

Interrater variabilita v hodnocení míry atrofie hipokampů pomocí Scheltensovy škály

Inter-rater Variability in Assessing Hippocampal Atrophy Using Scheltens Scales

Souhrn

Úvod: V poslední době se klade velký důraz na časnou diagnostiku Alzheimerovy choroby. Do nových diagnostických kritérií byly zařazeny nové diagnostické metody, které pomocí stanovení strukturálních a metabolických změn mozku tuto časnou diagnostiku umožňují. Patří mezi ně i vyšetření objemu hipokampů. Scheltens et al. zavedli vizuální škálu na hodnocení atrofie hipokampů. Cílem studie bylo zjistit interrater variabilitu a ověřit praktické použití této škály. **Metoda:** Osmi hodnotiteli s různým stupněm erudice v užívání Scheltensovy vizuální škály bylo vyhodnoceno 70 magnetických rezonancí pacientů s kognitivní poruchou a zdravých kontrol vyšetřených v Kognitivním centru při Neurologické klinice FN v Motole. Po krátkém zaškolení vyšetřující vyhodnotili tréninkový vzorek 20 pacientů s následným rozбором nálezů u každého z hodnotitelů a poté provedli vlastní zaslepené hodnocení atrofie hipokampů Scheltensovou škálou za použití snímků MR v koronární rovině. Variabilita byla posuzována pro všech osm hodnotitelů dohromady a pak ještě po jejich rozdělení na čtyři zkušené a čtyři nezkušené (před účastí v této studii neměli se škálou žádnou zkušenost). **Výsledky:** Nalezli jsme velmi vysoký koeficient shody pro všechny hodnotitele dohromady pro pravý ($K = 0,87$) i levý ($K = 0,88$) hipokampus. Velmi vysoký koeficient shody byl nalezen i po rozdělení do skupin zkušený/nezkušený vyšetřující v obou skupinách, a to pro pravý i levý hipokampus. **Závěr:** Scheltensova vizuální škála je jednoduše zvladatelný nástroj k hodnocení atrofie hipokampů i při malé erudici vyšetřovatele. Výsledná variabilita hodnocení je relativně malá a umožňuje doporučit tuto škálu pro rutinní klinické využití.

Abstract

Introduction: Recently, a great emphasis has been placed on early diagnosis of Alzheimer's disease (AD). The new diagnostic criteria for AD involve new methods such as detection of structural and metabolic changes in the brain. These include examination of hippocampal volume. Scheltens et al. introduced a visual rating scale for hippocampal atrophy assessment. Our aim was to determine inter-rater variability and to test practical application of this scale. **Methods:** MRI scans of 70 elderly persons with cognitive impairment and persons classified as cognitively normal were assessed by eight investigators. The investigators had a different degree of experience with Scheltens visual rating scale. Following a brief training, the investigators were asked to practice on 20 MRI brain scans in coronal plane. Correct scores were disclosed to all investigators. Subsequently, the investigators evaluated 70 study MRI scans. They were unaware of the participants' cognitive status. The variability was calculated for all investigators together and after dividing them into "experienced" ($n = 4$) and "inexperienced" ($n = 4$) group, where "inexperienced" referred to no previous knowledge of this visual rating scale. **Results:** Inter-rater agreement (kappa) was very high among all investigators for the right ($K = 0.87$) and left ($K = 0.88$) hippocampus. After dividing the raters into experienced and inexperienced, the inter-rater variability continued to be high in both groups for the right and left hippocampus. **Conclusion:** Scheltens visual rating scale is a simple and easy to use tool for hippocampal atrophy assessment even for inexperienced investigators. The final assessment variability was relatively low. We can recommend this visual rating scale for routine clinical use.

Děkujeme za finanční podporu grantu IGA NT 11225-4 a projektu FNUSA-ICRC (no. CZ.1.05/1.1.00/02.0123) z Evropského fondu regionálního rozvoje.

Autoři deklarují, že v souvislosti s předmětem studie nemají žádné komerční zájmy. The authors declare they have no potential conflicts of interest concerning drugs, products, or services used in the study.

Redakční rada potvrzuje, že rukopis práce splnil ICMJE kritéria pro publikace zasílané do biomedicínských časopisů.

The Editorial Board declares that the manuscript met the ICMJE "uniform requirements" for biomedical papers.

