

dr hab. n. o zdr. Emilia Sitek
Zakład Pielęgniarstwa Neurologiczno-Psychiatrycznego
Instytut Pielęgniarstwa i Położnictwa
Wydział Nauk o Zdrowiu z Instytutem Medycyny Morskiej i Tropikalnej
Gdański Uniwersytet Medyczny

R e c e n z j a

rozprawy na stopień doktora nauk medycznych

mgr Moniki Ponichtery-Kasprzykowskiej

pt. „Test fluencji słownej w diagnostyce otępienia w chorobie Alzheimera”,

przygotowanej pod kierunkiem dr hab. n.med. Jakuba Kaźmierskiego

I. Uzasadnienie ważności podjętego tematu

Próby fluencji słownej, polegające na wymienianiu (ustnie lub - rzadziej - pisemnie) jak największej liczby słów zgodnych z podanym kryterium formalnym (fonemicznym, literowym) lub treściowym (kategorialnym, semantycznym) w określonej jednostce czasu (zazwyczaj jest to 60, 90 lub 120 sekund), należą do najpopularniejszych w neuropsychologii prób klinicznych. Mają zastosowanie w wielu kontekstach klinicznych, m.in. w diagnostyce różnicowej chorób otępiennych, w tym choroby Alzheimera. Analiza profilu wykonania prób fluencji słownej w zestawieniu z oceną nazywania konfrontacyjnego może dostarczyć wglądu nie tylko w sprawność funkcji językowych, ale też pamięci semantycznej i funkcji wykonawczych. Próby fluencji słownej mogą być analizowane zarówno ilościowo (liczba słów poprawnych, intruzji i powtórzeń), jak i jakościowo (skupienia, przetłoczenia między kategoriami). Mimo częstego wykorzystywania prób fluencji słownej przez polskich klinicystów brakuje polskiej standaryzacji i normalizacji Testu Fluencji słownej, a główny polski wydawca testów psychologicznych, Pracownia Testów Psychologicznych - Practest, jest (zgodnie z informacją na stronie internetowej na dzień 20.03.2020) jedynie w trakcie opracowywania Testu Płynności Ekspresyjnej, który to z uwagi na odmienną procedurę i przedmiot badania, niestety nie będzie stanowił ekwiwalentu anglojęzycznych testów fluencji słownej jak np. COWAT- ang. *Controlled Oral Word Association Test*.

Próby fluencji fonemicznej wymagają doboru liter stanowiących optymalne kryteria przeszukiwania zasobów leksykalnych dla danego języka. Zazwyczaj w Polsce stosuje się litery 'K', 'S' oraz 'P', które cechują się wysoką częstością użycia. W niektórych polskojęzycznych wersjach testów, np. Montrealskiej Skali Oceny Funkcji Poznawczych, zdecydowano się zachować literę 'F' z wersji oryginalnej (anglojęzycznej) narzędzia, co jest wyborem kontrowersyjnym, z uwagi na inną częstość występowania tej litery w języku polskim niż w języku angielskim. Doktorantka podjęła się porównania przydatności klinicznej prób fluencji fonemicznej z wykorzystaniem różnych liter. Temat ten jest bardzo aktualny z punktu widzenia polskiego klinicysty, który nie dysponuje znormalizowanym w populacji polskojęzycznej Testem Fluencji Słownej. Stąd też dobór tematu rozprawy uważam za właściwy i ważny dla polskiej neuropsychologii zarówno w wymiarze naukowym jak i praktycznym.

II. Ocena pracy pod względem merytorycznym

1. Ocena części teoretycznej

W części teoretycznej, zgodnie z oczekiwaniem Recenzenta, Doktorantka omawia zarówno ogólne zastosowanie testów fluencji słownej jak i zastosowanie tego testu w diagnostyce osób z podejrzeniem lub rozpoznaniem ch. Alzheimera. Kluczowe z punktu widzenia teoretycznego wstępu do badań empirycznych treści zajmują jedynie niecałe 2 strony (2.4.1 Fluencja słowna w chorobie Alzheimera, s. 55-57). Jest to zdecydowanie zbyt mało, aby wyczerpująco omówić piśmiennictwo na ten temat w sposób stanowiący solidną podstawę do wysunięcia następnie hipotez pracy. Część teoretyczna rozprawy jest nadmiernie rozbudowana względem części empirycznej. Wstęp teoretyczny zajmuje ok. 50 stron, zaś dyskusja wyników jedynie ok. 20 stron. Doktorantka rekompensuje jednak te braki w podrozdziale Dyskusja, który jednoznacznie wskazuje, iż wykonała Ona wyczerpujący przegląd

piśmiennictwa zarówno anglojęzycznego jak i polskojęzycznego, zarówno jeśli chodzi o prace poglądowe i metaanalizy, jak i o prace oryginalne związane z tematem badań własnych.

Bardzo skrótowne omówienie podstawowego dla pracy zagadnienia, tj. zastosowania testu fluencji słownej w ch. Alzheimerera, może wynikać z tego, iż Doktorantka omówiła we Wstępie wiele zagadnień, które nie są kluczowe dla tematu rozprawy i mają raczej charakter uzupełniający - np. Badania neuroobrazowe (rozdz. 1.2.2), inne niż ch. Alzheimerera przyczyny otępienia (rozdz. 1.3.2 Otępienie z ciałami Lewy'ego, rozdz. 1.3.3 Otępienie naczyniopochodne, rozdz. 1.3.4 Inne rodzaje otępienia, rozdz. 1.3.5. Otępienie mieszane i atypowe). Przedmiotem badań Doktorantki była przydatność testu fluencji słownej wyłącznie w diagnostyce otępienia w chorobie Alzheimerera, stąd też omawianie nasilenia czy wzorca zaburzeń fluencji słownej w innych chorobach nie było konieczne.

W części Wstępu poświęconej otępieniu i otępieniu w chorobie Alzheimerera zabrakło krótkiego, przyjaznego dla czytelnika, np. tabelarycznego, porównania kryteriów otępienia wg. ICD-10, DSM-IV, DSM-V oraz NIA/AA. Każdy zestaw kryteriów jest omówiony osobno. Ponadto, w omówieniu kryteriów diagnostycznych choroby Alzheimerera nie znalazłam tzw. kryteriów naukowych IWG-2 przedstawionych w pracy autorstwa Dubois i wsp. (*The Lancet Neurology*, 2014).

