROC również są wysokie i wynoszą 89,95% dla AD vs reszta oraz 85,36% dla NC vs reszta. Dla porównań parami bardzo wysoką wartość AUC uzyskano dla AD vs NC: 93,35%. Dla AD vs MCI wartość AUC również jest wysoka i wynosi 85,74%. Najniższą wartość AUC otrzymano dla MCI vs NC: 71,08%.



Rysunek 31 Krzywe ROC dla walidacji na zbiorze EDSD dla: AD vs reszta oraz NC vs reszta.



Rysunek 32 Krzywe ROC dla walidacji na zbiorze EDSD dla: AD vs NC, MCI vs NC oraz AD vs MCI.

## **5.6. Zgodność predykcji ze zmianą diagnozy podczas okresu obserwacji**

Podczas okresu obserwacji część pacjentów zmieniła status chorobowy z łagodnego upośledzenia funkcji poznawczych na chorobę Alzheimera, co jest naturalnym procesem postępu choroby. W tym rozdziale porównano otrzymane z modelu wyniki predykcyjne z informacjami o zmianie diagnozy podczas okresu obserwacji. Poniższa tabela zawiera liczbę i odsetek pacjentów, u których nastąpiła zmiana diagnozy w danych okresach obserwacji (6 miesięcy, 12 miesięcy, 18 miesięcy, 24 miesiące, oraz 36 miesięcy) i nowa diagnoza była zgodna z predykcją (dla zbioru danych ADNI) uzyskaną na podstawie modelu rozszerzonego o dodatkowe zmienne (RBV oraz wiek) a także dla modelu podstawowego (tylko ze zmienną niezależną MMSE). Ponadto w tabeli zamieszczono liczbę i odsetek pacjentów, u których nastąpiła zmiana diagnozy w ciągu 48 miesięcy obserwacji (dla zbiorów ADNI i EDSD), w tym przypadku przyjęto jako ostateczną diagnozę ostatnią dostępną diagnozę z okresu 48 miesięcy obserwacji. Po 12 miesiącach obserwacji w zbiorze ADNI dostępne są dane dla 30 pacjentów z początkową diagnozą MCI, którzy zostali zakwalifikowani przez rozszerzony model do grupy chorych z chorobą Alzheimera. Wśród tych pacjentów predykcja jest zgodna z diagnozą po 12 miesiącach obserwacji u 11 pacjentów (36,67%). Dla modelu podstawowego ten odsetek wynosi 32,14%. Podobne obliczenia zostały przeprowadzone dla 6, 18, 24, 36 i 48 miesięcznego okresu obserwacji.

Tabela 18 Zgodność predykcji ze zmianą diagnozy z MCI na AD podczas okresu obserwacji dla modelu rozszerzonego oraz podstawowego z podziałem na zbiór ADNI oraz EDSO.

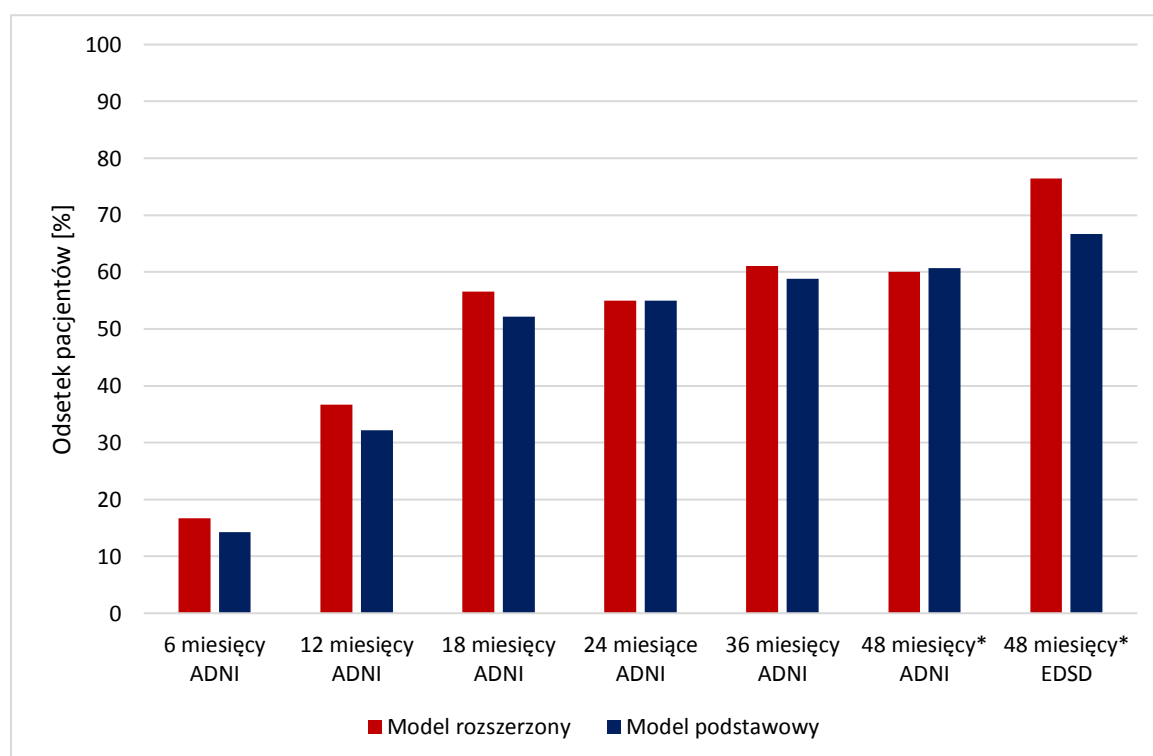
Zmiana diagnozy z MCI na AD	Zgodność predykcji AD ze zmianą diagnozy podczas okresu obserwacji	Predykcja AD wśród pacjentów z początkową diagnozą MCI (tylko pacjenci z wynikami po danym okresie obserwacji)	Zgodność predykcji AD ze zmianą diagnozy podczas okresu obserwacji	Predykcja AD wśród pacjentów z początkową diagnozą MCI (tylko pacjenci z wynikami po danym okresie obserwacji)	
	Liczba pacjentów (%)		Liczba pacjentów (%)		
Okres obserwacji	Model rozszerzony		Model podstawowy		
ADNI	6 miesięcy	5 (16,67%)	30 (100%)	4 (14,29%)	28 (100%)
	12 miesięcy	11 (36,67%)	30 (100%)	9 (32,14%)	28 (100%)
	18 miesięcy	13 (56,52%)	23 (100%)	12 (52,17%)	23 (100%)
	24 miesiące	11 (55,00%)	20 (100%)	11 (55,00%)	20 (100%)
	36 miesięcy	11 (61,11%)	18 (100%)	10 (58,82%)	17 (100%)
	48 miesięcy*	18 (60,00%)	30 (100%)	17 (60,71%)	28 (100%)
EDSO	48 miesięcy*	13 (76,47%)	17 (100%)	12 (66,67%)	18 (100%)

\* 48 miesięcy – przyjęto ostatnią dostępną diagnozę między 6 a 48 miesiącem obserwacji.

Predykcja była zgodna z diagnozą po 48 miesiącach obserwacji (przyjęto ostatnią dostępną diagnozę z 48 miesięcznego okresu obserwacji) dla 18 pacjentów (60%) spośród 30 pacjentów z początkową diagnozą: łagodne upośledzenie funkcji poznawczych, którzy zostali zakwalifikowani przez model do grupy chorych z chorobą Alzheimera (dla których dostępne są dane o diagnozie w okresie obserwacji). Ponadto, jeden pacjent zmienił diagnozę z choroby Alzheimera na łagodne upośledzenie funkcji poznawczych podczas okresu obserwacji i ta diagnoza jest zgodna z predykcją. Natomiast u 14 pacjentów początkowo zdrowych zdiagnozowano łagodne upośledzenie funkcji poznawczych, diagnoza z okresu obserwacji była zgodna z predykcją dla 5 pacjentów. Ponadto, wyniki dla modelu rozszerzonego o RBV oraz wiek pokazują, że odsetki poprawnych

predykcji (zgodnych z diagnozą z okresu obserwacji) są wyższe po 6, 12, 18 i 36 miesiącach obserwacji w porównaniu do modelu podstawowego.

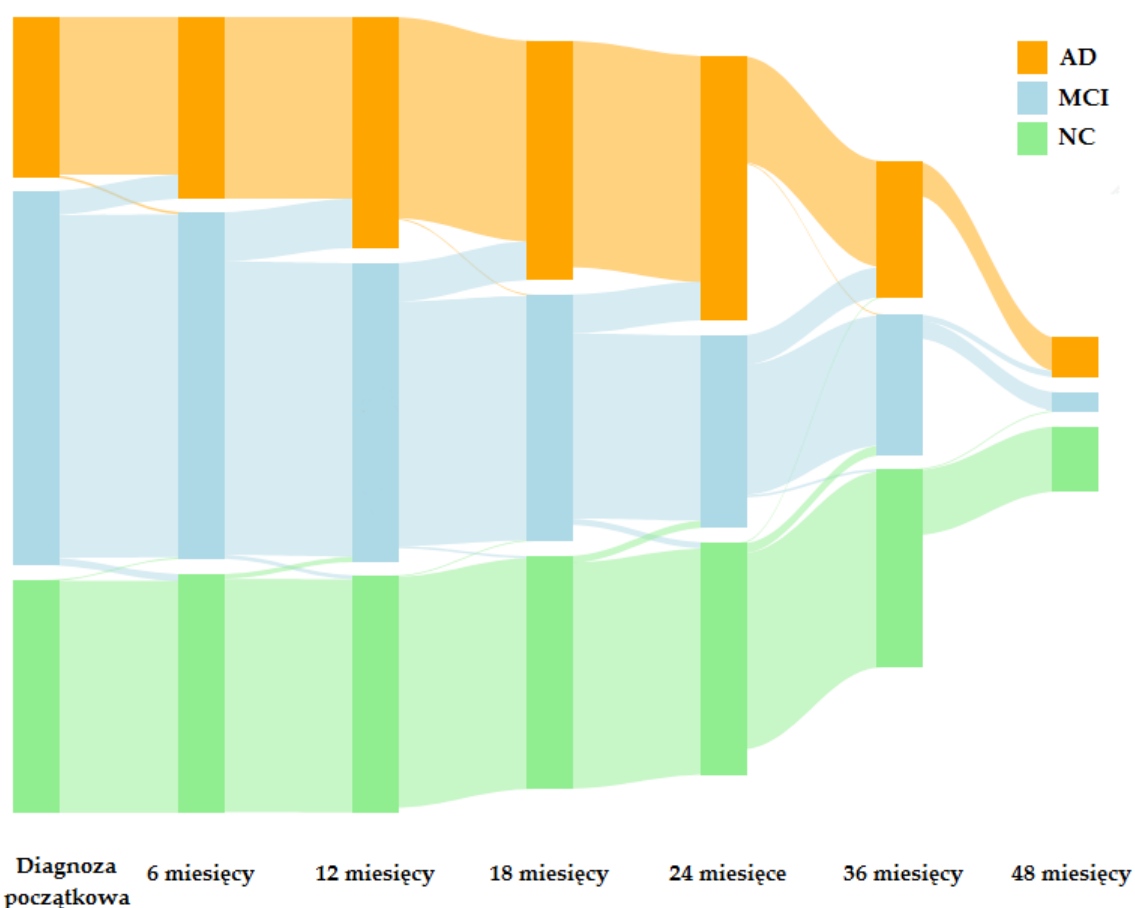
Predykcja na podstawie rozszerzonego wielomianowego modelu regresji logistycznej dla niezależnego zbioru EDSD pokazuje, że odsetek pacjentów, u których predykcja przejścia z MCI do AD została potwierdzona poprzez diagnozę w okresie obserwacji wynosi 76,47%. Dla modelu podstawowego na niezależnym zbiorze EDSD odsetek ten wynosi 66,67%. Wyniki potwierdzają, że dodanie do modelu podstawowego względnej objętości mózgu oraz wieku, jako zmiennej korygującej naturalny proces starzenia się mózgu, poprawia zdolności predykcyjne modelu. Poniższy wykres przedstawia wyniki zgodności zmiany diagnozy z predykcją dla zbiorów ADNI oraz EDSD z podziałem na model rozszerzony i podstawowy.