A. Kadlecová^{1,2}, M. Vyhnálek^{1,2}, J. Laczó^{1,2}, R. Andel^{2,3}, K. Sheardová⁴, B. Urbanová¹, Z. Nedělská^{1,2}, D. Hudeček², I. Gažová^{1,2}, J. Lisý⁵, D. Hořínek^{2,6}, J. Hort^{1,2}

¹ Kognitivní centrum, Neurologická klinika 2. LF UK a FN v Motole, Praha

² Mezinárodní centrum klinického výzkumu, FN u sv. Anny v Brně

³ School of Aging Studies, University of South Florida, Tampa, Florida, USA

⁴ Neurologická klinika, Mezinárodní centrum klinického výzkumu, FN u sv. Anny v Brně

⁵ Klinika zobrazovacích metod 2. LF UK a FN v Motole, Praha

⁶ Neurochirurgická klinika 1. LF UK a ÚVN – Vojenská fakultní nemocnice Praha



MUDr. Alexandra Kadlecová
Neurologická klinika
2. LF UK a FN v Motole
V Úvalu 84, 150 06 Praha
email: yooma@seznam.cz

Přijato k recenzi: 4. 7. 2012
Přijato do tisku: 12. 4. 2013

Klíčová slova

atrofie – hipokampus – vizuální škála – demence – Alzheimerova choroba

Key words

atrophy – hippocampus – visual scale – dementia – Alzheimer disease

Úvod

Alzheimerova choroba (ACH) je nejčastější příčina demence ve stáří [1]. V poslední době se výzkum zaměřuje na velmi časná stadia onemocnění, kdy neuropatologický proces není tak výrazně vyjádřen a kdy se předpokládá, že kauzální terapie (např. očkování proti betaamyloidu) by měla mít největší efekt [2]. V současnosti se ve stanovování diagnózy ACH uplatňují biomarkery [3]. Mezi ně patří např. strukturální a funkční zobrazovací biomarkery (magnetická rezonanční /MR/ volumetrie, pozitronová emisní tomografie) a metabolické (např. měření metabolitů v mozkomíšním moku) markery [4]. Praktické využití těchto markerů může narážet na časovou, finanční i logistickou náročnost nebo jejich přílišnou invazivitu. Např. vyšetření likvoru je invazivní a výsledky likvorologického vyšetření podléhají velké variabilitě podle vyšetřující laboratoře, což limituje jejich širší využití [5,6]. Vzhledem k relativně dobré dostupnosti a neinvazivitě vstupuje v poslední době do popředí význam vyšetření MR mozku. Jedná se o vyšetření neinvazivní a vhodné jak pro časný záchyt ACH, tak i pro monitorování průběhu tohoto onemocnění [5].

Vyšetření MR mozku nám u pacientů s demencí degenerativní etiologie umožňuje přesně změřit objemy anatomicky specifických oblastí mozku (MR volumetrie) a objemy takto naměřených oblastí posléze vztáhnout k výsledkům neuropsychologických testů [7–10]. Volumetrické techniky lze rozdělit na manuální, u nichž se určují hranice oblasti zájmu podle anatomického atlasu a následně se manuálně

obkreslují, a dále automatické, které využívají speciálních softwarových metod k automatickému propočtu objemů na základě rozdílu intenzit [11]. Manuální volumetrické techniky nám pomohou zjistit objem struktury, který víceméně odpovídá skutečnosti, ale vyžadují hlubokou znalost dané struktury, velkou zkušenost vyšetřujícího a jsou relativně časově náročné [12].

Již v časném stadiu ACH atrofuje hipokampus – struktura klíčová pro deklarativní paměť. Vzhledem k velké časové náročnosti manuální volumetrie hipokampu je tato metoda v klinické praxi nepoužitelná. Scheltens a jeho spolupracovníci

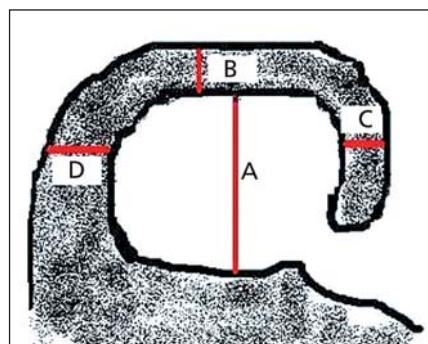
proto pro rutinní klinické využití vyvinuli Vizuální škálu hodnocení atrofie hipokampu [13]. Autoři prokázali, že pomocí této metody jsou schopni s velkou přesností detekovat atrofii hipokampů [14].