Przedstawiając we wstępie proces diagnostyczny (s.15-18), Doktorantka nakreśliła schemat postępowania diagnostycznego od rozpoznania otępienia do rozpoznania jednostki chorobowej i opisała go jako zgodny z obowiązującymi standardami postępowania. Według mojej wiedzy zdecydowanie jest to schemat nie tylko uniwersalnie stosowany w diagnostyce różnicowej zespołów otępiennych, ale możliwy do zastosowania i adekwatny w przypadku niektórych sytuacji klinicznych. Rozpoznanie jednostki chorobowej może bowiem poprzedzać rozpoznanie otępienia, tak jak np. w chorobie Parkinsona. Następnie, Doktorantka cytuje schemat diagnostyki różnicowej otępienia, który nie obejmuje badań neuroobrazowych, elektrofizjologicznych, neurochemicznych ani laboratoryjnych (str. 15). Świadczy to o pewnych trudnościach z selekcją piśmiennictwa poświęconego medycznej diagnostyce chorób otępiennych. Dalej z kolei, co nie pasuje do wcześniejszego wyboru schematu diagnostycznego, Autorka omawia w osobnych sekcjach zarówno badania laboratoryjne jak i neuroobrazowe. Takie szczegółowe omówienie, w pracy dotyczącej funkcji poznawczych, nie jest konieczne i nie jest wolne od nieścisłości. Omawiając badania merytorycznych np. echokardiografia jest wymieniona wśród badań laboratoryjnych (str. 17). Omawiając badania neuroobrazowe Doktorantka wskazuje, iż umożliwiają one ocenę metabolizmu i czynności kory mózgowej, co może sugerować, iż nie pozwalają one na ocenę funkcji struktur podkorowych. Wskazuje ponadto, iż badania obrazowe mogą „służyć wczesnemu wykrywaniu zaburzeń poznawczych”, podczas gdy te badania mogą w istocie służyć wczesnemu wykrywaniu zmian neurozwyrodnieniowych czy chorób otępiennych, ale nie wczesnemu wykrywaniu zaburzeń funkcji poznawczych *per se*, które można stwierdzić jedynie wykonując ocenę funkcjonowania poznawczego. Następnie, Autorka podaje, iż badanie CT „pozwała na wykluczenie organicznego podłoża funkcji poznawczych”. Brakuje to uzupełniającego stwierdzenia, iż chodzi o wykluczenie podłoża innego niż proces pierwotnie zwyrodnieniowy. Użyteczność badania MRI w diagnostyce różnicowej otępienia Autorka definiuje jako technikę wykorzystywaną głównie w razie wątpliwości diagnostycznych dotyczących patologicznych struktur o nieustalonym znaczeniu lub anomalii naczyniowej, podczas gdy badanie MRI jest - według mojej wiedzy - w tej sytuacji klinicznej również badaniem pozwalającym na określenie (jakościowe lub ilościowe) wzorca zaniku tkanki mózgowej. Doktorantka, psycholog z wykształcenia, mogła wykazać się nieco większą ostrożnością w raportowaniu informacji z literatury medycznej.

Nie jest jasne dlaczego Doktorantka omawia kryteria diagnostyczne MCI (str. 19), gdyż w obrębie grupy klinicznej nie wyodrębniano osób z MCI. W tej części tekstu pojawia się pewna niespójność pod względem terminologicznym. Najpierw Doktorantka, bardzo słusznie, pisze o ryzyku konwersji do otępienia, a następnie cytując badania dotyczące tego procesu pisze, iż u „32% osób z MCI w ciągu pięciu lat rozwija się ch. Alzheimerera”. Cytowana praca dotyczy konwersji do otępienia w ch. Alzheimerera. Osoby, u których stwierdzono konwersję do otępienia w ch. Alzheimerera, od początku właśnie na tę chorobę chorowały.

W opisie zakresu badania neuropsychologicznego- który z racji wąskiego, jasno sprecyzowanego, tematu pracy nie był niezbędny - nie znalazłam oceny praktyki, która jest kluczowym elementem każdego kompleksowego badania neuropsychologicznego w kierunku choroby otępiennej. Omawiając ocenę pamięci semantycznej (s. 33) Doktorantka omyłkowo podaje jako metody oceny: Test Żetonów, który służy do oceny kompetencji syntaktycznej (rozumienia komunikatów o różnej długości i różnym stopniu złożoności gramatycznej). Omyłkowo podano tam też podtest Podobieństwa ze Skali WAIS-R, podczas gdy w tym miejscu powinien znaleźć się podtest Wiadomości z tej samej skali.

Testami, które zostały stworzone do oceny kompetencji semantycznej są narzędzia takie jak: *Cambridge Semantic Battery*, *Sydney Language Battery*, *The Pyramids and Palm tree test*, czy test sprawdzający rozumienie czasowników i czynności o łatwo wpadającej w ucho nazwie *The Kissing and Dancing Test*. Do zakresu oceny funkcji wzrokowo-przestrzennych omyłkowo zaliczono fluencję figuralną (która wpisuje się w ocenę funkcji wykonawczych) oraz ocenę procesów zapamiętywania i odtwarzania materiału wzrokowo-przestrzennego (które są oceniane w ramach diagnostyki zdolności uczenia się). Omawiając badania przesiewowe Doktorantka omyłkowo umieściła na liście skal do oceny nasilenia otępienia Skalę Demencji Mattisa. Jest to skala psychometryczna pozwalająca na ocenę nasilenia zaburzeń poznawczych, a nie nasilenia objawów otępienia, tak jak skale CDR czy GDS, wśród których została omyłkowo umieszczona. Biorąc pod uwagę fakt, iż w wielu przypadkach Doktorantka miała trudności z przyporządkowaniem narzędzia do obszaru diagnozy, a wiele narzędzi wymaga zaangażowania nie pojedynczego procesu poznawczego, a wielu funkcji poznawczych, wystarczające dla czytelnika byłoby w tym miejscu po prostu wymienienie funkcji, które są oceniane w kompleksowym badaniu neuropsychologicznym w kierunku choroby otępiennej.

Część Wstępu poświęcona prezentacji testu fluencji słownej jako narzędzia używanego w diagnostyce neuropsychologicznej, w przeciwieństwie do części dot. diagnostyki otępienia w kontekście medycyny i neuropsychologii, jest zwarta i właściwie uporządkowana. Doktorantka wykazała się w tej części bardzo dobrym zrozumieniem wielorakich czynników, które mogą wpływać na poziom wykonania prób fluencji słownej. Oczywiście ta część jest kluczowa z punktu widzenia oceny pracy. W części, w której Doktorantka wymienia rodzaje badań fluencji słownej (s. 37) zabrakło fluencji naprzemiennej (ang. *alternating fluency*), a pojawiły się mniej popularne i przydatne z punktu widzenia praktyki klinicznej typy fluencji, np. płynność afektywna. Pisząc o doborze liter w próbach fluencji literowej na s. 38 Autorka nie dodała jednoznacznie, iż klasyfikacja do jednego z trzech poziomów trudności dotyczyła j. angielskiego. Informacja na ten temat pojawia się nieco później, co nie ułatwia czytelnikowi lektury. Z kolei szczegółowa charakterystyka wskaźników skupień i przełączeń w jakościowej ocenie fluencji słownej przedstawiona w pracy nie była konieczna, gdyż Doktorantka w swojej pracy podjęła się tylko ilościowych analiz. Wśród mechanizmów poznawczych warunkujących wykonanie zadań fluencji słownej Doktorantka na trzeciej pozycji omyłkowo umieściła pamięć epizodyczną. Z powodu zapomnienia instrukcji pacjent może oczywiście przerwać wymienianie słów lub zacząć wymieniać słowa spoza kategorii, natomiast próby fluencji słownej nie służą do pamięci epizodycznej.

Mimo wymienionych powyżej uchybień, niewątpliwym atutem części teoretycznej jest dość wyczerpująca analiza piśmiennictwa polskojęzycznego, w tym prac, które nie są dostępne w wersji elektronicznej, jak również omówienie znaczenia czynników socjodemograficznych dla poziomu wykonania prób fluencji słownej.