\* 48 miesięcy – przyjęto ostatnią dostępną diagnozę między 6 a 48 miesiącem obserwacji.

Rysunek 33 Zgodność predykcji ze zmianą diagnozy z MCI na AD podczas okresu obserwacji dla ADNI oraz EDSD z podziałem na model rozszerzony i podstawowy.

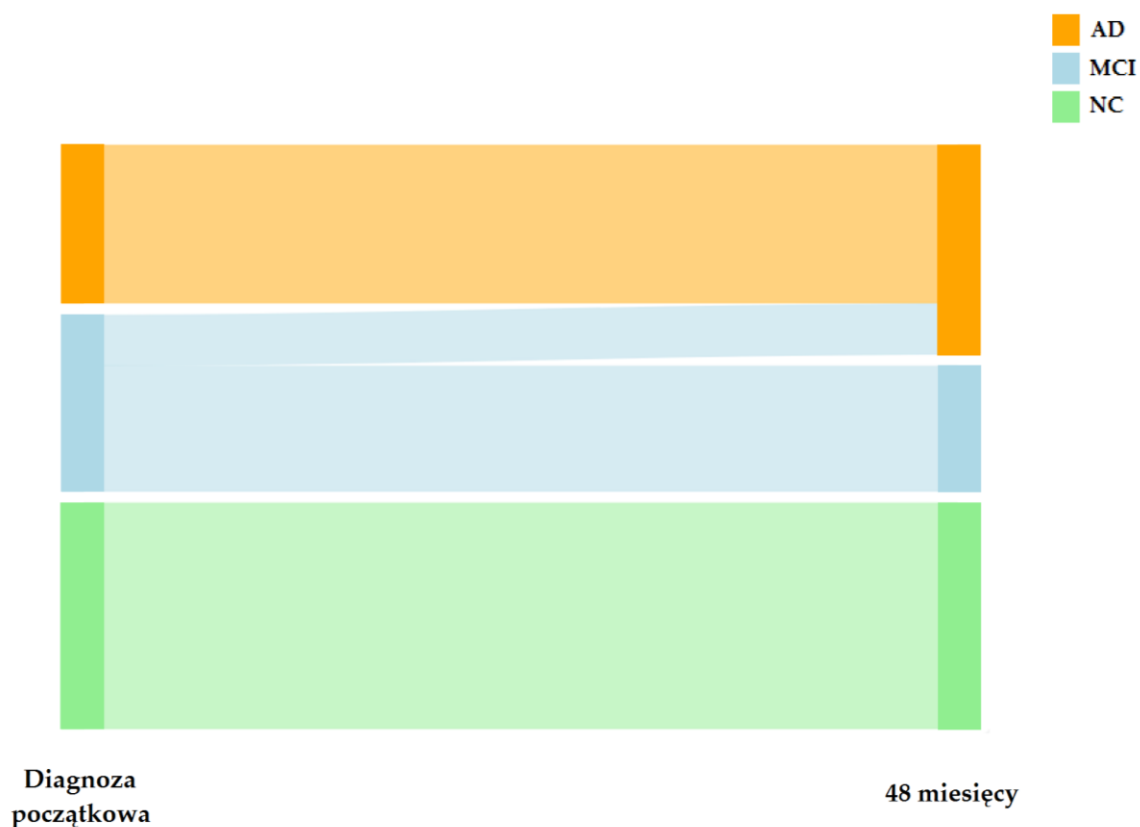
Zmiana w statusie diagnozy po 6, 12, 18, 24, 36 i 48 miesiącach została przedstawiona w postaci wykres wstęgowego (ang. Sankey plot). W przypadku brakujących danych dla konkretnych okresów obserwacji postępowano w następujący sposób: jeśli były dostępne dane dla co najmniej jednego późniejszego okresu obserwacji, to brakującą informację o diagnozie uzupełniano ostatnią dostępną diagnozą poprzedzającą okres, dla którego danej brakuje, a jeśli nie było dostępnych danych dla późniejszych okresów obserwacji, to brakującej danej nie uzupełniano.



Rysunek 34 Wykres wstęgowy dla zbioru ADNI: zmiana w statusie diagnozy po 6, 12, 18, 24, 36 i 48 miesiącach.



W przypadku zbioru danych EDSD nie posiadamy danych o diagnozie dla konkretnych punktów czasowych jak w ADNI. Ponadto informacja o zmianie diagnozy jest tylko dla pacjentów z początkową diagnozą AD lub MCI, nie ma informacji, czy nastąpiła zmiana diagnozy u osób zdrowych. W celu stworzenia wykresu wstęgowego dla zbioru EDSD przyjęto, że jako diagnozę po 48 miesiącach będzie brana ostatnia dostępna diagnoza z okresu obserwacji. Jeśli dla danego pacjenta nie ma informacji o zmianie diagnozy w okresie obserwacji, to przyjmowano, że zmiana diagnozy nie nastąpiła, zatem diagnoza końcowa jest taka sama jak diagnozą początkową. Brak danych o zmianie diagnozy jest dużym ograniczeniem w przypadku danych EDSD i powoduje, że wykres wstęgowy (ang. Sankey plot) jest bardzo uproszczony.



Rysunek 35 Wykres wstęgowy dla zbioru EDSD: zmiana w statusie diagnozy od diagnozy w momencie początkowym do diagnozy końcowej w 48 miesięcznym okresie obserwacji.

Ponadto został zbudowany model regresji logistycznej na całym analizowanym zbiorze ADNI w celu zbadania, które zmienne najlepiej opisują przejście ze statusu chorobowego MCI do AD. Zmiana statusu chorobowego z łagodnego upośledzenia funkcji poznawczych do choroby Alzheimera podczas okresu obserwacji (przyjęto ostatnią dostępną diagnozę z okresu obserwacji między 6 a 48 miesiącem) była zmienną zależną w modelu (zmiana z MCI na AD – 1, poziom referencyjny: stabilne MCI – 0). Wyniki modelu pokazują, że sposób analizowanych zmiennych niezależnych (wiek, płeć, liczba lat edukacji, liczba punktów MMSE, względna objętość mózgu bez płynu mózgowo-rdzeniowego, współczynnik pofałdowania oraz względna objętość komór bocznych mózgu) progresję statusu choroby z MCI do AD najlepiej opisują następujące zmienne: względna objętość mózgu bez płynu mózgowo-rdzeniowego (RBV), liczba punktów MMSE oraz wiek, gdzie wiek pełni funkcję korygującą naturalny proces starzenia się mózgu. Wynik porównania modelu podstawowego bez zmiennych RBV oraz wiek (czyli z tylko jedną zmienną niezależną: liczba punktów MMSE) z modelem rozszerzonym, z dodanymi zmiennymi RBV oraz wiek, jest istotny statystycznie ( $p = 0,00004$ ; BIC = 402,93 dla modelu rozszerzonego vs BIC = 411,80 dla modelu podstawowego), co oznacza, że model rozszerzony istotnie lepiej opisuje zmianę statusu chorobowego. Poniższa tabela przedstawia wartości współczynników modelu regresji logistycznej.

Tabela 19 Wyniki modelu regresji logistycznej dla zmiany statusu chorobowego z MCI do AD zbudowanego na całym zbiorze ADNI.

<b>Model regresji logistycznej</b>	<b>Wartość współczynnika</b>	<b>SE</b>	<b>Wartość p</b>
<b>Stała regresji</b>	21,54206	4,2627	<0,00001
<b>RBV</b>	-0,17156	0,04053	<0,00001
<b>Wiek</b>	-0,05739	0,02157	0,007811
<b>MMSE</b>	-0,26507	0,07177	0,000221

Dla współczynników obliczono wartości dopasowanego ilorazu szans wraz z 95% przedziałem ufności i wyniki zaprezentowano w poniższej tabeli.

Tabela 20 Skorygowany iloraz szans dla zmiennych predykcyjnych w modelu regresji logistycznej dla zmiany statusu choroby z MCI do AD (ADNI).

<b>Model regresji logistycznej</b>	<b>Skorygowany iloraz szans OR</b>	<b>Przedział ufności 95% CI</b>
<b>RBV</b>	0,842	[0,778; 0,912]
<b>Wiek</b>	0,944	[0,905; 0,985]
<b>MMSE</b>	0,767	[0,666; 0,883]

Wartości dopasowanego ilorazu szans pokazują, że dla każdego spadku o 1 punkt procentowy we względnej objętości mózgu bez płynu mózgowo-rdzeniowego (RBV) szansa progresji choroby z MCI do AD wzrasta 1,19 krotnie (95% CI [1,10; 1,29] oraz dla każdej redukcji o 1 punkt MMSE szansa progresji choroby wzrasta 1,30 krotnie (95% CI [1,13; 1,50]).

## 6. Dyskusja

Wyniki klasyfikacji z modelu wielomianowej regresji logistycznej na niezależnym zbiorze EDSO zostały porównane z wynikami przedstawionymi w literaturze. W tabeli zamieszczonej poniżej przedstawiono wyniki pochodzące z publikacji dla analiz przeprowadzonych na zbiorze ADNI. W 11 badaniach zbior ADNI był zbiorem uczących, w jednym badaniu zbiorem uczącym był lokalny zbiór danych a zbiór ADNI pełnił funkcję niezależnego zbioru danych do walidacji. Ponadto 6 spośród 12 badań posiada wyniki z testowania na niezależnym zbiorze danych.

