

**A head-to-head comparison of the sensitivity of two screening tests for dementia
Mini-Mental-State-Examination (MMSE) and the Test for the Early Detection of
Dementia with Discrimination from Depression (TE4D)**

**Równoległe porównanie czułości dwóch testów skriningowych w diagnostyce demencji
Mini-Mental-State-Examination (MMSE) oraz the Test for the Early Detection of
Dementia with Discrimination from Depression (TE4D)**

Ralf Ihl, Achim Biesenbach, Sarah Brieber, Brigitte Grass-Kapanke, Thomas Salamon
Department of Psychiatry University of Duesseldorf

Address for correspondence: Prof. Dr. med. Ralf Ihl
Alexianer Hospital
Kölner Str. 64
D-51149 Köln
Tel.: +49 22031271
FAX: +49 2203
E-Mail: r.ihl@alexianer-koeln.de

Key words: dementia, validity, sensitivity, specificity, comparison, MMSE, TE4D, diagnosis, test

Słowa kluczowe: demencja, trafność, czułość, swoistość, porównanie, MMSE, TE4D

Abstract

Objective: Numerous tests have been developed to detect dementia. Concerning data on validity, head to head comparisons are rare. To evaluate, which test show higher sensitivity and specificity for the screening of dementia, the Mini-Mental-Status-Examination (MMSE) and the Test for The Early Detection of Dementia with Discrimination from Depression (TE4D) were compared in a German-speaking population.

Methods: MMSE and the TE4D were assessed in 139 patients with Alzheimer's Disease (AD), 93 with depressive disorder (DEP) and in 21 healthy controls. Diagnosis was made without knowledge of the test results. The ADAS-Cog and the SKT were additionally

administered to cover the required neuropsychological deficits according to the NINCDS/ADRDA-criteria.

Results: The correlation between MMSE and TE4D was $r_s = .88$. Sensitivity for the detection of AD was .99 for the TE4D. Depending on the cut-off score of the MMSE, sensitivity was .79 (cut-off 23) or .94 (cut-off 26). Specificity against DEP was .99 for the TE4D, and for the MMSE .93 (cut-off 23) respectively .64 (cut-off 26). In comparison to a MMSE cut-off score of 23, the TE4D correctly classified more patients with AD ($N = 139$, $n = 26$, chi-square = 24,03, $p < .001$). For the MMSE cut-off score of 26, the specificity of the TE4D for DEP was significantly higher than for the MMSE ($N = 93$, $n = 34$, Fig. 1, chi-square = 28.26, $p < .001$). In the ROC comparison, the area under the curve was .99 for the TE4D and .95 for the MMSE.

Discussion: The correlation between MMSE and TE4D was high and explained 77% of variance. However, in discriminating patients with AD from patients with depression and healthy controls, sensitivity and specificity of the TE4D was superior to the MMSE.

Streszczenie:

Wstęp: Opracowano wiele testów do diagnostyki demencji. W ocenie ich trafności bezpośrednie porównania należą do rzadkości. W celu oceny, który z testów skriningowych prezentuje wyższą czułość oraz swoistość w diagnostyce demencji oceniano the Mini-Mental-Status-Examination (MMSE) oraz the Test for The Early Detection of Dementia with Discrimination from Depression (TE4D) w populacji niemieckojęzycznej.

Metody: Wyniki MMSE i TE4D oszacowano u 139 pacjentów cierpiących z powodu choroby Alzheimera (AD), 93 z rozpoznaniem zaburzenia depresyjnego (DEP) oraz zdrowych, stanowiących grupę kontrolną. Diagnozę stawiano bez znajomości wyników testów. Dodatkowo zastosowano ADAS-Cog oraz the SKT celem określenia deficytów neuropsychologicznych według kryteriów NINCDS/ADRDA.