2. Ocena projektu badania

Celem badania była ocena przydatności różnych wariantów testu fluencji słownej w rozpoznawaniu otępienia w chorobie Alzheimera. Z punktu widzenia celu projektu, który ma być projektem o wartości aplikacyjnej dla szerokiego grona klinicystów - neuropsychologów, neurologów, psychiatrów i geriatrów, przydatne byłoby dookreślenie grupy z rozpoznaniem otępieniem i zawężenie jej do osób z otępieniem w stadium łagodnym. Jak wykazała analiza zebranych przez Doktorantkę wyników, aż 55% z grupy klinicznej miało zgodnie z wynikami skali CDR otępienie umiarkowane, a 10% - otępienie głębokie. W praktyce klinicznej próby fluencji słownej są używane głównie w badaniach pacjentów z subiektywnymi skargami na funkcjonowanie poznawcze, pacjentów z łagodnymi zaburzeniami poznawczymi i otępieniem w stadium łagodnym. Stadium otępienia nie określa się z użyciem prób fluencji słownej. Ponadto, celem badania neuropsychologicznego nie jest w praktyce klinicznej różnicowanie osób w normie od osób z otępieniem umiarkowanym czy też głębokim. Większość, jeśli nie wszystkie testy neuropsychologiczne, mogą doskonale różnicować osoby z normie od osób z otępieniem w stadiach umiarkowanym i głębokim. Realny problem kliniczny stanowi dobór punktów odcięcia przydatnych w różnicowaniu osób w normie od osób z MCI oraz w różnicowaniu osób z MCI od osób z otępieniem. Stąd też, uważam iż już na etapie doboru badanej grupy ograniczone zostały potencjalne możliwości wykorzystania wyników w praktyce klinicznej.

Przystępując do formułowania celu i hipotez pracy Doktorantka twierdzi, iż „W Polsce powszechnie wykorzystywany jest zarówno anglojęzyczny zestaw FAS, jak i alternatywne zestawy liter bądź też pojedyncze litery.” Jako praktykujący neuropsycholog żywię nadzieję, iż wykorzystywanie anglojęzycznego zestawu liter FAS nie jest powszechne. Brakuje w tym miejscu podstaw empirycznych wspierających twierdzenie Doktorantki, iż stosowanie

zestawu anglojęzycznego stanowi częstą praktykę w naszym kraju. W pracy poglądowej, którą Doktorantka opublikowała w roku 2014 (Ponichtera-Kasprzykowska M., Sobów T. Adaptacja i wykorzystanie testu fluencji słownej na świecie / Adaptation and usage of the verbal fluency test in the world. *Psychiatr. Psychol. Klin.* 2014, 14 (3), p. 178–187) znajdują się cytowania wybranych polskich prac z zastosowaniem zestawu 'FAS' i Doktorantka wraca do tego również w Dyskusji. Szkoda, że zabrakło tych informacji we Wstępie.

Cel pracy został określony w sposób jasny. Hipotezy 1-3 zostały sformułowane poprawnie. Jeśli chodzi o sformułowanie dalszych hipotez badawczych, to w hipotezach 4-7 zabrakło mi wskazania kierunku oczekiwanej zależności np. Hipoteza 5 brzmi: Istnieje zależność między wiekiem badanych, a wynikami testu fluencji słownej. W świetle danych z piśmiennictwa dot. czynników socjodemograficznych omówionych wyczerpująco we Wstępie pracy hipoteza kierunkowa mogłaby brzmieć np. Starszy wiek wiąże się z gorszym poziomem wykonania testu fluencji słownej.

W projekcie badania nie uwzględniono efektu wprawy oraz wpływu kolejności zadania na jego wynik. Z racji tego, iż projekt badania obejmował 11 prób fluencji fonemicznej, próby te mogłyby albo być wykonywane w losowej kolejności z każdym uczestnikiem albo też każda podgrupa mogłaby być podzielona na kolejne podgrupy z inną kolejnością 3 zestawów literowych. Brak zastosowania kontroli czynnika kolejności wykonania danej próby / zestawu prób, ogranicza to możliwość wykorzystania uzyskanych przez Doktorantkę wyników w praktyce klinicznej.

3. Ocena doboru badanych grup

Niewątpliwym atutem pracy jest wykorzystanie dwóch grup kontrolnych (starszej i młodszej). Takie postępowanie jest w pełni uzasadnione i bardzo cenne z punktu widzenia zupełnie nowej w warunkach polskojęzycznej analizy. W dotychczasowych badaniach porównywano - w sposób zaplanowany - według mojej wiedzy w obrębie fluencji fonemicznej próby z wykorzystaniem liter „K” i „F” w grupach osób z zaburzeniami psychicznymi (prace Daniluk i wsp.).

Kryteria doboru osób z otępieniem i bez otępienia w przedziale wiekowym 65-90 lat są bardzo ogólne. Analiza wykresu obrazującego wyniki skali CDR (s. 85) wskazuje, iż w grupie osób z rozpoznaniem otępieniem znalazło się 3,3% osób z wynikiem CDR=0,5. Można było rozważyć wykluczenie tych osób z dalszych analiz, gdyż wcześniejsze rozpoznanie otępienia mogło wymagać weryfikacji. Podobnie, jak wskazano w ocenie projektu pracy wątpliwe, w kontekście potencjalnej wartości aplikacyjnej badań, jest włączenie do badania osób z otępieniem w stadium głębokim i umiarkowanym.

W charakterystyce grupy klinicznej nie znalazłam informacji o stanie zdrowotnym pacjentów poza osobami z rozpoznaniem otępienia w chorobie Alzheimera. W grupie osób w wieku 65-90 lat trudno znaleźć osoby bez schorzeń przewlekłych. Wśród kryteriów wyłączenia mogłyby znaleźć się chociażby inne choroby ośrodkowego układu nerwowego. Jeśli chodzi o różne choroby somatyczne, częste w tej grupie wiekowej to w grupie kontrolnej można było dopasować częstość głównych chorób przewlekłych do grupy klinicznej. Taki sposób doboru grupy kontrolnej był wykorzystywany na gruncie polskim m.in. przez Harciarkę w badaniach dot. funkcjonowania poznawczego osób z niewydolnością nerek.

Grupy kontrolne nie zostały dobrane do grupy klinicznej pod względem płci. Argument Doktorantki, iż rozkład płci w grupie klinicznej oddaje realną strukturę socjodemograficzną grupy pacjentów będących pensjonariuszami domów pomocy społecznej, nie uzasadnia braku doboru grup kontrolnych do grupy klinicznej pod tym istotnym względem. Płeć może mieć znaczenie dla wyników prób fluencji fonemicznej. Niestety, grupy badane nie zostały dobrane również pod względem wykształcenia. Nie znalazłam jednoznacznych dowodów na to, iż grupa z rozpoznaniem otępieniem (65-90 lat) nie różni się istotnie pod względem rozkładu wieku od grupy bez zdiagnozowanych zaburzeń (65-90 lat). Doktorantka podaje jedynie zbiorczy wynik analizy ANOVA obejmujący porównanie 3 grup, bez wyników analiz *post hoc*. Istotne byłoby udokumentowanie, iż grupy z otępieniem i bez otępienia w przedziale wiekowym 65-90 lat nie różniły się istotnie statystycznie pod względem wieku. Wzrokowa analiza wykresu ze s. 74 obrazująca rozkład wieku w obrębie podgrup kobiet i mężczyzn może sugerować większe zróżnicowanie wiekowe w grupach kontrolnych niż w grupie klinicznej.