Wyniki uzyskane dla testowania na niezależnym zbiorze danych EDSO są nie gorsze, a w wielu przypadkach lepsze niż wyniki z publikacji z tabeli 21. Model osiągnął najlepszy wynik dla zbalansowanej dokładności: 95,16%, podczas gdy, najwyższa wartość zbalansowanej dokładności dla AD vs NC z przedstawionych badań wynosi 85,5% [25]. Najwyższa wartość AUC (pole pod krzywą ROC) z wyników z publikacji wynosi 96,8%, jednak w modelu brano pod uwagę dodatkowe dane dotyczące stężenie amyloidów w płynie mózgowo-rdzeniowym [25]. Druga najwyższa wartość AUC wynosi 95,74%, jednak to badanie dotyczy tylko dwóch kategorii: AD i NC, a tu mamy dodatkowo kategorię MCI [27]. Trzecia z kolei wartość AUC wynosi 95,3%. Ta wartość jest wyższa od wartości uzyskanej w naszym modelu (93,25%), jednak wartości pozostałych parametrów predykcyjnych, takich jak zbalansowana dokładność, czułość, swoistość, PPV i NPV dla AD vs NC są wyższe w naszym modelu [25]. Najniższa wartość AUC dla AD vs NC wśród wyników badań prezentowanych w tabeli 21 wynosi 69% [24]. Tylko jedno badanie z prezentowanych w tabeli zawiera wyniki NPV i PPV dla AD vs NC, wartości tych wskaźników wynoszą odpowiednio: 95,3% i 73,4% dla modelu opartego na dodatkowych danych dotyczących stężenia amyloidów,

oraz 96,5% i 66,2% dla analizy bez danych dotyczących stężenia amyloidów [25]. Nasze rezultaty są lepsze, wartość NPV jest równa 97,93% dla AD vs NC, podczas gdy wartość PPV wynosi 90,48%. Najwyższa wartość czułości dla AD vs NC wśród porównywanych badań wynosi 95,6%, dla analizy, w której brano pod uwagę stężenie amyloidów w płynie mózgowo-rdzeniowym, a dla modelu bez danych dotyczących stężenia amyloidów wartość czułości wynosi 94,2%, podczas, gdy wartość z naszego modelu jest wyższa i wynosi 96,94% [25]. Dla swoistości najwyższa wartość uzyskana w badaniach wynosiła 98,13%, jednak w tym badaniu analizowano wyłącznie dwie kategorie: AD vs NC, co oznacza, że łatwiej było uzyskać lepszy rezultat niż dla trzech kategorii [27]. Druga najwyższa wartość dla swoistości wynosi 89,8% i jest ona niższa niż wartość uzyskana w naszym modelu (93,42%) [61]. Najniższa wartość swoistości wśród porównywanych badań wynosi 68,33% [24].

Porównanie wyników modelu z 5-krotną walidacją krzyżową na zbiorze ANDI z wynikami walidacji na zależnym zbiorze danych (tabela 21) pokazuje, że nasz model osiąga lepsze wyniki dla AD vs NC niż wszystkie prezentowane badania. Wyniki predykcji dla AD vs NC z prezentowanych badań pokazują, że najwyższa wartość AUC wynosi 98,5% [25], podczas gdy wynik dla naszego modelu wynosi 99,65%. Najwyższe wartości czułości i swoistości z prezentowanych badań wynoszą odpowiednio 95,6% [25] i 98,51% [53], podczas gdy nasz model osiąga wartość 100% dla obu wskaźników. Tylko jedno badanie zawiera wyniki NPV i PPV dla AD vs NC, wartości tych wskaźników są odpowiednio równe 96,6% i 92,7% dla danych dotyczących stężenia amyloidów a dla modelu bez danych dotyczących stężenia amyloidów wartości te wynoszą 95,2% (NPV) przy 81,5%

(PPV) [25], podczas, gdy wyniki osiągnięte przez nasz model są równe 100% dla NPV oraz PPV.

Tabela 21 Przegląd literatury dotyczącej modeli klasyfikacyjnych zbudowanych na danych ze zbioru ADNI.

Publikacja	Wielkość próby	Metody klasyfikacyjne	Typ danych	Walidacja	Grupy	Parametry	Wyniki
<b>Agostinho i inni, 2022</b> [24]	Wewnętrzny lokalny zbiór danych (n=41): AD (n=20), NC (n=21).	SVM	MRI, PiB-PET i DTI	Wewnętrzny lokalny zbiór danych oraz niezależny zbiór danych (ADNI (n=330): AD (n=166), NC (n = 164))	AD, NC	AUC, ACC, SEN, SPEC, BACC	Walidacja na zbiorze zależnym: AD vs NC: MRI: AUC=96%, ACC=92,05%, SEN=86,78%, SPEC=86,78%, BACC=92,05%; PiB PET: AUC=93%, ACC=90,53%, SEN=92%, SPEC=82, BACC=79,84%; DTI: AUC=86%, ACC=76,84%, SEN=76,17%, SPEC=82,09%, BACC=79,84%. Walidacja na zbiorze niezależnym: AD vs NC: MRI: AUC=81%, ACC=78,02%, SEN=74,12%, SPEC=82,29, BACC=78,20%; PiB PET: AUC=81%, ACC=76,87%, SEN=87,9%, SPEC=68,33%, BACC=78,12%; DTI: AUC=69%, ACC=62,79%, SEN=54,31%, SPEC=71,98%, BACC=63,15%.
<b>Gao i inni, 2022</b> [27]	1134 pacjentów: AD (n=454), NC (n=680).	3DMgNet (wielosiatkowa i konwolucyjna sieć neuronowa)	MRI	10-krotna walidacja krzyżowa oraz niezależny zbiór danych (AD (n=75), NC (n=59))	AD, NC	AUC, ACC, SEN, SPEC	Walidacja na zbiorze zależnym: ACC=92,13%, AUC=94,43%, SEN=88,42%, SPEC=95%. Walidacja na zbiorze niezależnym: ACC=87,91%, AUC=95,74%, SEN=79,73%, SPEC=98,31%.
<b>Goenka i inni, 2022</b> [28]	769 pacjentów: AD (n=70), MCI (n=224), NC (475)	CNN	MRI	633 obrazów ze zbioru ADNI	AD, MCI, NC	AUC, ACC	Walidacja na zbiorze zależnym: AD vs NC: ACC= 97,83%, AD vs MCI: ACC=98,68%, NC vs MCI: ACC=99,10%, NC vs MCI vs AD: ACC=98,26%. AD vs NC: AUC=94%, AD vs MCI: AUC=97%, NC vs MCI: AUC=99%, NC vs MCI vs AD: AUC=98%.
<b>Tang i inni, 2021</b> [29]	560 pacjentów: AD (n=80), EMCI (n=230), LMCI (n=110), NC (n=140)	SVM, RF, DT	MRI	10-krotna walidacja krzyżowa	AD, EMCI, LMCI, NC	AUC, ACC, SEN, SPEC	RF: NC vs AD: ACC=96,14%, SEN=88,14, SPE=92,81%, AUC=92%. NC vs EMCI: ACC=77,45%, SEN=79,51%, SPE=33,54%, AUC=59%. NC vs LMCI: ACC=87,56%, SEN=64,71%, SPE=83,94%, AUC=81%. EMCI vs AD: ACC=90,15%, SEN=93,51%, SPE=92,43%, AUC=85%. LMCI vs AD: ACC=84,54%, SEN=67,91, SPE=72,46%, AUC=89%.
<b>Dyrba i inni, 2021</b> [25]	633 pacjentów: AD (n=189), MCI (n=220), NC (n=254)	CNN	MRI i PET	1-krotna walidacja krzyżowa oraz trzy niezależne zbiory danych: ADNI-3 (n=575), AIBL	AD, MCI, NC	AUC, ACC, SEN, SPEC, BACC, PPV, NPV	Walidacja na zbiorze zależnym: AD vs NC: BACC=88.9%, SEN=94,2%, SPE=83,6%, PPV=81,5%, NPV=95,2% AUC=94,9%. MCI vs NC: BACC=74,5%, SEN=65,5%, SPE=83,6%, PPV=78,1%, NPV=74,1%, AUC=78,5%.

					(n=606), DELCODE (n=474).			amyloid-pozytywne AD vs amyloid-negatywne NC: BACC=94,9%, SEN=95,6%, SPE=94,3%, PPV=92,7%, NPV=96,6%, AUC=98,5%. amyloid-pozytywne MCI vs amyloid-negatywne NC: BACC=86,7%, SEN=79%, SPE=94,3%, PPV=91,6%, NPV= 96,6%, AUC=92,5%. Walidacja na zbiorze niezależnym DELCODE: AD vs NC: BACC=85,5%, SEN=94,2%, SPE=76,7%, PPV=66,2%, NPV=96,5% AUC=95,3%. MCI vs NC: BACC=71%, SEN=65,2%, SPE=76,7%, PPV=66,9%, NPV=75,3%, AUC=77,5%. amyloid-pozytywne AD vs amyloid-negatywne NC: BACC=83,3%, SEN=95,9%, SPE=70,7%, PPV=73,4%, NPV=95,3%, AUC=96,8%. amyloid-pozytywne MCI vs amyloid-negatywne NC: BACC=72,2%, SEN=73,7%, SPE=70,7%, PPV=71,2%, NPV= 73,2%, AUC=84%.
<b>Marzban i inni, 2020 [33]</b>	406 pacjentów: NC (n=185), MCI (n=106), AD (n = 115)	CNN	MRI i DTI	10-krotna walidacja krzyżowa	AD,NC, MCI	AUC, ACC, SEN, SPEC	AD vs NC: AUC=94%, ACC=93,5%, SEN=92,5%, SPEC=93,9%. MCI vs NC: AUC=84%, ACC=79,6%, SEN=62,7%, SPEC=89%	
<b>Li i inni, 2020 [53]</b>	404 pacjentów: NC (n=268), AD (n=136)	SVM	MRI	10-krotna walidacja krzyżowa oraz niezależny zbiór danych (AD (n=41), NC (n=25))	AD, NC	ACC, SEN, SPEC	Walidacja na zbiorze zależnym: AD vs NC: ACC=93,07%, SEN=94,12%, SPEC=98,51%. Walidacja na zbiorze niezależnym: AD vs NC: ACC=84,85%, SEN=85,36%, SPEC=84%.	
<b>Bae i inni, 2020 [52]</b>	390 pacjentów: AD (n=195), NC (n=195)	CNN	MRI	5-krotna walidacja krzyżowa oraz niezależny zbiór danych (AD (n=195), NC (n=195))	AD, NC	AUC, ACC, SEN, SPEC	Walidacja na zbiorze zależnym: AD vs NC: AUC=94%, ACC= 89%, SEN= 88%, SPEC=91%. Walidacja na zbiorze niezależnym: AD vs NC: AUC=88%, ACC= 83%, SEN= 76%, SPEC= 89%.	
<b>Liu i inni, 2020 [30]</b>	449 pacjentów: AD (n=97), MCI (n=233), NC (n=119)	CNN	MRI	5-krotna walidacja krzyżowa oraz niezależny zbiór danych (AD (n=45), MCI (n=46) i NC pacjentów (n=44)).	AD, MCI, NC	AUC, ACC, SEN, SPEC	Walidacja na zbiorze zależnym: AD vs NC: ACC=88,9%, SEN=86,6%, SPE=90,8%, AUC=92,5%. MCI vs NC: ACC=76,2%, SEN=79,5%, SPE=69,8%, AUC=77,5%. Walidacja na zbiorze niezależnym: AD vs NC: AUC=89,8%, MCI vs NC: AUC=72,2%.	