Cílem naší studie bylo zjistit, zda je použití této škály možné v běžné klinické praxi, jaká je interrater variabilita a jak se liší jednotliví hodnotitelé v závislosti na zkušenosti s užitím této škály.

Metodika

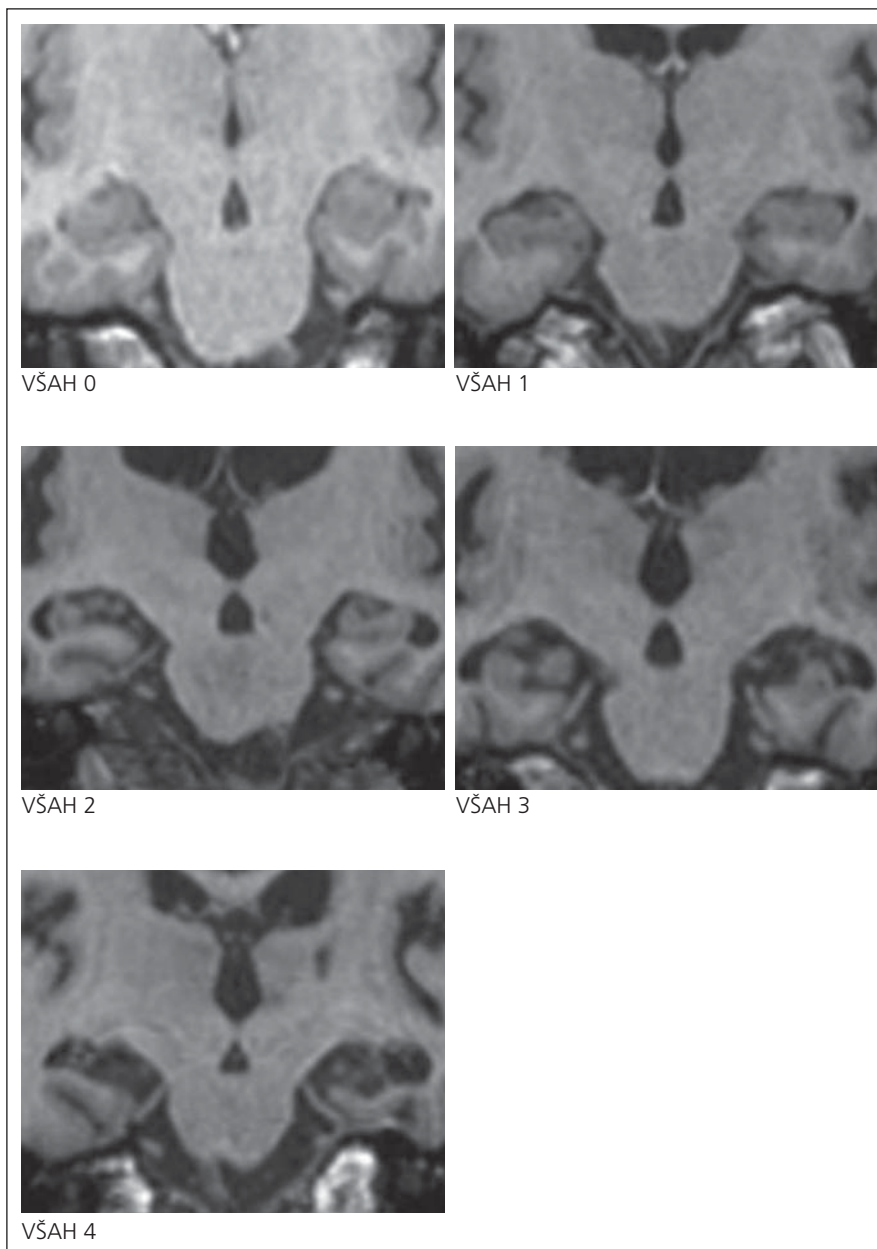
Pacienti

Do studie bylo zařazeno 70 MR snímků pacientů s kognitivní poruchou (demence



Obr. 1. Schéma koronálního řezu hipokampem v úrovni předního pontu.