Biorąc pod uwagę fakt, iż grupa kliniczna była najprawdopodobniej trudniej dostępna, można było dobrać względem tej grupy osoby kontrolne pod względem płci, wieku i wykształcenia. Dopiero taki dobór umożliwia podjęcie dalszych analiz i ich interpretację. Co prawda w sekcji Wyniki opisano wykorzystanie czynników

socjodemograficznych jako kontrolowanych w analizach wieloczynnikowych, jednak już samo wykorzystanie tych metod analizy w tak zróżnicowanych grupach budzi pewien niepokój czytelnika.

4. Ocena opisu metod badania

Opis metod badania jest zbyt skrótowy. Rozumiem, że mogło to wynikać z chęci skrócenia tej części metodyki neuropsychologicznej, która jest istotna z punktu widzenia rozprawy z zakresu psychologii, a już w mniejszym stopniu z punktu widzenia rozprawy z zakresu medycyny. Można było natomiast poszerzony opis procedury i kryteriów oceny umieścić w Aneksie.

W zakresie oceny fluencji fonemicznej nie jest dla mnie jasne dlaczego omawiając drugi etap badania na s. 62, jak również w streszczeniu, Doktorantka pisze o trzech zestawach liter: 'FAS', 'KPM', 'WZL', podczas gdy późniejsze analizy statystyczne odwołują się do pojedynczych liter, a nie ich zestawów. Ponadto, Doktorantka pisze, iż „litery z zestawu WZL, pojawiają się rzadziej w literaturze przedmiotu”, mając na myśli literaturę polskojęzyczną, brakuje tu jednak odnośnika bibliograficznego.

Instrukcja stosowana przez Doktorantkę w próbach fluencji fonemicznej (s. 199) odbiega od standardowej instrukcji stosowanej w krajach anglojęzycznych i tym samym utrudnia porównanie uzyskanych wyników z danymi z literatury anglojęzycznej. W instrukcji Doktorantka zawężyła zakres możliwych do wymienienia słów do rzeczowników. W próbach fluencji fonemicznej, zarówno FAS, jak i innych stosowanych w krajach anglojęzycznych nie stosuje się takiego zawężenia. Zastosowanie ograniczenia słów do rzeczowników może powodować zmianę wzorca wyników, w przypadku niektórych liter większa frekwencja wyrazów może dotyczyć czasowników i przymiotników. Oczywiście w przypadku braku ograniczenia słów do rzeczowników niezbędne byłoby poszerzenie instrukcji o informację, iż pacjent nie może wymieniać różnych form gramatycznych tego samego wyrazu (np. bawić, bawi), czy też nazw liczb (przy literach, w przypadku których istnieje takie ryzyko w j. polskim, np. P, S). Przydatne w instrukcji byłoby też podanie jednoznacznych przykładów słów niepoprawnych np. Barbara, czy Białystok. Przykłady powinny oczywiście odwoływać się do innej litery niż litera bodźcowa. Tego rodzaju zabieg jest stosowany m.in. w próbach fluencji z baterii D-KEFS.

Doktorantka nie wyjaśnia przesłanek, które skłoniły ją do selekcji kryteriów do prób fluencji semantycznej. Pomijając tradycyjne użycie tej kategorii w badaniach polskich w ciągu ostatnich kilkudziesięciu lat, Badaczka zdecydowała się tylko na dwie kategorie semantyczne i to takie kategorie, które nie mogą być traktowane jako równoważne już na etapie hipotez. Kategoria „zwierzęta” jest najpopularniejszą na świecie kategorią w ramach oceny fluencji semantycznej i jej użycie jest zasadne. Natomiast wykorzystanie kategorii „przedmioty ostre” jest wątpliwe z kilku powodów. Nie spotkałam się z badaniami naukowymi czy zaleceniami eksperckimi, które rekomendowałyby zastosowanie tej kategorii w diagnostyce różnicowej otępienia. Przydatność tej kategorii budzi kontrowersje nie tylko z tego powodu. Po pierwsze, w zależności od sposobu oceny może to być kategoria wąska (jako taką kwalifikuje ją np. Piskunowicz ze wsp. w pracy cytowanej przez Doktorantkę). W tym ujęciu przedmiotami ostrymi będą np. nożyczki, igła, czy scyzoryk. Kategoria „przedmioty ostre” może być jednak również traktowana jako kryterium ponadkategorialne, jeśli będzie zdefiniowane, np. jako wszystkie rzeczy, o które można się skaleczyć. Wtedy jako elementy poprawne byłyby zaliczone słowa takie jak np. „blacha”, „szkło”, czy „kolczyk”. Podczas gdy w przypadku liter weryfikację ich równoważności można uznać za zasadną, to zastosowane kryteria semantyczne już *a priori* równoważne nie są. Jako kategorie potencjalnie równoważne w piśmiennictwie podawane są np. imiona męskie / imiona żeńskie, warzywa / owoce, czy też ubrania / zawody. W części metodycznej pracy zabrakło informacji o sposobie oceniania próby „przedmioty ostre” wraz z przykładami odpowiedzi uznawanych za poprawne i błędne, co uniemożliwia replikację badania. W pracy nie znalazłam również kryteriów oceny fluencji semantycznej w kategorii zwierzęta. Zwyczajowo, w opisie protokołu badania znajduje się informacja odnośnie tego jaką punktację stosowano w przypadku wymieniania przez osoby badaną: 1) terminów nadrzędnych i podrzędnych (np. ryba- karp; 1 czy 2 pkt.?), 2) samca i samicy tego samego gatunku (np. kura, kogut), 3) zwierząt na różnych etapach rozwoju (np. pies, szczeniak), 4) zwierząt prehistorycznych (np. dinozaur), 5) zwierząt mitologicznych (np. jednorożec) i wreszcie 6) ras (np. pies, jamnik). Podobnie jak w przypadku próby „przedmioty ostre” utrudnia to ew. replikację badania. Nie jest jasne dlaczego Doktorantka zrezygnowała z analizy jakościowej typów błędów, tj. wtrąceń słów spoza kategorii i powtórzeń. Byłaby ona cenna w dwóch grupach osób starszych, w których się najprawdopodobniej pojawiały.

5. Ocena sposobu przeprowadzenia analiz, prezentacji wyników i zasadności wniosków

Zastosowane metody porównań międzygrupowych z kontrolą czynników socjodemograficznych oraz analiza porównawcza krzywych ROC odpowiadają celowi projektu oraz przyjętym hipotezom badawczym. W opisie metod statystycznych zabrakło natomiast informacji o zastosowanych w przypadku >2 grup metod porównań międzygrupowych na etapie analiz *post hoc*. Dotyczy to zarówno analizy wariancji jak i testu H Kruskalla-Wallisa. Nie podano też informacji o ilościowych sposobach porównania krzywych ROC. W opisie metod statystycznych brakuje testu korelacji liniowej Pearsona, który jest wymieniony w sekcji Wyniki (s.91).