<b>Zhang i inni, 2019</b> [60]	857 pacjentów: NC (n=322), MCI (n=322), AD (n = 213)	Analiza oparta na teorii grafów	MRI	Losowy podział danych na 80% i 20% dla uczenia i testowania.	AD, MCI, NC	AUC	AD vs MCI + NC: AUC=73%, NC vs AD + MCI: AUC=72%, MCI vs AD + NC: AUC= 69%.
<b>Westman i inni, 2012</b> [37]	369 pacjentów: AD (n=96), MCI (n=162) i NC (n=111).	OPLS	MRI, PET, CSF	7-krotna walidacja krzyżowa	AD, MCI, NC	AUC, ACC, SEN, SPEC, PPV, NPV	AD vs NC: MRI z CSF: ACC= 91,8%, SEN=88,5%, SPEC=94,6%, PPV=93,4%, NPV=90,5% i AUC=95,8%. Tylko MRI: ACC=87%, SEN=83,3%, SPEC=90,1%, PPV=87,9%, NPV=86,2% i AUC=93%. Tylko CSF: ACC=81,6%, SEN=84,4%, SPEC=79,3%, PPV=77,9%, NPV=85,4% i AUC=86,1%. MCI vs NC: MRI z CSF: ACC=77,6%, SEN=72,8%, SPEC=84,7%, PPV=87,4%, NPV=68,1% i AUC=87,6%. Tylko MRI: ACC=71,8%, SEN=66,7%, SPEC=79,3%, PPV=82,4%, NPV=62,0% i AUC=81,5%. Tylko CSF: ACC=70,3%, SEN=66,7%, SPEC=75,7%, PPV=80,0%, NPV=60,9% i AUC=74,9%.
<b>Eskildsen i inni, 2012</b> [61]	808 pacjentów: AD (n=194), NC (n=226), pMCI (n=161), sMCI (n=227)	LDA	MRI (grubość kory i wiek)	Walidacja krzyżowa z pominięciem jednego wyjścia (ang. leave-one-out validation, LOO)	AD, NC, pMCI, sMCI	AUC, ACC, SEN, SPEC	Zależny zestaw cech: AD vs NC: ACC=85,5%, SEN=80,4%, SPEC=89,8%, AUC=92%. pMCI vs sMCI: ACC=67,8%, SEN=64,6%, SPEC=70%, AUC=68,2%. Niezależny zestaw cech: AD vs NC: ACC=87,4%, SEN=82,5%, SPEC=91,6%, AUC=93,1%. pMCI vs sMCI: ACC=68,3%, SEN=67,7%, SPEC=68,7%, AUC=74,7%.

Skróty: AIBL, Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing; DELCODE, DZNE multicenter observational study on Longitudinal Cognitive Impairment and Dementia; EMCI, early mild cognitive impairment – wczesne łagodne upośledzenie funkcji poznawczych; LMCI, late mild cognitive impairment – późne łagodne upośledzenie funkcji poznawczych; pMCI, progressive MCI – progresywne MCI; sMCI, stable MCI – stabilne MCI; CNN, convolutional neural network - konwolucyjne sieci neuronowe; LDA, linear discriminant analysis - liniowa analiza dyskryminacyjna; SVM, support vector machine - metoda wektorów nośnych; RF, random forest – las losowy; DT, decision tree – drzewa decyzyjne; AUC, area under receiver operating characteristic curve – pole powierzchni pod krzywą ROC; ACC, accuracy - dokładność; SEN, sensitivity - czułość; SPEC, specificity - swoistość; BAAC, balanced accuracy – zbalansowana dokładność; PPV, positive predictive value – wartość predykcyjna dodatnia; NPV, negative predictive value – wartość predykcyjna ujemna.

## 7. Podsumowanie

Głównym celem pracy było zbudowanie modeli predykcyjnych, które będą wspomagały diagnostykę choroby Alzheimera oraz łagodnych zaburzeń funkcji poznawczych z wykorzystaniem obrazowania metodą rezonansu magnetycznego.

Zaproponowano usprawnienie procesu klasycznej diagnostyki opartej na punktacji MMSE poprzez znalezienie biomarkerów, których wartości są powszechnie dostępne w dokumentacji medycznej, a także zaproponowano test przesiewowy wspierający diagnostykę choroby Alzheimera oraz łagodnego upośledzenia funkcji poznawczych na podstawie zbudowanego modelu predykcyjnego. Zaproponowano dwa własne biomarkery oparte na obrazowaniu T1-zależnym rezonansem magnetycznym: względną objętość mózgu bez płynu mózgowo-rdzeniowego, która opisuje kurczenie się mózgu oraz współczynnik pofałdowania, który opisuje zmiany w pofałdowaniu mózgu powstałe w skutek rozwoju choroby.

We wstępnym przetwarzaniu danych zastosowano algorytm N4ITK do korekty niejednorodności pola obrazu, następnie dokonano ekstrakcji mózgu. W kolejnym kroku dostosowany algorytm MiMSeg zastosowano do segmentacji mózgu na płyn mózgowo-rdzeniowy oraz tkankę mózgową. Algorytm ten w pierwotnej wersji służył do detekcji tkanki nowotworowej na podstawie obrazów DWI. W niniejszej pracy algorytm ten został dostosowany do naszej sytuacji, w której segmentujemy obrazy MR T1-zależne.

W celu wyznaczenia wartości zaproponowanych biomarkerów opartych na MRI (względna objętość mózgu oraz współczynnik pofałdowania) zastosowano własne algorytmy wyznaczania konturu tkanki mózgowej oraz detekcji i niwelacji „dziur” oraz „plam” na obrazach MRI.

W głównej analizie ze względu na to, iż dane pogrupowane są ze względu na status chorobowy, który ma trzy kategorie: zdrowy, łagodne upośledzenie funkcji poznawczych oraz choroba Alzheimera, zastosowano wielomianowy model regresji logistycznej. Model zbudowano w dwóch wersjach. Pierwsza wersja to model 5-krotnej walidacji krzyżowej na zbiorze ADNI, a druga wersja to model zbudowany na całym analizowanym zbiorze ADNI i walidowany na niezależnym zbiorze EDSD.

Zbudowane na zbiorze ADNI wielomianowe modele regresji logistycznej z 5-krotną walidacją krzyżową oraz model zbudowany na całym zbiorze danych ADNI i testowany na niezależnym zbiorze danych EDSD mają taką samą strukturę. Status choroby był najlepiej opisywany przez względną objętość mózgu bez płynu mózgowo-rdzeniowego (RBV), punktację MMSE oraz wiek. Rozważono również model podstawowy, w którym analizowano tylko jedną zmienną niezależną: liczbę punktów MMSE (jest to obecnie stosowany test przesiewowy). Wyniki porównania tych dwóch modeli: podstawowego tylko ze zmienną niezależną MMSE oraz rozszerzonego z dodanymi zmiennymi RBV (względna objętość) i wiek pokazują, że dodanie zmiennych RBV oraz wiek (jako zmiennej korygującej naturalne starzenie mózgu) poprawia własności predykcyjne modelu. Silnym dowodem przemawiającym również za wyborem modelu rozszerzonego jest wysoka wartość wskaźnika Bayesa. Można zauważyć, że jakość detekcji MCI wzrosła po dodaniu zmiennych RBV i wiek (AUC: 67.04% vs 71.08%), przy równoczesnym zachowaniu poziomu jakości dla detekcji AD (AUC: 93.35% vs 93.25%). Ostatecznym zatem modelem zastosowanym do predykcji statusu chorobowego jest model z następującymi zmiennymi niezależnymi: RBV, MMSE i wiek (gdzie wiek pełni funkcję korygującą naturalny proces starzenia się mózgu). Średnie wartości

współczynników wielomianowego modelu regresji logistycznej dla 5-krotnej walidacji krzyżowej oraz wartości dla modelu zbudowanego na całym zbiorze danych ADNI są podobne, co potwierdza jednorodność zbioru uczącego i spójność procesu diagnozy. Średnie wartości statystyk predykcji uzyskane w 5-krotnej walidacji krzyżowej dla ADNI pokazują, że mamy wybitne wyniki klasyfikacji AD vs NC oraz AD vs reszta, gdzie AUC wynosi odpowiednio 99,65% i 94,18%. Ponadto, wartość AUC dla AD vs MCI oraz NC vs reszta są również bardzo wysokie (90,78% i 90,01%, odpowiednio). Średnie wartości AUC uzyskano dla klasyfikacji MCI vs NC (79,30%), ale jest to nadal dobry wynik, szczególnie, jeśli weźmiemy pod uwagę, że grupa MCI jest niejednorodna i u części pacjentów z tej grupy rozwija się choroba Alzheimera a część pacjentów pozostaje na stabilnym poziomie łagodnych upośledzeń funkcji poznawczych. Średnia wartość zbalansowanej dokładności dla trzech klas (AD vs reszta, NC vs reszta i MCI vs reszta) wynosi 76,10% dla 5-krotnej walidacji krzyżowej. Ponieważ celem jest rozwój wspierającego testu diagnostycznego, gdzie detekcja pacjentów z chorobą jest najważniejsza, zatem wskaźniki NPV oraz PPV odgrywają tutaj kluczową rolę. Dla 5-krotnej walidacji krzyżowej dla klasyfikacji AD vs NC wartość wskaźnika NPV wynosi 100%, przy PPV wynoszącym również 100%. Dla klasyfikacji AD vs reszta wyniki te przedstawiają się następująco: NPV - 90,91%, przy PPV - 74,41%.