Výška hipokampu (A), šířka choroidální štěrbině (B), šířka temporálního rohu (C), vzdálenost mezi mozkovým kmenem a hipokampem (D – nehodnotí se).



Obr. 2. Stupně vizuální škály atrofie hipokampu (VŠAH).

při ACH, mírná kognitivní porucha či subjektivní porucha paměti [3]) a zdravotních kontrol (z řad příbuzných pacientů a frekventantů Univerzity třetího věku UK v Praze), kteří byli vyšetřeni v Kognitivním centru při Neurologické klinice 2. LF UK a FN v Motole v období od 1. 1. 2011 do 1. 8. 2011. Všichni vyšetřovaní podepsali informovaný souhlas.

MR

Pacienti byli vyšetřeni 1,5T magnetickou rezonancí (Gyrosan, Philips Medical Systems). Atrofie byla hodnocena na koronárních řezech v 3D MP-RAGE (three-dimensional Magnetization-Prepared RAPid Gradient Echo) sekvenci o tloušťce řezu 1 mm. Celkový čas sekvence byl 14 min; TR = 25 ms; TE = 5 ms; úhel = 30°; pole 256 mm; a matrix 256 × 256.

Scheltensova vizuální škála hodnocení atrofie hipokampusu

Škála bere v úvahu výšku hipokampu (A), šířku choroidální štěrbiny (B) a šířku temporálního rohu (C) (obr. 1).

Škála se pohybuje od 0 (žádná atrofie) až po 4 (těžká atrofie). Přičemž při skóre 1 je pouze rozšířena choroidální štěrbina, při skóre 2 je rozšířen i temporální roh a lehce snížena výška hipokampu, při skóre 3 jsou výrazně rozšířeny choroidální štěrbina i temporální roh a snížen hipokampus, při skóre 4 je již masa hipokampu nepatrná [13]. K hodnocení se doporučují řezy v úrovni předního pontu (obr. 2).

Postup hodnocení:

- informační schůzka s prezentací škály (60 min),
- zadáním snímků 20 pacientů k trénování hodnocení během jednoho týdne,
- schůzka po týdnu s rozбором hodnocení pacientů jednotlivými hodnotiteli (1 hod), s následným zadáním seznamu (rodné číslo) 70 jiných pacientů,
- hodnotitelé samostatně v průběhu jednoho týdne zhodnotili pomocí Scheltensovy vizuální škály atrofii hipokampusu u všech 70 pacientů,
- po týdnu sběr dat.

Hodnotitelé

Hodnotilo osm lékařů: sedm neurologů, z nichž tři mají s vizuální škálou atrofie hipokampusu již předchozí zkušenost (od roku 2010), čtyři bez předchozí praktické zkušenosti a jeden specialista z oboru

Tab. 1. Hodnocení shody u zkušených a nezkušených hodnotitelů.

| | Coeff of concordance | F | Num DF | Denom DF | Prob > F |
|-------------------|----------------------|-------|--------|----------|----------|
| celkem/dx hip | 0,86717 | 19,58 | 63,5 | 190,5 | < 0,0001 |
| celkem/sin hip | 0,88173 | 22,37 | 63,5 | 190,5 | < 0,0001 |
| nezkušený/dx hip | 0,88039 | 22,08 | 63,5 | 190,5 | < 0,0001 |
| nezkušený/sin hip | 0,87644 | 21,28 | 63,5 | 190,5 | < 0,0001 |
| zkušený/dx hip | 0,87612 | 21,22 | 68,5 | 205,5 | < 0,0001 |
| zkušený/sin hip | 0,88740 | 23,64 | 68,5 | 205,5 | < 0,0001 |

dx hip – pravý hipokampus, sin hip – levý hipokampus.

Tab. 2. Aritmetické průměry skóre Scheltensovy škály u jednotlivých diagnóz u zkušených a nezkušených hodnotitelů.

| | Norma Sch. s. průměr | Smc Sch. s. průměr | MCI Sch. s. průměr | ACH Sch. s. průměr |
|-----------|----------------------|--------------------|--------------------|--------------------|
| zkušený | 1,20 | 1,17 | 1,52 | 1,94 |
| nezkušený | 1,15 | 0,94 | 1,34 | 1,76 |

zobrazovacích metod s předchozí zkušeností (od roku 2010).