W sekcji Wyniki nie podano wyników analiz *post hoc* w przypadku prób fluencji fonemicznej. Wyniki przedstawiono zbiorczo „ANOVA; wszędzie $p < 0,001$. Uczestnicy badania w wieku 65-9 lat ze zdiagnozowanym otępieniem osiągnęli, co nie budzi zaskoczenia, najniższe wyniki” (s.98). Również w tabelach i na wykresach umieszczonych na kolejnych stronach pracy nie znalazłam wyników analiz *post hoc* (s. 100-121). Nie jestem przekonana co do uzasadnionego wykorzystania w analizach międzygrupowych analizy wariancji. Czy rzeczywiście rozkład zmiennych i homogeniczność wariancji na to pozwalały w przypadku każdej z analizowanych liter? Wzrokowa analiza wykresów i tabel może wzbudzać wątpliwości czytelnika w tej kwestii.

Ponadto, z punktu widzenia celów projektu zaskakujący jest brak analizy statystycznej wyników umieszczonych w tabelach na s. 122 i 123, tj. porównań pomiędzy wynikami poszczególnych prób fluencji fonemicznej w każdej podgrupie osób badanych oraz porównań zbiorczych wyników trzech zestawów literowych (FAS, KPM i WZL). Biorąc pod uwagę fakt, iż w badaniu miała być oceniona przydatność zestawów literowych, brakuje też późniejszego porównania krzywych ROC dla trzech zestawów literowych (ujętych sumarycznie). Nie sprecyzowano również metody porównań 11 krzywych umieszczonych na ryc. 43. Nie jest jasne do jakiej statystyki testowej odnosi się $p = 0,500$ (cytowane w wielu miejscach pracy, np. s.143, 146, 150), a to wydaje się jednym z najważniejszych dla projektu wyników. Połączenie na w/w wykresie prób fluencji fonemicznej (9) i semantycznej (2) nie znajduje w mojej ocenie uzasadnienia merytorycznego z perspektywy neuropsychologicznej. Mam też pewne wątpliwości, czy dostępne ilościowe metody porównania krzywych ROC pozwalają na porównanie aż 11 krzywych ROC jednocześnie. Osobiście nie spotkałam się z dedykowanymi dla takiej sytuacji testami *post hoc*, pozwalającymi - przy stwierdzeniu różnic - ocenić, które 2 krzywe rzeczywiście różnią się między sobą pod względem AUC.

W tabelach z wynikami statystyk opisowych brakuje legendy odnośnie terminów statystycznych (np. SE). W Tabeli XVIII niektóre komórki wyróżniono pogrubioną czcionką, ale nie znalazłam dla tego wyróżnienia objaśnienia. Każda tabela powinna być czytelna bez potrzeby poszukiwania dodatkowych objaśnień w tekście.

Wykresy, choć z racji zastosowanych kolorów i wielkości, sprawiają na pierwszy rzut oka wrażenie bardzo czytelnych i przyjaznych dla czytelnika (prawie każdy wykres jest prezentowany na oddzielnej stronie), niestety przy bardziej wnikliwej analizie ujawniają braki. W pracy zostały zastosowane głównie 3 typy wykresów: wykres „ramka-wąsy”, wykres liniowy, wykres krzywej ROC. Większość wyników porównań międzygrupowych jest przedstawiona w formie wykresów typu ramka-wąsy (s.74, s.77, s.84, s.101, s.103, s.105, s.107, s.109, s.111, s.113, s.115, s.117, s.119, s.121). W przypadku każdego z tych wykresów brakuje legendy dot. ramki, wąsów i definiowania obserwacji odstających. Ten brak jest istotny z punktu widzenia analizy wyników, gdyż w zależności od tego, co przedstawiają wąsy np. w przypadków wąsów odpowiadających $\pm 2SD$ można by oczekiwać, iż wąsy obrazujące wyniki grupy z otępieniem i starszej grupy kontrolnej bez otępienia nie będą się istotnie zająbiać dla prób, które mają być przydatne klinicznie w diagnostyce różnicowej.

Wzrokowa analiza tabel z wynikami fluencji fonemicznej w poszczególnych grupach wskazuje na to, iż dla liter ‘F’, ‘A’, ‘S’, ‘M’, ‘P’, ‘W’, ‘Z’ i ‘L’ obserwowano efekt podłogi (wynik min. 0) w grupie osób z rozpoznaniem otępieniem, ale już nie dla litery „K” (najniższy wynik w tej grupie to 1pkt.). Nie można wykluczyć, iż gdyby w grupie klinicznej nie było osób z otępieniem głębokim i umiarkowanym, uzyskane wyniki mogłyby wyraźniej ukazać różnice pomiędzy zastosowanymi w badaniu literami. Kluczowe z punktu widzenia tematu pracy są wartości podane w tabeli XVIII, które nie są omówione szczegółowo w tekście. Analiza tabeli wskazuje, iż jedynie w przypadku liter ‘K’ oraz ‘M’ czułość oraz PPV wynosiła 100%. W analizie porównawczej opartej na krzywych ROC oczekiwałabym: wskazania kryterium Akaike (AIC), wskazania przedziałów ufności dla wartości każdego AUC oraz wyników testów statystycznych porównujących AUC, które można wykonać np. w ramach dodatków do oprogramowania STATISTICA lub programu MedCalc. Analiza typu „*non-inferiority*” mogłaby np. wykazać, czy wykonanie prób fluencji z użyciem zestawu ‘WZL’ nie różnicuje grup słabiej od wykonania prób fluencji z użyciem zestawu ‘KMP’. Dodatkowym atutem byłoby przeprowadzenie analogicznych analiz ze zróżnicowanymi kosztami błędnych klasyfikacji (tj. innymi kosztami klasyfikacji zdrowych jako

osób z otępieniem i osób z otępieniem jako zdrowych). Uzyskane przez Doktorantkę wyniki mogą stanowić punkt wyjścia do dodatkowych analiz.

Hipotezy pracy nie uzasadniają bezpośrednio prezentacji wyników Testu Rysowania Zegara, gdyż nie są one związane z postawionymi hipotezami badawczymi. Zastosowanie korelacji liniowej Pearsona w przypadku wyników CDT (s.91) wydaje się dyskusyjne, gdyż zmienna ma charakter porządkowy.

Przedstawienie wyników Testu Rysowania Zegara (s.79), Kwestionariusza CDR (s.82, s.85, s.86, s.87, s.88, s.89, s.90) poprzez trzy krzywe na wykresie liniowym może być mylące dla czytelnika. Wykres określa częstości występowania wyników punktowych i zamiast łączenia tego rodzaju punktów w krzywą można było np. zastosować wykres kolumnowy.