Wyniki walidacji modelu na niezależnym zbiorze danych EDSD potwierdzają wysoką jakość predykcyjną modelu. Wartości pola powierzchni pod krzywą ROC (AUC) dla klas AD vs reszta oraz NC vs reszta są wysokie i wynoszą odpowiednio 89,95% i 85,36%. Bardzo wysoką wartość AUC osiągnięto dla AD vs NC (93,25%), co oznacza, że model bardzo dobrze rozróżnia pacjentów z chorobą Alzheimera od starszych zdrowych osób. Dla AD vs MCI uzyskana wartość AUC wynosi

85,74%. Wartość zbalansowanej dokładności dla trzech klas: AD vs reszta, NC vs reszta, MCI vs reszta jest równa 76,83%. Uzyskano również wysokie wartości wskaźników NPV i PPV. Najwyższe wartości dotyczą AD vs NC, gdzie wartość NPV jest bardzo wysoka i wynosi 97,93% przy wartości PPV równej 90,48%. Również wysokie wartości uzyskano dla AD vs reszta, gdzie NPV jest równe 88,39%, przy PPV równym 73,64%. Model dobrze sobie radzi również z rozpoznawaniem pacjentów cierpiących na łagodne upośledzenie funkcji poznawczych, PPV dla MCI vs NC wynosi 82,08% przy PPV równym 69,78%. Swoistość dla klasy AD vs NC osiąga wartość 93,42%, przy czułości wynoszącej 96,94%, i jest to najwyższa wartość swoistości dla grup analizowanych parami. Dla AD vs reszta swoistość wynosi 90,17% przy czułości równej 69,85%.

Wyniki parametrów predykcyjnych uzyskane dla modelu podstawowego (wyłącznie ze zmienną niezależną MMSE) potwierdzają, że dodanie do modelu zmiennych: RBV oraz wiek (jako zmienną pełniącą funkcję korygującą naturalny proces starzenia się mózgu) poprawia jakość predykcyjną modelu. Wartość pola pod krzywą ROC (AUC) dla klasyfikacji MCI vs NC wzrasta z 67,04% (dla modelu podstawowego z jedną zmienną niezależną: MMSE) do 71,08% (dla modelu rozszerzonego z dodanymi zmiennymi RBV oraz wiek). Oznacza to, że dodanie do modelu naszego biomarkera opartego na T1-zależnym MRI, jakim jest względna objętość mózgu istotnie poprawia efektywność modelu.

Ponadto, co bardzo istotne, wśród pacjentów z początkową diagnozą MCI prawidłowo przewidziano diagnozę AD u 60% pacjentów z ADNI oraz 76,47% z EDS (spośród pacjentów z początkową diagnozą MCI i predykcją AD; przejście ze statusu MCI do AD nastąpiło w ciągu 48 miesięcy obserwacji).

## 8. Wnioski

Praca pokazuje, że zaproponowany rozszerzony model predykcyjny może stać się użytecznym w codziennej praktyce lekarskiej testem przesiewowym. Proponowany biomarker oparty na danych obrazowych MRI (T1-zależnych) w połączeniu z punktacją MMSE oraz korektą dla wieku posiada bardzo dobre zdolności predykcyjne w odniesieniu do statusu chorobowego AD. Ponadto, zaproponowana metoda oparta na danych MRI nie wymaga inwazyjnych i drogich badań laboratoryjnych oraz bazuje na klasycznych metodach uczenia maszynowego, a więc nie wymaga dużej mocy obliczeniowej.

Wiele publikacji opisuje raczej proces diagnozy niż testy przesiewowe i tylko jedno z porównywanych badań zawiera wyniki dotyczące NPV oraz PPV, podczas gdy wyniki dotyczące czułości i swoistości dostępne są prawie we wszystkich pracach. Nasz model na niezależnym zbiorze danych osiąga lepsze wyniki NPV i PPV zarówno dla AD vs NC, jak i dla MCI vs NC, co oznacza, że jako test przesiewowy spełnia bardzo dobrze swoje zadanie. Publikowanych jest wiele zaawansowanych metod (np. CNN) z bardzo wysokimi wynikami, jednak te metody nie są proste do zastosowania w codziennej praktyce lekarskiej. Metody te wymagają dużej mocy obliczeniowej, a ponadto są czułe na protokoły pomiaru oraz wstępne przetwarzanie danych. Potrzeba jeszcze wiele czasu, zanim metody te będą mogły być stosowane w codziennej praktyce klinicznej. Nasz model bazuje na łatwo dostępnych danych, które można znaleźć w historii choroby pacjentów z podejrzeniem choroby Alzheimera. T1-zależne MRI jest standardowym, w przypadku pacjentów z demencją, powszechnie dostępnym i niedrogim w porównaniu z PET badaniem, również kwestionariusz MMSE jest standardowym testem neuropsychologicznym. Klasyczne metody statystyczne mają również

przewagę interpretacyjną. Dzięki uzyskanym, w wyniku modelowania, wartościom skorygowanego ilorazu szans można wyznaczyć szansę wystąpienia choroby dla konkretnego pacjenta (w danym wieku, z danymi punktami MMSE oraz daną względną objętością mózgu). Ponadto zastosowanie klasycznych metod uczenia maszynowego, które nie wymagają dużej mocy obliczeniowej powoduje, że metody są praktycznie gotowe do zastosowania w codziennej praktyce lekarskiej.

W pracy udowodnione zostało, że włączenie biomarkera opartego na MRI do standardowych klinicznych predyktorów AD prowadzi do powstania użytecznego w codziennej praktyce klinicznej modelu usprawniającego proces diagnozy poprzez wstępną detekcję pacjentów cierpiących na chorobę Alzheimera lub łagodne upośledzenie funkcji poznawczych. Ponadto, wykazane zostało, że model rozpoznaje pacjentów zmieniających status choroby z MCI do AD podczas okresu obserwacji jako pacjentów z chorobą Alzheimera już na kilka lat wcześniej niż standardowy, medyczny proces diagnozy, co pozwala na wcześniejsze podjęcie leczenia i opóźnienie postępu choroby. W obliczu ostatnich doniesień o trzeciej fazie badań nad nowym lekiem (Lecanemab) spowalniającym tempo pogarszania się funkcji poznawczych u pacjentów we wczesnym stadium choroby Alzheimera, kluczową kwestią jest jak najszybsze rozpoznanie choroby, które daje nadzieję na przedłużenie życia i opóźnienie pogorszenia jakości życia.

Podsumowując, zaproponowane modele predykcyjne posiadają bardzo wysoką efektywność w rozpoznawaniu pacjentów z chorobą Alzheimera oraz łagodnymi zaburzeniami poznawczymi i są możliwe do zastosowania, jako testy przesiewowe, w codziennej praktyce lekarskiej.

## 9. Podziękowania

W latach 2014-2016 autorka rozprawy była stypendystą programu „DoktoRIS - Program stypendialny na rzecz innowacyjnego Śląska” finansowanego przez Unię Europejską.



Projekt współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego

Podziękowania dla twórców baz ADNI oraz EDSD. Analizę przeprowadzono na danych ANDI pochodzących z wieloośrodkowego badania obserwacyjnego „The Alzheimer’s Disease Neuroimaging Initiative”, którego celem jest zebranie serii danych z obrazowania rezonansem magnetycznych (MRI), pozytonowej tomografii emisyjnej (PET), oraz danych dotyczących biomarkerów biologicznych a także danych klinicznych i neuropsychologicznych pozwalających na ocenę progresji łagodnych zaburzeń poznawczych (MCI) i wczesnej choroby Alzheimera. Model testowano na danych pochodzących z bazy EDSD. Badanie EDSD („The European DTI Study on Dementia”) rozpoczęło się w 2010 roku. Koordynatorem badania jest Niemieckie Centrum Chorób Neurodegeneracyjnych (DZNE).



## 10. Spis skrótów

ACC	accuracy – dokładność;
AD	Alzheimer's disease – choroba Alzheimera;
ADNI	Alzheimer's Disease Neuroimaging Initiative;
AIBL	Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing;
ALVIN	Automatic Lateral Ventricle delineatioN;
AUC	area under the curve – pole powierzchni pod krzywą;
BAAC	balanced accuracy – zbalansowana dokładność;
BIC	Bayesian information criterion – bayesowskie kryterium informacyjne;
CI	confidence interval – przedział ufności;
CNN	convolutional neural network – konwolucyjne sieci neuronowe;
CSF	cerebrospinal fluid – płyn mózgowo-rdzeniowy;
CT	computed tomography – tomografia komputerowa;
DELCODE	DZNE multicenter observational study on Longitudinal Cognitive Impairment and Dementia;
DT	decision tree – drzewo decyzyjne;
DTI	diffusion tensor imaging – obrazowanie tensora dyfuzji;
DWI	diffusion-weighted imaging – obrazowanie dyfuzyjne;
EDSD	European DTI Study on Dementia;
EMCI	early mild cognitive impairment – wczesne stadium łagodnego upośledzenia funkcji poznawczych;
FDG-PET	fluorodeoxyglucose positron emission tomography imaging - pozytonowa tomografia emisyjna po podaniu fluorodeoksyglukozy;
fMRI	functional MRI – funkcjonalne obrazowanie metodą rezonansu magnetycznego;
LDA	linear discriminant analysis – liniowa analiza dyskryminacyjna;

LMCI	late mild cognitive impairment – późne stadium łagodnego upośledzenia funkcji poznawczych;
MCI	Mild Cognitive Impairment – łagodne upośledzenie funkcji poznawczych;
MMSE	Mini-Mental State Examination score – Krótka Skala Zaburzeń Poznawczych;
MRI	magnetic resonance imaging – obrazowanie metodą rezonansu magnetycznego;
MTA	medial temporal lobe atrophy – atrofia przyśrodkowej części płata skroniowego;
NC	Normal Control – grupa kontrolna złożona z osób zdrowych;
NPV	negative predictive value – wartość predykcyjna ujemna;
PET	positron emission tomography – pozytonowa tomografia emisyjna;
pMCI	progressive MCI – postępujące łagodne upośledzenie funkcji poznawczych;
PPV	positive predictive value – wartość predykcyjna dodatnia;
RBV	relative brain volume – względna objętość mózgu;
RF	random forest – las losowy;
ROC	receiver operating characteristic curve – wykres charakterystyczny klasyfikatora;
SEN	sensitivity – czułość
SF	shrinkage factor – współczynnik pofałdowania;
sMCI	stable MCI – stabilne łagodne upośledzenie funkcji poznawczych;
SPEC	specificity – swoistość
SVM	support vector machine – metoda wektorów nośnych.