Wyniki: Korelacja pomiędzy testami MMSE i TE4D wyniosła $r_s = .88$. Czulość w diagnostyce AD wyniosła .99 dla testu TE4D. W zależności od punktu odcięcia w teście MMSE czulość wyniosła .79 (punkt odcięcia 23) i .64 (punkt odcięcia 26). W porównaniu do testu MMSE (punkt odcięcia 23) test TE4D prawidłowo sklasyfikował większą ilość pacjentów z rozpoznaniem AD ($N = 139$, $n = 26$, chi-kwadrat = 24,03, $p < .001$). W porównaniu do testu MMSE przy punkcie odcięcia 26 swoistość testu TE4D dla DEP była znacząco wyższa w porównaniu do testu MMSE ($N = 93$, $n = 34$, Rys. 1, chi-kwadrat = 28.26,

$p < .001$). W porównaniu ROC pole poniżej wykresu wyniosło .99 dla testu TE4D i .95 dla testu MMSE.

Wnioski: Korelacja pomiędzy testami MMSE oraz TE4D była wysoka i wyrażała 77 % wariancji. Niemniej jednak w różnicowaniu pacjentów z rozpoznaniem AD od pacjentów z zaburzeniem depresyjnym i zdrowymi członkami grupy kontrolnej, czułość i swoistość testu TE4D była wyższa od testu MMSE.

Wprowadzenie

Każdego roku wzrasta liczba testów skriningowych przeznaczonych do diagnostyki demencji. Niewiele z nich poddanych zostało prawidłowej ocenie, a tylko w kilku przypadkach wykazano ich wyższość nad już istniejącymi skalami [1]. Zagadnienie trafności używanych testów utrudnia wybór odpowiedniego narzędzia w przesiewowej diagnostyce demencji. Pomimo niekorzystnych doniesień przedstawionych w kilku publikacjach [2-9], test the Mini-Mental-State-Examination [10] jest często wykorzystywany w badaniach. Do jednym z najważniejszych mankamentów skali MMSE należy jej niska czułość w diagnostyce mniej zaawansowanych postaci otępień [11-12]. Test MMSE pominie znaczną liczbę pacjentów cierpiących z powodu mniej zaawansowanych stadiów demencji. Dlatego też opracowano skalę the “Test for the Early Detection of Dementia with Discrimination from Depression” umożliwiającą diagnostykę otępień we wczesnym etapie choroby, z wysoką czułością i swoistością [13, appendix]. W odniesieniu do części poznawczej, rzetelność oceniana metodą porównywalności (inter-rater reliability) wyniosła .99, metodą powtórnego testowania (test-retest reliability) .87, metodą zgodności wewnętrznej (internal consistency) (Cronbach alfa) .88 [13]. Ostatnio przedstawiono zbieżność jej rzetelności z badaniem EEG [14]. Jednakże dotychczas nie przeprowadzono równoległego porównania obu wspomnianych skal – MMSE i TE4D. Z tego powodu autorzy zbadali, za pomocą obu testów, grupę kontrolną osób zdrowych, grupę pacjentów z rozpoznaniem depresji oraz grupę chorych cierpiących z powodu niezaawansowanej formy otępienia, kładąc nacisk na MMSE i część poznawczą TE4D.

Metody

W badaniu, po wyrażeniu świadomej zgody, wzięło udział 139 pacjentów cierpiących z powodu choroby Alzheimera (AD), 93 z rozpoznaniem zaburzenia depresyjnego (DEP) oraz 21 osób zdrowych, stanowiących grupę kontrolną (CON). Diagnoza prawdopodobnego rozpoznania demencji typu Alzheimera postawiona została zgodnie z kryteriami McKanna [15]. Deficyty neuropsychologiczne oszacowano na podstawie skal the ADAS-cog [16] oraz the SKT [17]. Zaburzenie depresyjne zdiagnozowano według wytycznych ICD-10 (F32-F33). W rozpoznawaniu depresji nie stosowano skal MMSE i TE4D. Punktacja uzyskana podczas

badania grupy kontrolnej za pomocą ADAS-cog oraz SKT winna była zawierać się w przedziale wartości prawidłowych. Pacjenci rekrutowani byli z kartotek ambulatorium. Grupę kontrolną stanowili krewni pracowników Kliniki.