6. Ocena dyskusji uzyskanych wyników

Dyskusja jest bardzo zwięzła w porównaniu ze wstępem teoretycznym, natomiast w zakresie odwołań do literatury przedmiotu uznaję ją za wyczerpującą. W tej części odwołania do piśmiennictwa zarówno anglojęzycznego jak i polskojęzycznego są adekwatne. Ponadto, warto zwrócić uwagę, że Doktorantka w sposób właściwy selekcjonuje treści z piśmiennictwa, które stanowią dobry punkt odniesienia dla dyskusji własnych wyników. W toku Dyskusji widać wyraźnie, iż (co sugerowałam wcześniej) zbiorcza analiza wyników prób fluencji fonemicznej i semantycznej nie jest optymalna dla prowadzenia dyskursu naukowego w neuropsychologii.

Nie wszystkie wnioski przedstawione przez Doktorantkę w Dyskusji wydają się w pełni uzasadnione w świetle wyników analiz, które zostały opisane w rozprawie. W mojej ocenie przedstawione w rozprawie analizy nie stanowią podstaw do jednoznacznie pozytywnej weryfikacji Hipotezy 1 dot. równoważności wariantów w populacji osób zdrowych. Doktorantka w dyskusji (s. 143) odwołuje się do ryc. 43, na której to zestawiono krzywe ROC dot. różnicowania osób z otępieniem i bez otępienia w grupie 65-90 lat. Hipotezę 1 dot. równoważności wariantów w populacji osób zdrowych można by zweryfikować jedynie poprzez porównanie wyników tych 11 prób w grupach osób zdrowych (starszej i młodszej). Przy czym w tym kontekście takie porównanie byłoby zasadne jedynie dla kryterium literowego, gdyż kryteria semantyczne są odrębne i włączanie ich do wspólnych analiz nie wydaje się uzasadnione w pracy dotyczącej oceny funkcji poznawczych.

Cześć pierwszego zdania dot. weryfikacji Hipotezy 2 (s.146) jest sformułowana w sposób nie dopasowany względem zastosowanych analiz „Uzyskane wyniki ujawniły całkowity brak istotnych statystycznie zależności pomiędzy poszczególnymi składowymi testu fluencji słownej ($p=0,500$)”. Jeśli dobrze rozumiem zacytowany wynik (nie wiadomo przy tym do jakiej statystyki testowej się odnosi, bo nie sprecyzowano tego w sekcji Wyniki) dotyczy drugiej części stwierdzenia tj. podobnej wartości diagnostycznej poszczególnych elementów. Hipoteza 2 nie mogła być weryfikowana poprzez analizę współzależności np. analizę korelacji. Jeśli wszystkie próby adekwatnie różnicowały badane grupy to raczej można by oczekiwać wysokich współczynników korelacji między nimi.

Kolejny wniosek Doktorantki dotyczący tego, iż „nie można wyodrębnić konkretnego zestawu, czy też pojedynczej próby, które byłyby bardziej użyteczne w diagnostyce osób z otępieniem powyżej 65 roku życia niż pozostałe” nie znajduje pełnego uzasadnienia w wynikach, gdyż krzywe ROC dla poszczególnych 3-literowych wariantów nie zostały w przedstawionej mi do oceny rozprawie ani porównane statystycznie, ani zaprezentowane zbiorczo w formie graficznej.

Następnie Doktorantka podaje, iż „obserwowane różnice między omawianymi kategoriami fonemicznymi nie były istotne statystycznie, zarówno wśród osób zdrowych, jak i chorych z otępieniem powyżej 65 roku życia. Nie znalazłam analizy potwierdzającej ten wniosek. W Tabeli XVI podano jedynie statystyki opisowe dla poszczególnych grup i liter. Doktorantka pisze również, że „podsumowując wyniki badań własnych, że „osoby chore uzyskały we wszystkich 11 kategoriach VFT gorsze wyniki niż ich zdrowi rówieśnicy” (s. 148). Wniosek najprawdopodobniej jest poprawny, gdyż trudno oczekiwać innych wyników w sytuacji, gdy grupę osób z otępieniem tworzyli nie tylko pacjenci z otępieniem łagodnym, ale też osoby z otępieniem umiarkowanym i głębokim. Natomiast nie znalazłam w pracy wyników analiz *post hoc* wykazujących różnice między grupami osób z otępieniem i bez otępienia. Podano jedynie wyniki ANOVA wskazujące o tym, iż były jakieś istotne różnice międzygrupowe, ale nie wiadomo, pomiędzy którymi grupami.

Doktorantka nie zawsze stara się wyjaśnić potencjalne powody rozbieżności uzyskanych wyników z danymi z piśmiennictwa, a jedynie stwierdza taką rozbieżność, np. „Jak można zauważyć, rezultaty uzyskane w badaniach własnych autorki nie są w pełni zgodne z tymi prezentowanym powyżej”, co pozostawia pewien niedosyt.

Główna hipoteza pracy, tj. Hipoteza 3 dotycząca tego że „różne warianty testu fluencji słownej odróżniają z podobną skutecznością populację osób zdrowych od chorych” została - w mojej ocenie - częściowo zweryfikowana pozytywnie. Jak wskazałam powyżej, w pracy nie porównano statystycznie wartości AUC. Uzasadniony jest wniosek, iż każda z zastosowanych w badaniu prób pozwoliła na w pełni satysfakcjonujące zróżnicowanie osób zdrowych od chorych.

Dyskusja Hipotezy 4 jest prawidłowa. Dyskusja Hipotezy 5 i 6 również nie budzą istotnych zastrzeżeń poza stwierdzeniem, że „Siła korelacji pomiędzy wiekiem, a kategorią Zwierzęta była najwyższa spośród 11 analizowanych wyników w grupie osób 30-50 lat i trzecia z kolei w przypadku osób w wieku 65-90 lat” (s.155). W metodach analiz statystycznych ani też w sekcji Wyniki nie udało mi się odnaleźć analiz weryfikujących istotność różnic w tym zakresie. Wnioski mogłyby być bardziej powściągliwe.. Dyskusja Hipotezy 7 związanej ze zmienną płeć wymagałaby moim zdaniem wskazania w tym miejscu na ograniczenie pracy związane z nierównomiernym rozkładem płci w badanych grupach.

W pierwszym zdaniu sekcji Wnioski Doktorantka wspomina o interpretacji ilościowej i jakościowej wyników. Nie znalazłam w sekcji Wyniki wyników jakościowych. W przypadku prób fluencji słownej analiza jakościowa obejmuje analizę skupień i przetęczeń lub też analizę popełnionych błędów (intruzji, powtórzeń).