## 11. Spis tabel

Tabela 1 Charakterystyka kliniczna analizowanych zbiorów z baz ADNI i EDSD. ....	31
Tabela 2 Statystyka opisowa zmiennej względna objętość mózgu (RBV) dla zbioru danych ADNI.....	52
Tabela 3 Wyniki testu post-hoc HSD Tukeya dla porównań parami. ....	52
Tabela 4 Statystyka opisowa dla współczynnika pośladowania SF dla danych ADNI. ....	59
Tabela 5 Wyniki testu post-hoc HSD Tukeya dla porównań parami. ....	60
Tabela 6 Statystyka opisowa dla zmiennej: względna objętość komór bocznych mózgu dla danych ze zbioru ADNI. ....	62
Tabela 7 Macierz błędów dla modelu predykcyjnego. ....	66
Tabela 8 Średnia wartość współczynników wielomianowego modelu regresji logistycznej dla 5-krotnej walidacji krzyżowej na zbiorze ADNI.....	70
Tabela 9 Skorygowany iloraz szans wraz z 95% przedziałem ufności dla uśrednionego modelu wielomianowej regresji logistycznej dla 5-krotnej walidacji krzyżowej na zbiorze ADNI.....	70
Tabela 10 Macierz błędów dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla dwóch porównywanych modeli: rozszerzonego oraz podstawowego (z dodatkowymi zmiennymi oraz bez) .....	71
Tabela 11 Wskaźniki jakości modelu predykcyjnego rozszerzonego (z dodatkowymi zmiennymi) dla każdej grupy z 5-krotnej walidacji krzyżowej na zbiorze ADNI. ADNI (Model rozszerzony: status chorobowy ~ RBV + MMSE + Wiek).....	73
Tabela 12 Wskaźniki jakości modelu predykcyjnego podstawowego (tylko ze zmienną niezależną MMSE, bez dodatkowych zmiennych) dla każdej grupy z 5-krotnej walidacji krzyżowej na zbiorze ADNI. ADNI (Model podstawowy: status chorobowy ~ MMSE).....	75
Tabela 13 Wskaźniki jakości (wraz z 95% przedziałem ufności) uśrednionego rozszerzonego modelu predykcyjnego z 5-krotną walidacją krzyżową na zbiorze ADNI.....	77

Tabela 14 Wartości współczynników i błędów standardowych wielomianowego modelu regresji logistycznej zbudowanego na całym zbiorze ADNI. ....	78
Tabela 15 Skorygowany iloraz szans dla zmiennych predykcyjnych w wielomianowym modelu regresji logistycznej dla całego zbioru ADNI. ....	78
Tabela 16 Wskaźniki jakości modelu predykcyjnego dla niezależnego zbioru testowego EDSD. ....	80
Tabela 17 Wskaźniki jakości modelu predykcyjnego dla podgrup 1,5T oraz 3T dla niezależnego zbioru testowego EDSD. ....	82
Tabela 18 Zgodność predykcji ze zmianą diagnozy z MCI na AD podczas okresu obserwacji dla modelu rozszerzonego oraz podstawowego z podziałem na zbiór ADNI oraz EDSD. ....	86
Tabela 19 Wyniki modelu regresji logistycznej dla zmiany statusu chorobowego z MCI do AD zbudowanego na całym zbiorze ADNI. ....	91
Tabela 20 Skorygowany iloraz szans dla zmiennych predykcyjnych w modelu regresji logistycznej dla zmiany statusu choroby z MCI do AD (ADNI). ....	91
Tabela 21 Przegląd literatury dotyczącej modeli klasyfikacyjnych zbudowanych na danych ze zbioru ADNI. ....	95

## 12. Spis rysunków

- Rysunek 1 Obraz z rezonansu magnetycznego: po lewej stronie mózg osoby zdrowej, po prawej mózg osoby z chorobą Alzheimera. Mózg osoby z chorobą Alzheimera charakteryzuje się większym pofałdowaniem kory mózgowej, powiększonymi komorami oraz zmniejszonym hipokampem. MTA (ang. medial temporal lobe atrophy): MTA = 0 oznacza brak atrofii w przyśrodkowej części płata skroniowego, MTA=4 oznacza poważną utratę objętości hipokampu. Źródło: [49]. .....23
- Rysunek 2 Wyniki obrazowania PDG-PET dla: A – zdrowej osoby, B – osoby z wczesnym stadium choroby Alzheimera, C – osoby z zaawansowaną chorobą Alzheimera. Źródło: [32] .....24
- Rysunek 3 Schemat przedstawiający kolejne kroki wstępnego przetwarzania danych prowadzące do głównej analizy. ....27
- Rysunek 4 Liczebność analizowanych zbiorów danych z baz ADNI oraz EDSD z podziałem na status diagnozy. ....30
- Rysunek 5 (a) Wiek pacjentów ze zbioru ADNI w momencie diagnozy dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (b) Wiek pacjentów ze zbioru EDSD w momencie diagnozy dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (c) Liczba lat edukacji pacjentów ze zbioru ADNI dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (d) Liczba lat edukacji pacjentów ze zbioru EDSD dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (e) Liczba punktów MMSE pacjentów ze zbioru ADNI dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (f) Liczba punktów MMSE pacjentów ze zbioru EDSD dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. (g) Odsetek kobiet i mężczyzn z analizowanego zbioru ADNI dla grup: NC, MCI, AD. Wykres kolumnowy. (h) Odsetek kobiet i mężczyzn z analizowanego zbioru EDSD dla grup: NC, MCI, AD. Wykres kolumnowy. ....33
- Rysunek 6 Mapa cieplna przedstawiająca wynik porównania obrazów MRI mózgu przed korektą niejednorodności natężenia pola i po korekcie metodą N4ITK. ....35
- Rysunek 7 a) Obraz MRI przed ekstrakcją mózgu; b) Wynik ekstrakcji mózgu za pomocą FSL-BET. ....36
- Rysunek 8 Wynik ekstrakcji mózgu za pomocą narzędzia FSL-BET. a) Zaciemnione na pomarańczowo fragmenty prezentują wynik ekstrakcji mózgu na tle oryginalnego

obrazu przed ekstrakcją. b) Zacieniowana na czerwono transparentne fragmenty prezentują wynik ekstrakcji mózgu na tle oryginalnego obrazu przed ekstrakcją, .....	36
Rysunek 9 Rezultaty ekstrakcji mózgu: a) bez korekty niejednorodności, b) po korekcie niejednorodności (kolor szary) oraz dodatkowo z zastosowaniem estymacji środka ciężkości (kolor czerwony). .....	37
Rysunek 10 Wyniki ekstrakcji mózgu za pomocą narzędzia SPM 12 (rezultat ekstrakcji zaznaczono na kolor biały oraz transparenty czerwony). .....	38
Rysunek 11 Schemat przedstawiający kolejne kroki segmentacji mózgu za pomocą dostosowanego algorytmu MiMSeg. ....	39
Rysunek 12 Rezultat dekompozycji za pomocą modelu mieszanin gaussowskich: a) model mieszanin gaussowskich z trzema komponentami dla 90-tej warstwy obrazu MRI wpisany w histogram, b) zależność pomiędzy BIC a liczbą komponentów dla 90-tej warstwy obrazu MRI, c) model mieszanin gaussowskich z czterema komponentami dla 100-tej warstwy obrazu MRI, d) zależność pomiędzy BIC a liczbą komponentów dla 100-tej warstwy obrazu MRI, e) model mieszanin gaussowskich z pięcioma komponentami dla 110-tej warstwy obrazu MRI, f) zależność pomiędzy BIC a liczbą komponentów dla 110-tej warstwy obrazu MRI. ....	41
Rysunek 13 Rezultat klasyfikacji wartości parametrów składowych GMM za pomocą algorytmu k-średnich. ....	43
Rysunek 14 Funkcje gęstości prawdopodobieństwa komponentów GMM z podziałem na dwa klastry. ....	45
Rysunek 15 Funkcje gęstości prawdopodobieństwa dla pierwszego i drugiego klastra obrazu X. ....	46
Rysunek 16 Krzywe prawdopodobieństwa warunkowego z podziałem na klastry.....	47
Rysunek 17 Rezultat segmentacji obrazu MRI na płyn mózgowo-rdzeniowy oraz tkankę mózgu za pomocą dostosowanego algorytmu MiMSeg. ....	48
Rysunek 18 Schemat uzyskania obrazu tkanki mózgu poprzez segmentację i oddzielenie płynu mózgowo-rdzeniowego od mózgu.....	50
Rysunek 19 Rezultat tworzenia maski dla całego mózgu oraz dla tkanki mózgowej uzyskanej w wyniku segmentacji po oddzieleniu płynu mózgowo-rdzeniowego.....	51

Rysunek 20 Względna objętość mózgu bez płynu mózgowo-rdzeniowego dla grup: NC, MCI i AD. Wykres typu ramka-wąsy.....	53
Rysunek 21 Warstwa obrazu MRI przedstawiająca kontur tkanki mózgowej (linia jasnoczerwono po lewej stronie) oraz tkankę mózgową (kolor czerwony po prawej stronie). .....	54
Rysunek 22 Maska binarna tkanki mózgu dla analizowanej warstwy. ....	55
Rysunek 23 Fragment macierzy reprezentującej analizowaną warstwę. Na granatowo zaznaczono komórkę reprezentującą element co, do którego podejmowana jest decyzja, czy należy do konturu tkanki mózgowej, a na niebiesko zaznaczono komórki sąsiadujące, których suma wartości daje odpowiedź na pytanie, czy środkowy element należy do konturu czy nie.....	55
Rysunek 24 Kontur tkanki mózgu dla analizowanej warstwy obrazu MRI. ....	56
Rysunek 25 Kolejne kroki procedury usuwania drobnych konturów z obrazu MRI.....	57
Rysunek 26 Rezultat procedury usuwania drobnych konturów oraz procedury wyznaczania konturu tkanki mózgu.....	58
Rysunek 27 Współczynnik pofałdowania dla grup: NC, MCI i AD. Wykres typu ramka-wąsy. ....	60
Rysunek 28 Względna objętość komór bocznych mózgu dla grup: NC, MCI i AD. Wykres typu ramka-wąsy.....	62
Rysunek 29 Krzywe ROC dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla: AD vs reszta oraz NC vs reszta. Na wykresie zamieszczono również 95% przedziały ufności dla krzywych ROC. ....	83
Rysunek 30 Krzywe ROC dla 5-krotnej walidacji krzyżowej na zbiorze ADNI dla: AD vs NC, MCI vs NC oraz AD vs MCI. Na wykresie zamieszczono również 95% przedziały ufności dla krzywych ROC.....	83
Rysunek 31 Krzywe ROC dla walidacji na zbiorze EDSD dla: AD vs reszta oraz NC vs reszta.....	84
Rysunek 32 Krzywe ROC dla walidacji na zbiorze EDSD dla: AD vs NC, MCI vs NC oraz AD vs MCI. ....	84

Rysunek 33 Zgodność predykcji ze zmianą diagnozy z MCI na AD podczas okresu obserwacji dla ADNI oraz EDSD z podziałem na model rozszerzony i podstawowy.....	87
Rysunek 34 Wykres wstępowy dla zbioru ADNI: zmiana w statusie diagnozy po 6, 12, 18, 24, 36 i 48 miesiącach. ....	88
Rysunek 35 Wykres wstępowy dla zbioru EDSD: zmiana w statusie diagnozy od diagnozy w momencie początkowym do diagnozy końcowej w 48 miesięcznym okresie obserwacji. ....	89