Punktacja skali MMSE zawierała się w zakresie od 0 do 30, przy czym 30 stanowiło najlepszy wynik. Zgodnie z doniesieniami przedstawionymi w szeregu publikacji zaleca się określenie punktu odcięcia na poziomie 23 lub 26 punktów [7, 18]. Skala TE4D bada pamięć, orientację, przyporządkowywanie, wykonywanie poleceń, rysowanie zegara i płynność wysławiania się, na podstawie 9 pozycji, ocenianych łącznie do 50 punktów, przy czym wynik ten stanowi najwyższą i zarazem najlepszą wartość (porównaj appendix). Punkt odcięcia w diagnostyce otępień wynosił 35 i poniżej. Ponadto wprowadzono dwie 11-punktowe skale Likerta w celu oceny nasilenia zaburzenia depresyjnego. Wszystkich pacjentów oraz grupę kontrolną zbadano, podczas jednej sesji, obiema skalami – MMSE i TE4D. Kolejność wykorzystania odpowiednich narzędzi poddana została randomizacji z prawdopodobieństwem 0.5, by uniknąć efektu serii.

Korelacja między dwoma narzędziami została wyliczona za pomocą Spearman rho. Zbadanie hipotezy, iż skalę TE4D cechuje wyższa czułość w diagnostyce wczesnych etapów demencji niż test MMSE, wymagało potwierdzenia rozbieżności w przyporządkowywaniu do grup rozpoznań. W zależności od wielkości grupy wyliczono test McNemar lub Binominal. Z powodu dwóch punktów odcięcia skali MMSE powyższe wartości zostały wyliczone dwukrotnie.

W celu przedstawienia czułości i swoistości skal MMSE i TE4D wyliczono krzywe ROC. W obliczeniach, poza testami McNemar i Binominal, wykorzystano test SPSS 11.0.1. Łącznie obliczono 9 testów. Poziom znamienności określono na poziomie $p < .01$. Wyliczono korekcję Bonferroniego otrzymując poziom znamienności wynoszący $p < .001$.

Wyniki

Tabela 1. prezentuje charakterystykę pacjentów oraz grupy kontrolnej. Średnia wieku oraz poziom wykształcenia różniły się znamienne między grupami.

Korelacja pomiędzy skalami MMSE i TE4D wyniosła $r_s = 0.88$ ($N = 253$, $p < 0.001$, Rys.1 A). W odniesieniu do pacjentów cierpiących z powodu choroby Alzheimera korelacje

wyniosły $r_s = 0.78$ ($N = 139$, $p < 0.001$, Rys.1 B), chorych z zaburzeniem depresyjnym $r_s = 0.46$ ($N = 93$, $p < 0.001$, Rys.1 C), natomiast dla grupy kontrolnej $r_s = 0.48$ ($N = 21$, $p < 0.05$, n.s., Rys.1 D). Model linearny przedstawił najwyższą zgodność dla wszystkich grup z wyjątkiem zbioru osób z otępieniem, w odniesieniu do którego najwyższą zgodność wyników wykazała regresja nieliniarna.

Przy punkcie odcięcia wynoszącym 35 test TE4D poprawnie sklasyfikował 97.8 % pacjentów z chorobą Alzheimera. Gdy punkt odcięcia, dla skali MMSE, ustalono na poziomi 26, oba testy poprawnie sklasyfikowały 131 (94.22 %) z grupy pacjentów z chorobą Alzheimera. Diagnostyka przeprowadzona za pomocą skali MMSE była mylna w trzech, natomiast skali TE4D w pięciu przypadkach (test Binominal, $N = 139$, $n = 8$, n.s.). Nie stwierdzono pacjentów, w przypadku których zawiodły obie skale. Badanie pacjentów cierpiących z powodu zaburzenia depresyjnego skalą MMSE, przy punkcie odcięcia wynoszącym 26, doprowadziło do postawienia 33 fałszywie pozytywnych diagnoz demencji, dla porównania test TE4D do 1 (35.5% vs. 0.72 %, 33 vs. 1, McNemar-Test, $N = 93$, $n = 34$, chi-kwadrat 28.26, $p < .001$).