Statistická analýza

Vzhledem k ordinálnímu charakteru hodnot byla k hodnocení míry shody použita neparametrická metoda – vážený koeficient Cohenovo kappa ve variantě pro více pozorovatelů. Pro výpočet bylo použito MAGREE macro pro SAS. Výpočet rozdílu mezi skupinami zajistil Kruskal-Wallisův a Mann-Whitneyův U-test, k hodnocení vztahu mezi výsledky byla použita Spearmanova korelace. Data byla zpracována statistickým softwarem SAS pro Windows.

Autoři prohlašují, že studie na lidských subjektech popsána v manuskriptu byla provedena v souladu s etickými standardy příslušné komise (institucionální a národní) odpovědné za provádění klinických studií a Helsinskou deklarací z roku 1975, revidovanou v r. 2000.

Výsledky

Analýza potvrdila velmi dobrou shodu mezi vyšetřujícími – vážený koeficient kappa byl 0,87 pro pravý a 0,88 pro levý hipokampus ($p < 0,0001$).

Při samostatném hodnocení shody u zkušených a nezkušených hodnotitelů zvlášť nebyly zjištěny výraznější rozdíly

v míře shody mezi oběma skupinami vyšetřujících (tab. 1).

Z porovnání absolutních hodnot měření vyplynulo, že zkušený vyšetřující hodnotil snímky přísněji než vyšetřující nezkušený. Přesněji řečeno, zprůměrování všech hodnot zkušených a nezkušených hodnotitelů aritmetickým průměrem a jejich následné porovnání ukázalo rozdíly v hodnocení mezi těmito dvěma skupinami (1,74 vs 1,25; $p < 0,0001$). Přičemž nezkušený hodnotitelé podhodnocovali výsledek ve všech skupinách, ale ne o více než 0,5 bodu (tab. 2)

Pacienti s ACH se velikostí hipokampusu dle Scheltensovy škály lišili od kontrol ($p = 0,045$) i od pacientů se subjektivní poruchou paměti ($p = 0,003$), pacienti ze skupiny norem, subjektivních poruch paměti a MCI se navzájem nelišili ($p \geq 0,189$).

MMSE slabě pozitivně korelovalo s velikostí hipokampusu v celém vzorku subjektů ($r = 0,225$; $p = 0,034$); jinak se v MMSE ACH lišily od všech skupin ($p < 0,001$), ale ostatní skupiny se mezi sebou v MMSE nelišily (tab. 2).

Diskuze

Z našich výsledků vyplývá, že i při relativně rozdílné délce zkušenosti s užíváním vizuální škály hodnocení atrofie hipokampusu (VŠAH) byla míra shody mezi hodnoti-

Tab. 3. Průměrné hodnoty Scheltensovy škály a MMSE u jednotlivých skupin.

| | Norma n = 15 průměr ± SD | Smc n = 12 průměr ± SD | MCI n = 22 průměr ± SD | ACH n = 21 průměr ± SD |
|---------------|-----------------------------|---------------------------|---------------------------|---------------------------|
| skóre Schelt. | 1,16 ± 0,51 | 1,06 ± 0,40 | 1,50 ± 0,75 | 1,86 ± 0,72 |
| MMSE | 29 ± 1,41 | 28,91 ± 0,86 | 27,82 ± 1,90 | 20,63 ± 3,94 |

teli velmi vysoká a byla srovnatelná s výsledky, které byly hodnoceny jako uspokojivé i u jiných škál [14–16]. Škála je pro svou jednoduchost a jednoznačnost dobře uchopitelná i pro méně zkušené hodnotitele a k dosažení dostatečného stupně erudice stačil krátký trénink.

Z naměřených výsledků vyplývá, že zkušenější hodnotitelé byli všeobecně přesnější oproti nezkušeným, kteří se přiklání spíše k nižšímu skóru, ale v průměru tento rozdíl nepřesahoval 0,5 bodu, přičemž k tomuto jevu docházelo u všech skupin pacientů i kontrol.