### 13. Spis literatury

- [1] G. B. Frisoni, N. C. Fox, C. R. Jack Jr, P. Scheltens i P. M. Thompson, „The clinical use of structural MRI in Alzheimer disease,” *Nat Rev Neurol*, tom 6, nr 2, pp. 67-77, 2010.
- [2] H. Zetterberg i B. B. Bendlin, „Biomarkers for Alzheimer's disease-preparing for a new era of disease-modifying therapies,” *Mol Psychiatry*, tom 26, nr 1, pp. 296-308, 2021.
- [3] C. Oboudiyat, H. Glazer, C. Greer, A. Seifan i R. Isaacson, „Alzheimer's disease,” *Semin Neurol.*, tom 33, nr 4, pp. 313-329, Wrzesień 2013.
- [4] „Alz.org,” [Online]. Available: <https://www.alz.org/media/Documents/alzheimers-dementia-about-alzheimers-disease-ts.pdf>. [Data uzyskania dostępu: 24 Październik 2021].
- [5] R. Taliyan, V. Kakoty, K. C. Sarathlal, S. S. Kharavtekar, C. R. Karennanavar, Y. K. Choudhary, G. Singhvi, Y. Riadi i inni, „Nanocarrier mediated drug delivery as an impeccable therapeutic approach against Alzheimer's disease,” *Journal of Controlled Release*, tom 343, pp. 528-550, 2022.
- [6] „Alzheimer's Disease Facts and Figures 2022,” tom 18, nr 4, pp. 700-789, 2022.
- [7] M. Gawęł i A. Potulska-Chromik, „Choroby neurodegeneracyjne: choroba Alzheimera i Parkinsona,” *Postępy Nauk Medycznych*, tom 28, nr 7, pp. 468-476, 2015.
- [8] R. Roberts i D. S. Knopman, „Classification and epidemiology of MCI,” *Clin Geriatr Med*, tom 29, nr 4, pp. 753-772, Listopad 2013.
- [9] J. N. Moody, K. E. Valerio, A. N. Hasselbach, S. Prieto, M. W. Logue, S. M. Hayes i inni, „Alzheimer's Disease Neuroimaging Initiative (ADNI). Body Mass Index and Polygenic Risk for Alzheimer's Disease Predict Conversion to Alzheimer's Disease,” *J Gerontol A Biol Sci Med Sci*, tom 76, nr 8, pp. 1415-1422, Lipiec 2021.
- [10] J. A. Soria Lopez, H. M. González i G. C. Léger, „Alzheimer's disease,” *Handb Clin Neurol*, tom 167, pp. 231-255, 2019.
- [11] F. Bature, B. Guinn, D. Pang i Y. Pappas, „Signs and symptoms preceding the diagnosis of Alzheimer's disease: a systematic scoping review of literature from 1937 to 2016,” *BMJ Open*, tom 7, nr e015746, 2017.
- [12] A. Atri, „The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management,” *Medical Clinics of North America*, tom 103, nr 2, pp. 263-293, 2019.

- [13] M. F. Folstein, S. E. Folstein i P. R. McHugh, „Mini-mental state". A practical method for grading the cognitive state of patients for the clinician," *J Psychiatr Res*, tom 12, nr 3, pp. 189-198, 1975.
- [14] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings i H. Chertkow, „Montreal Cognitive Assessment (MoCA)," *APA PsycTests*, 2005.
- [15] B. Hausz-Piskorz i K. Buczkowski, „Diagnostyka i leczenie choroby Alzheimera w warunkach praktyki lekarza rodzinnego," *Forum Medycyny Rodzinnej*, tom 7, nr 4, pp. 198-207, 2013.
- [16] A. M. Association, „ICD-10-CM 2019. The Complete Official Codebook," *American Medical Association*, 2018.
- [17] A. P. Association, „Diagnostic and statistical manual of mental disorders (DSM-5®)," *American Psychiatric Pub*, 2013.
- [18] B. Dubois, H. H. Feldman, C. Jacova, S. T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte i inni, „Research criteria for the diagnosis of Alzheimer's disease:," *Lancet Neurol*, tom 6, nr 8, p. 734–746, 2007.
- [19] C. Swanson, Y. Zhang, S. Dhadda i i inni, „A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-A $\beta$  protofibril antibody," *Alz Res Therapy*, tom 13, nr 80, 2021.
- [20] C. Swanson, Y. Zhang, S. Dhadda i i inni, „Correction: A randomized, double-blind, phase 2b proof-of-concept clinical trial in early Alzheimer's disease with lecanemab, an anti-A $\beta$  protofibril antibody," *Alz Res Therapy*, tom 14, nr 70, 2022.
- [21] A. H. Simonsen, S. K. Herukka, N. Andreasen, I. Baldeiras, M. Bjerke, K. Blennow i inni, „Recommendations for CSF AD biomarkers in the diagnostic," *Alzheimer's & Dementia*, tom 13, nr 3, pp. 274-284, 2017.
- [22] G. R. Grøntvedt, T. N. Schröder, S. B. Sando, L. White, G. Bråthen i C. F. Doeller, „Alzheimer's disease," *Current Biology*, tom 28, nr 11, pp. PR645-R649, 2018.
- [23] Y. N. Ou, W. Xu, J. Li i inni, „FDG-PET as an independent biomarker for Alzheimer's biological diagnosis: a longitudinal study," *Alz Res Therapy*, tom 11, nr 57, 2019.
- [24] D. Agostinho, F. Caramelo, A. P. Moreira, I. Santana, A. Abrunhosa i M. Castelo-Branco, „Combined Structural MR and Diffusion Tensor Imaging Classify the Presence of Alzheimer's Disease With the Same Performance as MR Combined With Amyloid Positron Emission Tomography: A Data Integration Approach," *Front Neurosci*, tom 15, p. 638175, 2022.
- [25] M. Dyrba, M. Hanzig, S. Altenstein, S. Bader, T. Ballarini, F. Brosseon i inni, „ADNI, AIBL, DELCODE study groups. Improving 3D convolutional neural network comprehensibility via interactive visualization of relevance maps: evaluation in Alzheimer's disease," *Alzheimers Res Ther*, tom 13, nr 1, p. 191, 2021.

- [26] J. C. Lee, S. J. Kim, S. Hong i Y. Kim, „Diagnosis of Alzheimer's disease utilizing amyloid and tau as fluid biomarkers,” *Exp Mol Med*, tom 51, nr 5, pp. 1-10, 2019.
- [27] Y. Gao, H. Huang i L. Zhang, „Predicting Alzheimer's Disease Using 3DMgNet,” *arXiv*, tom arXiv2201.04370v1 [eess.IV], 2022.
- [28] N. Goenka i S. Tiwari, „AlzVNet: A volumetric convolutional neural network for multiclass classification of Alzheimer's disease through multiple neuroimaging computational approaches,” *Biomed Signal Process Control*, tom 74, nr 103500, 2022.
- [29] X. Tang i J. Liu, „Comparing different algorithms for the course of Alzheimer's disease using machine learning,” *Ann Palliat Med*, tom 10, nr 9, pp. 9715-9724, 2021.
- [30] M. Liu, F. Li, H. Yan, K. Wang, Y. Ma, Alzheimer's Disease Neuroimaging Initiative, L. Shen i M. Xu, „A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease,” *Neuroimage*, tom 208, p. 116459, 2020.
- [31] S. H. Nozadi, S. Kadoury i The Alzheimer's Disease Neuroimaging Initiative, „Classification of Alzheimer's and MCI Patients from Semantically Parcelled PET Images: A Comparison between AV45 and FDG-PET,” *Int J Biomed Imaging*, tom 2018, p. 1247430, 2018.
- [32] L. Rice i S. Bisdas, „The diagnostic value of FDG and amyloid PET in Alzheimer's disease-A systematic review,” *Eur J Radiol*, tom 94, pp. 16-24, 2017.
- [33] E. N. Marzban, A. M. Eldeib, I. A. Yassine, Y. M. Kadah i Alzheimer's Disease Neurodegenerative Initiative, „Alzheimer's disease diagnosis from diffusion tensor images using convolutional neural networks,” *PLoS One*, tom 15, nr 3, p. e0230409, Marzec 2020.
- [34] I. Becerra-Laparra, D. Cortez-Conradis, H. G. Garcia-Lazaro, M. Martinez-Lopez i E. Roldan-Valadez, „Radial diffusivity is the best global biomarker able to discriminate healthy elders, mild cognitive impairment, and Alzheimer's disease: A diagnostic study of DTI-derived data,” *Neurol India*, tom 68, nr 2, pp. 427-434, 2020.
- [35] A. Chandra, G. Dervenoulas, M. Politis i Alzheimer's Disease Neuroimaging Initiative, „Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment,” *J Neurol*, tom 266, nr 6, pp. 1293-1302, 2019.
- [36] O. Hansson, J. Seibyl, E. Stomrud, H. Zetterberg, J. Q. Trojanowski, T. Bittner, V. Lofke, V. Corradini, i inni, Swedish BioFINDER study group i Alzheimer's Disease Neuroimaging Initiative, „CSF biomarkers of Alzheimer's disease concord with amyloid- $\beta$  PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts,” *Alzheimers Dement*, tom 14, nr 11, pp. 1470-1481, 2018.