Przy punkcie odcięcia wynoszącym 23 dla testu MMSE obie skale poprawnie sklasyfikowały 110 (79.1 %) ze 139 pacjentów z chorobą Alzheimera. Oba testy niepoprawnie sklasyfikowały 3 pacjentów jako przedstawicieli grupy kontrolnej. Badanie jedynie za pomocą skali MMSE dało wynik fałszywy w 26 przypadkach, natomiast zastosowanie jedynie testu TE4D w żadnym (26 vs. 0, McNemar-Test, $N = 139$, $n = 26$, chi-kwadrat 24,03, $p < .001$). Badanie pacjentów cierpiących z powodu zaburzenia depresyjnego skalą MMSE, przy punkcie odcięcia wynoszącym 23, pozwoliło na poprawne sklasyfikowanie 87 z 93 pacjentów z zaburzeniem depresyjnym, jako nie cierpiących z powodu otępienia (93.5 %). Wynik badania przeprowadzonego jedynie za pomocą skali MMSE był niepoprawny w 4 przypadkach, natomiast zastosowanie jedynie testu TE4D dało wynik błędny w 2, klasyfikując chorych z zaburzeniem depresyjnym jako chorych z otępieniem (Binominal test, $N = 93$, $n = 6$, n.s.).

W ocenie swoistości dla grupy kontrolnej nie uchwycono znamienych różnic (1 błędna klasyfikacja przy użyciu skali MMSE, 2 skali TE4D).

Pole pod krzywą ROC wyniosło .95 dla skali MMSE, .99 dla skali TE4D (Rys. 3).

Omówienie

Korelacja pomiędzy skalami MMSE i TE4D wyniosła $r_s = 0.88$. Podczas obserwacji jedynie podgrupy pacjentów cierpiących z powodu choroby Alzheimera (AD) korelacja pozostała wysoka. Z powodu niższej wariacji w podgrupie DEP i CON korelacje okazały się w nich niższe. Przytoczone dane mogłyby prowadzić do założenia, iż nie stwierdzi się znacznej różnicy pomiędzy oboma testami. Podczas oceny swoistości obraz uległ zmianie. Wykazano, że wskaźniki swoistości skali TE4D, we wszystkich trzech grupach, wyniosły ponad 90 %. Przy punkcie odcięcia, dla skali MMSE, wynoszącym 23 jej czułość w diagnostyce AD była niska w porównaniu do skali TE4D. Zdiagnozowała ona jedynie 79 % pacjentów, w porównaniu do 99 % sklasyfikowanych za pomocą skali TE4D. Niemniej jednak w tym przypadku czułość w różnicowaniu depresji od demencji była wysoka, osiągając 93.5 %. Gdy punkt odcięcia podwyższono do 26 punktów, czułość w diagnozowaniu AD może być wyższa. Jednakże swoistość w różnicowaniu AD od DEP obniżyła się do 64.5 %. Tym samym, niezależnie od wysokości ustalonego punktu odcięcia, skala MMSE nie osiągnęła wartości wskaźników czułości i swoistości oszacowanych dla testu TE4D. Badacze nie byli w stanie określić, czy w populacyjnych badaniach skriningowych różnicowanie byłoby równie trafne. Różnice w wieku i wykształceniu były znaczące między grupami. Pacjenci z otępieniem stanowili grupę mniej wykształconą i starszą. Wydaje się jednak, że wpływ tych czynników na porównanie skal może być pominięty. Po uwzględnieniu wspomnianych ograniczeń dopuszczalne jest dalsze wnioskowanie.

Niska czułość testu MMSE została uprzednio w wykazana w porównaniach ze skalami ADAS, the SKT, oraz BCRS [19]. Udokumentowano także, że test MMSE jest nieprzydatny w ocenie przebiegu demencji [1-2]. Z obu powyżej wspomnianych powodów jego użyteczność w praktyce ogólnolekarskiej jest ograniczona [18-19]. Wszystko to wyjaśnia, że w zależności od punktu odcięcia skali MMSE ustalonego na poziomie 23 lub 26, jej interpretacja prowadzi będzie do błędnych decyzji w różnej liczbie przypadków [3-5, 7-9, 18]. W zależności od punktu odcięcia zmienia się przede wszystkim rodzaj błędów. Oznacza to, że użyteczność testu MMSE w diagnostyce wczesnych stadiów AD jest dyskusyjna [11-12].