V klinické praxi by to mohlo znamenat, že méně erudovaní hodnotitelé by při stanovování diagnózy ACH mohli mít při posuzování atrofie hipokampů falešně negativní výsledky. Atrofie mezeitemporálních struktur se vyskytuje i u jiných typů neurologických onemocnění [17–19], např. typicky u pacientů s frontotemporální demencí, jejíž klinické projevy jsou z části podobné jako u ACH, nebo i u jedinců s normálním stárnutím [20]. Z tohoto důvodu je třeba interpretovat hodnocení této škály obezřetně a v kontextu dalších klinických vyšetření [5]. I v naší skupině MCI se najdou pacienti, kteří mají vyšší skóre než 2, někdy pouze na jedné straně, a to by již mělo vyšetřujícího navést k časťším kontrolám a sledování jak kliniky, tak vývoje atrofie, aby byla případná konverze do demence odhalena časně.

Ve studii prováděné Koedamem et al [15] byla u této škály hodnocena variabilita mezi třemi zkušenými hodnotiteli, která se pohybovala mezi 0,82 a 0,9, což se příliš neliší od výsledků v naší skupině zkušených hodnotitelů. V původních studiích Scheltense při zavádění této škály se variabilita pohybovala mezi 0,34 (95% CI = 0,34–0,54) a 0,57 (95% CI = 0,41–0,61) [21], což je horší výsledek než v naší skupině.

Nález variability u VŠAH jsou relativně vyšší než opakovaná měření automatickými softwary – např. FreeSurferem,

kde se variační koeficient pohybuje mezi 1,3 a 2,6 % [22]. Nedávná studie však poukázala na limity automatické volumetrie v tom smyslu, že použití různých verzí FreeSurferu a různých operačních systémů (Apple Macintosh a Windows) poskytovalo variabilitu v hodnocení objemů více než 15 % [23]. Nevýhodou automatické volumetrie je pak rovněž časová a logistická náročnost praktického provedení, náklady na vybavení a erudici personálu, který program ovládá. VŠAH naproti tomu může použít každý klinický pracovník.

Původně byla škála VŠAH vyvinuta pro CT metodiku, kde lze snímky rovněž hodnotit. Celkově je však vypovídací hodnota MR hodnocení vyšší než CT hodnocení. Teoretická možnost hodnocení na CT však znamená, že je opravdu využitelná v klinické praxi velmi dobře [24].

Naše MR snímky byly zhotoveny po 1 mm ve speciálním protokolu našeho centra. Rutinně se na řadě pracovišť používají většinou řezy 4 mm, proto je zacílení hipokampu složitější a teoreticky méně přesné, navíc hrubší řezy většinou atrofii nadhodnocují. Je také důležité provést řez v přední části pontu, na kterém je zachycena interpedunkulární cisterna, což může napomoci orientaci (obr. 2).

VŠAH se ukázala být rychlá a dostupná metoda pro posouzení stupně atrofie v klinickém prostředí ve srovnání s volumetrickými metodami, jež vyžadují mnohem větší technickou zdatnost a jsou časově náročné. VŠAH je jedním z nejdůležitějších markerů ACH při hodnocení MR. V nově vytvořených diagnostických kritériích ACH je přítomnost postižení epizodické paměti a detekce atrofie hipokampů pomocí VŠAH stěžejní pro stanovení této diagnózy [4].

Literatura

- Scarpini E, Scheltens P, Feldman H. Treatment of Alzheimer's disease: current status and new perspectives. *Lancet Neurol* 2003; 2(9): 539–547.
- Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC et al. The diagnosis of mild cognitive

impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; 7(3): 270–279.

3. Hort J, O'Brien JT, Gainotti G, Pirttilä T, Popescu BO, Rektorova I et al. EFNS Scientist Panel on Dementia. EFNS guidelines for the diagnosis and management of Alzheimer's disease. *Eur J Neurol* 2010; 17(10): 1236–1248.

4. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* 2007; 6(8): 734–746.

5. Hort J, Glosová L, Vyhnálek M, Bojar M, Škoda D, Hladíková M. The liquor tau protein and beta amyloid in Alzheimer's disease. *Cesk Slov Neurol N* 2007; 70/103(1): 30–36.

6. Hort J, Bartos A, Pirttilä T, Scheltens P. Use of cerebrospinal fluid biomarkers in diagnosis of dementia across Europe. *Eur J Neurol* 2010; 17(1): 90–96.