Wnioski 3-6 sformułowane przez Doktorantkę są w pełni uzasadnione. Pierwszy i drugi wniosek jest sformułowany w sposób nieprecyzyjny. Nie wykluczam, iż Doktorantka mogła przeprowadzić dodatkowe analizy uzasadniające wyciągnięcie takiego wniosku, ale nie zamieściła ich w pracy (dlaczego?). Doktorantka twierdzi, iż „Wykazano, że poszczególne warianty prób fluencji słownej stosowane w Polsce (wykorzystywane indywidualnie lub w dowolnym zestawie) są równoważne wśród chorych z otępieniem oraz osób bez zaburzeń funkcji poznawczych w wieku 65-90 lat” oraz iż „Analizowane warianty odróżniają z podobną skutecznością populację osób zdrowych od chorych w wyżej wymienionym przedziale wiekowym”. Według mojej wiedzy uzasadniony zbiorczy wniosek mógłby brzmieć np. „Analizowane warianty prób fluencji fonemicznej z satysfakcjonującą skutecznością pozwalają na zróżnicowanie osób zdrowych w wieku 65-90 lat od osób z otępieniem w przypadku mieszanej grupy chorych (od stadium łagodnego po głębokie)”. Sformułowanie Doktorantki jest nieprecyzyjne, gdyż: 1) w pracy nie znalazłam analiz statystycznych weryfikujących przydatność poszczególnych zestawów prób, 2) zdolność różnicowania dwóch grup nie jest tożsama z równoważnością zadań w jednej próbie (poziom wykonania różnych prób był zróżnicowany, na co wskazują m.in. zróżnicowane punkty odcięcia w przypadku różnych prób), 3) nie przedstawiono wartości statystyk testowych porównujących wykonanie testu w 3 grupach z analizą post hoc, 4) nie znalazłam też statystyki testowej dla porównania 11 krzywych ROC, które -wydaje się - iż powinno być wykonywane bez łączenia prób fluencji fonemicznej z próbami fluencji semantycznej. Chciałam jednocześnie zauważyć, iż wnioski sformułowane w artykule przygotowanym na podstawie tej rozprawy doktorskiej są bardziej powściągliwe i zasadne w kontekście opisanych wyników analiz (Ponichtera-Kasprzykowska M., Sobów T, Kaźmierski J. Variants of the verbal fluency test in the diagnosis of dementia – does the selection of letters matter? / Warianty testu fluencji słownej w diagnostyce otępień – czy wybór liter ma znaczenie? *Psychiatr Psychol Klin* 2019, 19 (2), p. 149–157)..

Doktorantka w toku dyskusji przyznaje, iż wzorzec uzyskanych wyników może wynikać ze struktury grupy klinicznej, w obrębie której istotną podgrupę stanowiły osoby z otępieniem umiarkowanym (s. 151) oraz, że rozkłady zmiennych demograficznych w badanych grupach nie były jednorodne (s.162). Natomiast można by w tym miejscu oczekiwać wypunktowania również innych ograniczeń metodologicznych pracy.

Chciałam jednocześnie zwrócić uwagę, że część sformułowanych przeze mnie tutaj zastrzeżeń jest wtórna względem sposobu prezentacji wyników. Doktorantka w sposób adekwatny zestawia dane z badań własnych z danymi z piśmiennictwa.

7. Uwagi terminologiczno-językowe

W części empirycznej pracy Doktorantka błędnie używa terminu „grupa badana”, tak jakby grupa kontrolna nie była grupą badaną. Bardziej adekwatnym terminem w tym miejscu byłaby „grupa kliniczna”. Ponadto w dyskusji i we wnioskach zbyt często Doktorantka posługuje się terminami statystycznymi, które są zarezerwowane głównie dla sekcji Wyniki, np. we wniosku nr pisze, iż „Liczba lat edukacji korelowała dodatnio, istotnie statystycznie z większością kategorii literowych, nie wykazano natomiast zależności między wyżej wymienioną zmienną, a wariantami semantycznymi VFT” (s.161). We wnioskach należałoby raczej pisać o pozytywnej zależności między czasem edukacji

a wynikiem większości prób fluencji literowej. To jasne, iż we wnioskach oczekujemy raportowania tylko tych wyników, które były istotne statystycznie.

Na s.41 błędnie odmieniono nazwisko Sharon Abrahams, Doktorantka najprawdopodobniej uznała, że to badacz - mężczyzna. Analogiczny błąd dotyczy Angeli Troyer (s.57), Edith Kaplan (s.35-40) oraz Gisele Wolf-Klein, których nazwiska zostały odmienione w formie męskiej. Odwrotny błąd pojawił się w kontekście nazwiska Johna Hodgesa (s.152), które zostało potraktowane jako nazwisko żeńskie.

Na liście skrótów można znaleźć nieścisłości terminologiczne dot. zarówno nazw własnych jak i nazw pospolitych np.:

- Skala Funkcjonowania Poznawczego Addenbrooke'a, podczas gdy nazwa skali nie pochodzi od autora testu tylko od nazwy szpitala Addenbrooke (ten sam błąd pojawia się na s. 30)
- skrót D-KEFS wyjaśniony jako Skala Oceny Funkcji Wykonawczych Delisa i Kaplana, podczas gdy powinno być „Kaplan”, gdyż współautorką testu jest kobieta: Edith Kaplan (ten sam błąd w odmianie nazwiska autorki testu pojawia się na s. 35 oraz na s. 40)
- skrót CFT jako pochodzący od Rey-Osterrieth Complex Figure Test, podczas, gdy wtedy zasadny byłby skrót ROCFT
- APOE jako gen apolipoproteiny E, a nie sama apolipoproteina E
- fMRI jako funkcjonalny magnetyczny rezonans jądrowy, a nie obrazowanie metodą funkcjonalnego rezonansu magnetycznego
- WAIS-R, w polskojęzycznym rozwinięciu skrótu zabrakło wyjaśnienia „R” jako wersji zrewidowanej

Na str. 18 Doktorantka pisze o „tylnym zakręcie obręczy” mając na myśli najprawdopodobniej tylną część zakrętu obręczy, gdyż w każdej półkuli mózgowej jest jeden zakręt obręczy, w obrębie którego można wyróżnić poszczególne części w tym jego część tylną.

W tekście pracy, szczególnie w Dyskusji pojawiają się sformułowania niejasne pod względem formalno-treściowym np. „litery M i K, oraz w mniejszym stopniu L, okazały się istotne w odniesieniu do zmiennej edukacji”. Z powyższego zdania trudno się domyślić, że w przypadku prób z użyciem tych liter nie stwierdzono pozytywnego związku z czasem edukacji. We wniosku nr 4 sformułowanie iż „Biorąc pod uwagę liczbę lat edukacji, wyodrębniono trzy warianty fonemiczne (M, K, L) oraz jeden wariant semantyczny *Przedmioty ostre*, istotne statystycznie dla wspomnianej zmiennej” (s.161) jest niejasne. Niektóre zdania są nieprecyzyjne np. „wymienione litery mają wyższy stopień trudności w stosunku do wersji anglojęzycznej” (s. 144), tymczasem chodzi o próby z wykorzystaniem liter lub też kryteria literowe. Na s. 46 opisując wyniki badań zespołu Baldo Doktorantka pisze o „pojęciach”, podczas gdy we fluencji wg kryterium formalnego wymieniane są słowa, a nie pojęcia.

W pracy można znaleźć niezręczne językowo sformułowania, np. we Wstępie Doktorantka określa ‘guzy’ jako ‘choroby „potencjalnie odwracalne”’(str. 15). Autorka pisząc o wymienianiu słów przez chorych używa określenia „produkują przeciętnie X słów” (dwukrotnie na s. 56).