W żadnym z dotychczasowych badań nie wyliczono porównania, wobec pozostałych krótkich testów, które rozpatrywałoby parametry czułości i swoistości w diagnostyce AD. Czułość i swoistość innych testów skriningowych, w stosunku do skali MMSE, zostały wykazane w szeregu badań. Parametry te w odniesieniu do testu TE4D oszacowano w niniejszym badaniu. Zaprezentowane wyniki wspierają założenie, iż skala TE4D posiada wyższą czułość w diagnostyce AD oraz wyższą swoistość w różnicowaniu AD od DEP oraz grupy kontrolnej, co zaprezentowano także za pomocą krzywych ROC.

Wnioskowanie może być ograniczone przez projekt badania. Liczba przedstawicieli grupy kontrolnej była relatywnie niewielka, a wyniki mogłyby ulec zmianie po jej wzroście. Dalsze ograniczenia pojawić się mogą w związku z metodą klasyfikacji pacjentów z AD, zgodną z kryteriami NINCDS/ADRDA oraz ADAS-cog i the SKT. Jednakże nie istnieją przesłanki, które wskazywałyby na możliwość innego pogrupowania pacjentów w wyniku zastosowania alternatywnych testów.

Przedstawione wyniki powinny przekonać badaczy do zaniechania używania skali MMSE w klasyfikacji AD szczególnie we wczesnych stadiach. Wykazano, że skala TE4D jest jej alternatywą, charakteryzującą się wyższą czułością i swoistością.

Piśmiennictwo

1. **Ihl R, Grass-Kapanke B, Jänner M, Weyer G.** Neuropsychometric tests in cross sectional and longitudinal studies – a regression analysis of ADAS-cog, SKT and MMSE. *Pharmacopsychiatry* 1999; 32: 248-254.
2. **Clark SM, Sheppard L, Fillenbaum GG, Galasko D, Morris JC, Koss E, Mohs R, Heyman A.** Variability in annual Mini-Mental State Examination score in patients with probable Alzheimer disease: a clinical perspective of data from the consortium to Establish a Registry for Alzheimer's Disease. *Arch Neurol* 1999; 56 :857-862.
3. **Fillenbaum GG, Hughes DC, Heyman A, George LK, Blazer DG.** Relationship of health and demographic characteristics to Mini- Mental State examination score among community residents. *Psychol Med* 1988; 18: 719-26.
4. **O'Connor DW, Pollitt PA, Treasure FP, Brook CP, Reiss BB.** The influence of education, social class and sex on Mini-Mental State scores. *Psychol Med* 1989; 19: 771-6.
5. **Blessed G, Black SE, Butler T, Kay DV.** The diagnosis of dementia in the elderly. A comparison of CAMCOG (the cognitive section of CAMDEX] the AGECAAT program, DSM-III, the Mini-Mental State Examination and some short rating scales. *Br J Psychiatry* 1991; 159: 193-8.
6. **Ihl R, Froelich L, Dierks T, Martin EM, Maurer K.** Differential validity of psychometric tests in dementia of Alzheimer type. *Psychiatry Res* 1992; 2: 93-106.
7. **Monsch A-U, Foldi N-S, Ermini-Funfschilling D-E, Berres M, Taylor K-I, Seifritz E, Stahelin H-B, Spiegel R.** Improving the diagnostic accuracy of the Mini-Mental State Examination. *Acta Neurol Scand* 1995; 92: 145-50.
8. **Monsch AU.** Neuropsychological examination in evaluating dementia. *Schweiz Rundschau Med Prax* 1997; 86: 1340-1342.