- [37] E. Westman, J. S. Muehlboeck i A. Simmons, „Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion,” *Neuroimage*, tom 62, nr 1, 2012.
- [38] R. Wolz, V. Julkunen, J. Koikkalainen, E. Niskanen, D. P. Zhang, D. Rueckert, H. Soininen, J. Lötjönen i Alzheimer's Disease Neuroimaging Initiative, „Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease,” *PLoS One*, tom 6, nr 10, p. e25446, 2011.
- [39] R. Tadeusiewicz i J. Śmietański, „Pozyskiwanie obrazów medycznych oraz ich przetwarzanie, analiza, automatyczne rozpoznawanie i diagnostyczna interpretacja,” Kraków, Wydawnictwo Studenckiego Towzystka Naukowego, 2011.
- [40] M. Kouzuki, M. Nagano, T. Suzuki, Y. Katsumata, S. Nakamura, A. Takamura i K. Urakami, „Cerebrospinal fluid biomarkers of Alzheimer's disease are associated with carotid plaque score and hemodynamics in intra- and extra-cranial arteries on ultrasonography,” *Journal of Clinical Neuroscience*, tom 49, pp. 32-36, 2018.
- [41] E. C. Lin, „Radiation risk from medical imaging,” *Mayo Clin Proc*, tom 85, nr 12, pp. 1142-1146, 2010.
- [42] A. Burns, „Brain imaging: Computed tomography in Alzheimer's disease,” *The Lancet*, tom 341, nr 8845, pp. 601-602, 1993.
- [43] F. Pasquier, L. Bail, F. Lebert, J. P. Pruvo i H. Petit, „Determination of medial temporal lobe atrophy in early alzheimer's disease with computed tomography,” *The Lancet*, tom 343, nr 8901, p. 861, 1994.
- [44] K. A. Johnson, N. C. Fox, R. A. Sperling i W. E. Klunk, „Brain Imaging in Alzheimer Disease,” *Cold Spring Harb Perspect Med*, tom 2:a006213, 2012.
- [45] N. Borof, A. Marcinkowska, A. Sabisz i E. Szurowska, „Nowoczesne techniki rezonansu magnetycznego w obrazowaniu guzów mózgu,” *Forum Medycyny Rodzinnej*, tom 11, nr 5, p. 216–224, 2017.
- [46] N. Amoroso, A. Monaco, S. Tangaro i Alzheimer's Disease Neuroimaging Initiative, „Topological Measurements of DWI Tractography for Alzheimer's Disease Detection,” *Computational and Mathematical Methods in Medicine*, tom 17, 2017.
- [47] S. Genc, C. E. Steward, C. B. Malpas, D. Velakoulis, T. J. O'Brien i P. M. Desmond, „Short-Term White Matter Alterations in Alzheimer's Disease Characterized by Diffusion Tensor Imaging,” *J. MAGN. RESON. IMAGING*, tom 43, p. 627–634, 2016.
- [48] M. A. Oghabian, S. A. Batouli, M. Norouziyan, M. Ziaei i H. Sikaroodi, „Using functional Magnetic Resonance Imaging to differentiate between healthy aging subjects, Mild Cognitive Impairment, and Alzheimer's patients,” *J Res Med Sci*, tom 15, nr 2, pp. 84-93, 2010.

- [49] W. M. van Oostveen i E. C. de Lange, „Imaging Techniques in Alzheimer’s Disease: A Review of Applications in Early Diagnosis and Longitudinal Monitoring,” *Int. J. Mol. Sci.*, tom 22, nr 2110, 2021.
- [50] M. A. Kolanko, Z. Win, F. Loreto, N. Patel, C. Carswell, A. Gontsarova, R. J. Perry i P. A. Malhotra, „Amyloid PET imaging in clinical practice,” *Practical Neurology*, tom 20, pp. 451-462, 2020.
- [51] C. Zheng, Y. Xia, Y. Pan i J. Chen, „Automated identification of dementia using medical imaging: a survey from a pattern classification perspective,” *Brain Inform*, tom 3, nr 1, pp. 17-27, 2016.
- [52] J. B. Bae, S. Lee, W. Jung, S. Park, W. Kim, H. Oh i inni, „Identification of Alzheimer’s disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging,” *Sci Rep*, tom 10, nr 1, p. 22252, 2020.
- [53] B. Li, M. Zhang, J. Riphagen, K. Morrison Yochim, B. Li, J. Liu i D. H. Salat, „Alzheimer’s Disease Neuroimaging Initiative. Prediction of clinical and biomarker conformed Alzheimer’s disease and mild cognitive impairment from multi-feature brain structural MRI using age-correction from a large independent lifespan sample,” *Neuroimage Clin*, tom 28, p. 102387, 2020.
- [54] F. Di Stefano, S. Epelbaum, N. Coley, C. Cantet, P. J. Ousset, H. Hampel, inni i GuidAge study group, „Prediction of Alzheimer’s Disease Dementia: Data from the GuidAge Prevention Trial,” *J Alzheimers Dis*, tom 48, nr 3, pp. 793-804, 2015.
- [55] R. F. Buckley, P. Maruff, D. Ames, P. Bourgeat, R. N. Martins, C. L. Masters, S. Rainey-Smith, N. Lautenschlager, C. C. Rowe, G. Savage, V. L. Villemagne, K. A. Ellis i AIBL study, „Subjective memory decline predicts greater rates of clinical progression in preclinical Alzheimer’s disease,” *Alzheimers Dement*, tom 12, nr 7, pp. 796-804, 2016.
- [56] I. Ampuero, R. Ros, A. Royuela, V. Abaira, T. del Ser, G. García-Ribas i inni, „Risk factors for dementia of Alzheimer type and aging-associated cognitive decline in a Spanish population based sample, and in brains with pathology confirmed Alzheimer’s disease,” *J Alzheimers Dis*, tom 14, nr 2, pp. 179-191, 2008.
- [57] L. Buratti, S. Balestrini, C. Altamura, G. Viticchi, L. Falsetti, S. Luzzi i inni, „Markers for the risk of progression from mild cognitive impairment to Alzheimer’s disease,” *J Alzheimers Dis*, tom 45, nr 3, pp. 883-890, 2015.
- [58] M. Haris, S. K. Yadav, A. Rizwan, A. Singh, K. Cai, D. Kaura i inni, „T1rho MRI and CSF biomarkers in diagnosis of Alzheimer’s disease,” *Neuroimage Clin*, tom 7, pp. 598-604, 2015.
- [59] S. Iddi, D. Li, P. S. Aisen, M. S. Rafii, W. K. Thompson, M. C. Donohue i Alzheimer’s Disease Neuroimaging Initiative, „Predicting the course of Alzheimer’s progression,” *Brain Inform*, tom 6, nr 1, p. 6, 2019.

- [60] R. Zhang, L. Giancardo, D. Pena, Y. Kim, H. Tong, X. Jiang i Alzheimer's Disease Neuroimaging Initiative, „From Brain Imaging to Graph Analysis: a study on ADNI's patient cohort,” 2019.
- [61] S. F. Eskildsen, P. Coupé, D. García-Lorenzo, V. García-Lorenzo, J. C. Pruessner, D. L. Collins i Alzheimer's Disease Neuroimaging Initiative, „Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning,” *Neuroimage*, tom 65, pp. 511-521, 2013.
- [62] F. L. Seixas, B. Zadrozny, J. Laks, A. Conci i Muchaluat Saade DC, „A Bayesian network decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild cognitive impairment,” *Comput Biol Med*, tom 51, pp. 140-158, 2014.
- [63] I. Beheshti, H. Demirel, H. Matsuda i Alzheimer's Disease Neuroimaging Initiative, „Classification of Alzheimer's disease and prediction of mild cognitive impairment-to-Alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm,” *Comput Biol Med*, tom 83, pp. 109-119, 2017.
- [64] B. T. Wyman, D. J. Harvey, K. Crawford, M. A. Bernstein, O. Carmichael, P. E. Cole, i inni i Alzheimer's Disease Neuroimaging Initiative, „Standardization of analysis sets for reporting results from ADNI MRI data,” *Alzheimers Dement*, tom 9, nr 3, pp. 332-337, 2013.
- [65] K. Brueggen, M. J. Grothe, M. Dyrba, A. Fellgiebel, F. Fischer, M. Filippi i inni, „The European DTI Study on Dementia - A multicenter DTI and MRI study on Alzheimer's disease and Mild Cognitive Impairment,” *Neuroimage*, tom 144, nr Pt.B, pp. 305-308, 2017.
- [66] M. Brant-Zawadzki, G. D. Gillan i W. R. Nitz, „MP RAGE: a three-dimensional, T1-weighted, gradient-echo sequence--initial experience in the brain.,” *Radiology*, tom 182, nr 3, 1992.
- [67] J. Cohen, „Eta-Squared and Partial Eta-Squared in Fixed Factor Anova Designs,” *Educational and Psychological Measurement*, tom 33, nr 1, pp. 107-112, 1973.
- [68] P. D. Ellis, *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*, Cambridge: Cambridge University Press, 2010, p. 12.
- [69] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich i inni, „N4ITK: improved N3 bias correction,” *IEEE Trans Med Imaging*, tom 29, nr 6, pp. 1310-1320, 2010.
- [70] J. G. Sled, A. P. Zijdenbos i A. C. Evans, „A nonparametric method for automatic correction of intensity nonuniformity in MRI data,” *IEEE Trans Med Imag*, tom 17, nr 1, p. 87-97, 1998.

- [71] S. M. Smith, „Fast robust automated brain extraction,” *Hum. Brain Mapp*, tom 17, pp. 143-155, 2002.
- [72] W. D. Penny, J. Ashburner, S. Kiebel, R. Henson, D. E. Glaser, C. Phillips i inni, *Statistical Parametric Mapping: An Annotated Bibliography*. Wellcome Department of Cognitive Neurology, University College London, 2001.
- [73] F. Binczyk, B. Stjelties, C. Weber, M. Goetz, K. Maier-Hein, H. P. Meinzer i inni, „MiMSEg - an algorithm for automated detection of tumor tissue on NMR apparent diffusion coefficient maps,” *Information Sciences*, tom 384, pp. 235-248, 2017.
- [74] A. P. Dempster, N. M. Laird i D. B. Rubin, „Maximum likelihood from incomplete data via the EM algorithm,” *J. Royal Stat. Soc. Series B (methodological)*, tom 1, pp. 1- 38, 1977.
- [75] J. A. Hartigan i M. A. Wong, „Algorithm AS 136: A K-Means Clustering Algorithm,” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, tom 28, nr 1, pp. 100-108, 1979.
- [76] M. J. Kempton, T. S. Underwood, S. Brunton, F. Stylios, A. Schmechtig, U. Ettinger i inni, „A comprehensive testing protocol for MRI neuroanatomical segmentation techniques: Evaluation of a novel lateral ventricle segmentation method,” *Neuroimage*, tom 58, nr 4, pp. 1051-1059, 2011.
- [77] S. A. Czepiel, „Maximum Likelihood Estimation of Logistic,” [Online]. Available: <https://czep.net/stat/mlelr.pdf>.
- [78] A. Agresti, „On logit confidence intervals for the odds ratio with small samples,” *Biometrics*, tom 55, nr 2, pp. 597-602, 1999.
- [79] M. Stone, „Cross-Validatory Choice and Assessment of Statistical Predictions,” *J. Royal Statistical Soc. Series B*, tom 36, pp. 111-147, 1974.
- [80] E. Wit, E. Heuvel i J. W. Romeyn, „'All models are wrong...': an introduction to model uncertainty,” *Statistica Neerlandica*, tom 66, nr 3, p. 217–236, 2012.
- [81] L. García Barrado, E. Coart, T. Burzykowski i Alzheimer's Disease Neuroimaging Initiative, „Development of a diagnostic test based on multiple continuous biomarkers with an imperfect reference test,” *Stat Med*, tom 35, nr 4, pp. 595-608, 2016.
- [82] D. S. Nagin, „Analyzing developmental trajectories: A semiparametric, group-based approach,” *Psychological Methods*, tom 4, nr 2, p. 139–157, 1999.