Ponadto, w pracy pojawia się przestarzały termin „płasawica Huntingtona” (str. 18, 48, 54, 57), którego odpowiednikiem w aktualnej światowej terminologii medycznej jest „choroba Huntingtona”. Przykłady innych niezręcznych sformułowań to: „trudności w rysowaniu konstrukcyjnym” zamiast „obniżone zdolności wzrokowo-konstrukcyjne” lub „trudności w rysunkowych próbach prakcji konstrukcyjnej” (s.26), niezrozumiałe sformułowanie „zaawansowane obniżenie modalności w różnych obszarach funkcjonowania (s.26). W części poświęconej występowaniu napadów padaczkowych i mioklonii w zaawansowanym stadium otępienia powinno być „i / lub mioklonie”, gdyż możliwe jest występowanie łączne napadów padaczkowych oraz mioklonii.

Cytując kryteria diagnostyczne, które mają swoje oficjalne polskie tłumaczenie, jak np. kryteria rozpoznania otępienia wg ICD-10 czy DSM-IV, wskazane było konsekwentne stosowanie jednego sposobu cytowania, tj. albo cytowania dosłownego, zgodnie z polskim tłumaczeniem, albo przytaczani ich w treści pracy dokonując przeformułowań językowych, które uczynią te kryteria bardziej zrozumiałymi. Doktorantka niekiedy cytuje kryteria wprost np. podając „prymitywizację zachowań społecznych”, co brzmi dość niezręcznie, choć taki zwrot funkcjonuje w polskim tłumaczeniu, a niekiedy dokonuje modyfikacji tłumaczenia. Doktorantka tłumacząc kategorie diagnostyczne z DSM-V podaje nazwy podtypów z wielkiej litery, idąc śladem wersji anglojęzycznej, co jednak w języku polskim stanowi błąd językowy. Dla przykładu w j. polskim „zaburzenia neuropoznawcze” powinny być pisane małą literą. Podobnie, słowa takie jak otępienie czy depresja powinny być pisane małą literą (s.64). W opisie kryteriów

oceny TRZ wg Sunderlanda zabrakło informacji, czy przedstawiono kryteria oceny w tłumaczeniu własnym, czy na podstawie jednego z dostępnych tłumaczeń (w tej sytuacji niezbędne byłoby podanie właściwego źródła).

Ponadto w pracy pojawiają się nieuzasadnione anglicyzmy, np. „klastery” zamiast „skupienia”, mimo iż Doktorantka cytuje polskojęzyczną pracę na ten temat, autorstwa Jodzio, w której wprowadzono właściwy termin skupienia. Na s. 61 Doktorantka podaje, iż w badaniu brali udział „uczestnicy uniwersytetu trzeciego wieku”, a powinno być „słuchacze Uniwersytetu Trzeciego Wieku” ze wskazaniem nazwy Uniwersytetu.

W polskojęzycznym tłumaczeniu nazwy skali ADAS-Cog. brakuje określenia „podskala poznawcza”. Omawiając wykorzystanie badania neuropsychologicznego w diagnostyce zespołu otępiennego Doktorantka używa terminu „rzetelność” (s.31) mając na myśli - jak wskazują kolejne zdania - raczej trafność diagnostyczną.

W pracy są błędy literowe, w tym w nazwiskach badaczy: „Brickman” (s.49), Raskovsky (s.57, 153) Paula McHugh (s.62), Freedman (s.64), Ardila (s.145), interpunkcyjne (s. 13), leksykalne (s.78 powinno być „niższy wynik”), jak również dotyczące pisowni rozdzielnej [„przedchorobowy”(s. 16), „radiofarmaceutyki” (s.27)]. Zdarzają się też pominięcia przyimków (np. „w” we frazie „radiofarmaceutyków w PET”(s.27), czy słów („odgrywają istotną [rolę]”, s. 43). Drobne błędy językowe znalazły się również w streszczeniu anglojęzycznym.

III. Ocena pracy pod względem formalnym


Układ pracy jest bardzo czytelny i zawiera wszystkie części, które powinna zawierać rozprawa doktorska. Część teoretyczna jest nadmiernie rozbudowana względem części empirycznej, co omówiono pod względem treściowym w punkcie dotyczącym oceny merytorycznej pracy. Na pochwałę zasługuje konsekwentna, spójna struktura pracy, w której kolejność prezentowanych hipotez odpowiada następnie kolejności omawianych w dyskusji zagadnień. Szkoda, że ta kolejność nie została zachowana również w części poświęconej wynikom. Ponadto, atutem pracy jest spis zastosowanych skrótów, który może ułatwić czytelnikowi lekturę. Pewne wątpliwości budzi jedynie wielość skrótów i zasadność ich stosowania w sytuacji, gdy dany termin pojawia się w całej pracy tylko raz bądź kilka razy, np. MRS. W pracy brakuje numeru zgody Komisji Bioetycznej, zwyczajowo taki numer jest podawany w pracy.

Starając się przytaczać - tam gdzie to możliwe - piśmiennictwo polskojęzyczne Doktorantka odwołując się do badań epidemiologicznych cytuje polskie prace Opali oraz Pfeffer i Kłoszewskiej, jednak bez dodania „[za: X]”, gdyż nie cytuje oryginalnych badań epidemiologicznych, a prace pogładowe, w których te międzynarodowe badania epidemiologiczne były opisane.

W piśmiennictwie, generalnie przygotowanym w sposób spójny, zdarzają się jednak odstępstwa od przyjętego stylu cytowania (poz. 13, 83, 94, 52, 154, 155, 185, 205, 216, 260, 287, 292, 293) oraz nieliczne błędy (poz. 236 zawierająca tylko nr DOI). Wydaje się, że w poz. 64 i 65 tytuł czasopisma jest niepełny (*Annales UMCS?*), podobnie jak w poz. 220 (*Seminars in Cerebrovascular Diseases and Stroke*). Wątpliwości budzi cytowanie prac, które (w wersji oryginalnej) zawierają błędy w tytule (poz. 291- „*chronic nervous system*” zamiast „*central nervous system*”), co nie budzi zaufania do tych materiałów źródłowych.

IV. Podsumowanie i wniosek końcowy

Ogólnie rozprawę mgr Moniki Ponichtery-Kasprzykowskiej oceniam pozytywnie. Przygotowana rozprawa dowodzi, iż Doktorantka bez wątplenia posiadała podstawowe kompetencje jakich oczekuje się od kandydatów na stopień naukowy doktora dotyczące samodzielnego rozwiązania problemu naukowego oraz ogólnej wiedzy teoretycznej w zakresie obszaru badawczego. Przeprowadzone przez Autorkę badania stanowią interesujący głos w dyskusji nt. adaptacji prób fluencji słownej do warunków polskich oraz stosowania różnych wariantów tych prób w praktyce klinicznej, który zostanie jak sądzę zauważony w polskiej neuropsychologii dzięki publikacjom na ten temat, których współautorką jest Doktorantka (2014, 2019). Na szczególną pochwałę zasługuje temat rozprawy, istotny z zarówno dla nauki jak i praktyki klinicznej. Mimo wskazanych powyżej trudności metodologicznych, przedstawiona mi do oceny rozprawa doktorska spełnia zarówno warunki ustawowe jak i wymagania akademickie stawiane rozprawom doktorskim. Stąd, wnoszę o dopuszczenie Doktorantki do dalszych etapów przewodu doktorskiego.


dr hab. n. o zdr. Emilia J. Sitek
specjalista psycholog kliniczny
neuropsycholog
nr PMR 0019