9. **Cossa FM, Della Sala S, Musico M, Spinnler H, Uluzia MC.** Comparison of two scoring systems of the Mini Mental State Examination as a screening test for dementia. *J Clin Epidemiol* 1997; 50: 961-5.
10. **Folstein MF, Folstein SE, Mc Hugh PR.** "MINI MENTAL STATE" A practical method for grading the cognitive state of the patients for the clinician. *J Psychiatr Res* 1975; 12: 189-198.
11. **Riedel- Heller SG, Matschinger H, Schork A, Angermayer MC.** Do memory complaints indicate the presence of cognitive impairment? Results of a field study. *Eur Arch Psychiatry Clin Neurosci* 1999; 249: 197-204.
12. **Mezey M, Teresi J, Ramsey G, Mitty E, Bobrowitz T.** Decision-making capacity to execute a health care proxy: development and testing of guidelines. *J Am Geriatr Soc* 2000; 48: 179-87.
13. **Ihl R, Grass-Kapanke B, Lahrem P, Brinkmeyer J, Fischer S, Gaab N, Kaupmannsennecke C.** Entwicklung und Validierung eines Tests zur Früherkennung der Demenz mit Depressionsabgrenzung (TFDD). *Fortschr Neurol Psychiat* 2000; 68: 413-422.
14. **Brinkmeyer J, Grass-Kapanke B, Ihl R.** EEG and the Test for the Early Detection of Dementia with Discrimination from Depression (TE4D): a validation study. *Int J Geriatr Psychiatry* 2004; 19: 749-53.
15. **Mc Khann G, Folstein M, Katzman R, Price D, Stadlan EM.** Clinical diagnosis of Alzheimer disease: report of the NINCDS/ ADRDA work group under the auspices of Department of Health and Human services Task Force on Alzheimer's disease. *Neurology* 1984; 34: 939-944.
16. **Rosen WG, Mohs RC, Davis KL.** A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984; 11: 1356-1364.
17. **Erzigkeit H.** The development of the SKT project; in Hindmarch I, Hippus H, Wilcock GK (eds.): *Dementia: molecules, methods and measures*. New York, John Wiley, 1991.

18. **White N, Scott A, Woods RT, Wenger G-C, Keady J-D, Devakumar M.** The limited utility of the Mini-Mental State Examination in screening people over the age of 75 years for dementia in primary care. *Br J Gen Pract* 2002; 52: 1002-3.

19. **Wind AW, Schellevis F-G, Van Staveren G, Scholten RP, Jonker C, Van Eijk J-T.** Limitations of the Mini-Mental State Examination in diagnosing dementia in general practice. *Int J Geriatr Psychiatry* 1997; 12: 101-8.

Appendix

Test for the Detection of Dementia with Discrimination from Depression (TE4D]

1 Could you please read each of the words I will show you out loud and try to memorise them! What words do you remember?

seller comet news mirror fairy-tale steam adventure _____
sum 1 (7)

Could you please read the words again.

2 Could you please tell me the dates today? _____/_____/_____

Still correct if +/- 1 day _____ **sum 2 (3)**

3 Please tell me the seasons you know. S S A W _____
sum 3 (4)

4 What season is now?
Tolerance for the correct season is +/- 14 days _____
sum 4 (1)

5 What months belong to the present season?

<u>Spring</u>	<u>Summer</u>	<u>Autumn</u>	<u>Winter</u>
March <input type="checkbox"/>	June <input type="checkbox"/>	September <input type="checkbox"/>	December <input type="checkbox"/>
April <input type="checkbox"/>	July <input type="checkbox"/>	October <input type="checkbox"/>	January <input type="checkbox"/>
May <input type="checkbox"/>	August <input type="checkbox"/>	November <input type="checkbox"/>	February <input type="checkbox"/>
June <input type="checkbox"/>	September <input type="checkbox"/>	December <input type="checkbox"/>	March <input type="checkbox"/>

_____ **sum 5 (4)**

6 At first, please touch your right ear with your left hand
then touch your left ear with your right hand
and at the end clap your hands.

Sequence completely correct _____

sum 6 (4)

7 Please draw the dial of a clock including all numbers and set the hands to ten past eleven.

sum 7 (10)

8 At the beginning you read some words on cards and I asked you to memorise them. Could you please tell me the words you remember?

seller comet news mirror fairy-tale steam adventure _____
sum 8 (7)

9 Could you please name as many animals as you can in the next minute.

sum 9 (10)

Total score:

10. Self-rating

0 1 2 3 4 5 6 7 8 9 10

Not
depressed

Severly
depressed

11. Doctors rating

0 1 2 3 4 5 6 7 8 9 10

Not
depressed

Severly
depressed