7. Nedelska Z, Anđel R, Laczo J, Vlček K, Horinek D, Lisy J et al. Spatial navigation impairment is proportional to right hippocampal volume. *Proc Natl Acad Sci U S A* 2012; 109(7): 2590–2594.

8. Venneri A, Gorgoglione G, Toraci C, Nocetti L, Panzetti P, Nichelli P. Combining neuropsychological and structural neuroimaging indicators of conversion to Alzheimer's disease in amnesic mild cognitive impairment. *Curr Alzheimer Res* 2011; 8(7): 789–797.

9. Benoit M, Robert PH. Imaging correlates of apathy and depression in Parkinson's disease. *J Neurol Sci* 2011; 310(1–2): 58–60.

10. Heckers S, Konradi C. Hippocampal pathology in schizophrenia. *Curr Top Behav Neurosci* 2010; 4: 529–553.

11. Horinek D, Varjassyová A, Hort J. Magnetic resonance analysis of amygdalar volume in Alzheimer's disease. *Curr Opin Psychiatry* 2007; 20(3): 273–277.

12. Cavallin L, Bronge L, Zhang Y, Oksengård AR, Wahlund LO, Fratiglioni L et al. Comparison between visual assessment of MTA and hippocampal volumes in an elderly, non-demented population. *Acta Radiol* 2012; 53(5): 573–579.

13. Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* 1992; 55(10): 967–972.

14. Korf ES, Wahlund LO, Visser PJ, Scheltens P. Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. *Neurology* 2004; 63(1): 94–100.

15. Koedam EL, Lehmann M, van der Flier WM, Scheltens P, Pijnenburg YA, Fox N et al. Visual assessment of posterior atrophy development of a MRI rating scale. *Eur Radiol* 2011; 21(12): 2618–2625.

16. Kapeller P, Barber R, Vermeulen RJ, Adèr H, Scheltens P, Freidl W et al. European Task Force of Age Related White Matter Changes. Visual rating of age-related white matter changes on magnetic resonance imaging: scale comparison, interrater agreement, and correlations with quantitative measurements. *Stroke* 2003; 34(2): 441–445.

17. Weintraub D, Doshi J, Koka D, Davatzikos C, Siderowf AD, Duda JE et al. Neurodegeneration across stages of cognitive decline in Parkinson disease. *Arch Neurol* 2011; 68(12): 1562–1568.

18. Diniz PR, Velasco TR, Salmon CE, Sakamoto AC, Leite JP, Santos AC. Extratemporal damage in temporal lobe epilepsy: magnetization transfer adds information to volumetric MR imaging. *AJNR Am J Neuroradiol* 2011; 32(10): 1857–1861.

19. Hayashi K, Kurioka S, Yamaguchi T, Morita M, Kanazawa I, Takase H et al. Association of cognitive dysfunction with hippocampal atrophy in elderly Japanese people with type 2 diabetes. *Diabetes Res Clin Pract* 2011; 94(2): 180–185.
20. Laakso MP, Partanen K, Riekkinen P, Lehtovirta M, Helkala EL, Hallikainen M et al. Hippocampal volumes in Alzheimer's disease, Parkinson's disease with and without dementia, and in vascular dementia: an MRI study. *Neurology* 1996; 46(3): 678–681.
21. Scheltens P, Launer LJ, Barkhof F, Weinstein HC, van Gool WA. Visual assessment of medial temporal lobe atrophy on magnetic resonance imaging: interobserver reliability. *J Neurol* 1995; 242(9): 557–560.
22. Wu Y, Du H, Storey P, Glielmi C, Malone F, Sidharthan S et al. Comprehensive brain analysis with automated high-resolution magnetization transfer measurements. *J Magn Reson Imaging* 2012; 35(2): 309–317.
23. Gronenschild EH, Habets P, Jacobs HI, Mengelers R, Rozendaal N, van Os J et al. The effects of FreeSurfer version, workstation type, and Macintosh operating system version on anatomical volume and cortical thickness measurements. *PLoS One* 2012; 7(6): e38234.
24. Wattjes MP, Henneman WJ, van der Flier WM, de Vries O, Träber F, Geurts JJ et al. Diagnostic imaging of patients in a memory clinic: comparison of MR imaging and 64-detector row CT. *Radiology* 2009; 253(1): 174–183.