

UNIVERZITA KARLOVA

Filozofická fakulta

Katedra psychologie

SEMINÁRNÍ PRÁCE

Test hodin v diagnostice neurodegenerativních onemocnění

Michaela Vaškovicová, Jan Netík

Akademický rok: 2018/2019

Obor: Psychologie, 1. ročník nMgr.

Prohlášení:

Prohlašujeme, že jsme tuto práci zpracovali jako seminární práci v rámci předmětu Metodologická praktika v psychologii a nebyla ani nebude použita k účelům jiné atestace.

V Praze, dne 23. října 2018

.....
Michaela Vaškovicová

.....
Jan Netík

Abstrakt

Cílem práce je krátce popsat test hodin a zhodnotit jeho psychometrické vlastnosti z hlediska jeho využitelnosti v diagnostice neurodegenerativních onemocnění. První část se věnuje obecným aspektům testu, jeho administraci a skórovacím systémům. Druhá část je zaměřena na meta-analýzy zabývající se testem hodin a další studie zkoumající jeho validitu a reliabilitu, a to nejenom z hlediska detekce demence, ale i MCI.

Klíčová slova

Test hodin, demence, MCI, neurodegenerativní onemocnění, neuropsychologie, screening

Abstract

The aim of this paper is to briefly describe the Clock Drawing Test and its psychometric parameters when used in the diagnostics of neurodegenerative diseases. First part introduces general aspects of the test, its administration and scoring systems. Second part is focused on meta-analytic studies of the Clock Drawing test and other studies exploring its validity and reliability, not only for detection of dementia, but also of MCI.

Keywords

Clock Drawing Test, dementia, MCI, neurodegenerative diseases, neuropsychology, screening

OBSAH

1.	ÚVOD	2
2.	O TESTU HODIN.....	4
2.1	ADMINISTRACE TESTU.....	4
2.2	SKÓROVACÍ SYSTÉMY	5
3.	PSYCHOMETRICKÉ VLASTNOSTI TESTU.....	7
3.1	META-ANALÝZY SKÓROVACÍCH SYSTÉMŮ	7
3.1.1	<i>Park et al. (2018)</i>	7
3.1.2	<i>Tsoi et al. (2015)</i>	9
3.2	VALIDITA TESTU HODIN PRO DIAGNOSTIKU MCI.....	10
3.2.1	<i>Ehreke et al. (2009)</i>	10
3.2.2	<i>Vyhnálek et al. (2016)</i>	11
3.2.3	<i>Rubínová et al. (2014)</i>	12
3.2.4	<i>Ricci et al. (2016)</i>	12
3.2.5	<i>Rakusa, Jensterle, & Mlakar (2018)</i>	12
3.3	RELIABILITA TESTU.....	13
3.3.1	<i>Savaş, Yerlikaya, & Yener (2018)</i>	13
3.3.2	<i>Mazancova, Nikolai, Stepankova, Kopecek, & Bezdicek (2016)</i>	13
4.	ZÁVĚR.....	14
5.	SEZNAM POUŽITÉ LITERATURY.....	16

1. Úvod

V současné době se stále zvyšuje poptávka po časně diagnostice neurodegenerativních onemocnění, přičemž nejčastěji se jedná o Alzheimerovu nemoc (AN), jelikož svou prevalencí výrazně převyšuje ostatní typy neurodegenerativních onemocnění. V současnosti se stále více využívá neuropsychologických baterií pro identifikaci typických kognitivních profilů jednotlivých neurodegenerativních onemocnění, tedy pro zpřesnění diagnostiky. Tyto baterie by měly zejména zahrnovat testy zaměřené na následující domény: paměť, pozornost a pracovní paměť, exekutivní funkce, fatické funkce a vizuospeciální schopnosti (Nikolai, Štěpánková, Vyhnálek, & Kopeček, 2016).

V běžné klinické praxi se však s velkou četností využívají i jednodušší screeningové metody, které jsou schopny detekovat zejména až fázi demence, v lepších případech již předcházející fázi mírné kognitivní poruchy (MCI). Screeningové testy umožňují rychlou orientaci a zhodnocení stavu pacienta, což umožňuje lépe stanovit další diagnostický postup (Nikolai et al., 2016). Screeningové testy by měly splňovat několik základních podmínek: jednoduchá a krátká administrace, dobré přijetí testu pacienty, snadné skórování, relativní nezávislost na kultuře, jazyce a vzdělání, vysoká shoda mezi posuzovateli, test-retest reliabilita, souběžná validita, senzitivita a specifita a rovněž prediktivní validita (Shulman, 2000).

Jako zlatý standard se v klinické praxi v ČR běžně využívá screeningový test MMSE (vhodný zejména pro odlišení středně těžké demence od normálního stavu), na nějž je vázána preskripce kognitiv (Štěpánková et al., 2015). Administrace tohoto testu může u zdravého jedince trvat i 5 minut, u jedince s demencí i výrazně déle (Rakusa, Jensterle, & Mlakar, 2018). Mezi nevýhody testu dále patří, že je málo senzitivní pro MCI a je výrazně zaměřen na fatické funkce (těžištěm testu jsou verbální úlohy), zatímco téměř zcela opomíjí funkce vizuospeciální a zcela chybí položky zaměřené na exekutivní funkce. Vzhledem ke zaměřenosti na jazykové funkce je poměrně kulturně specifický a závislý na věku a vzdělání (Štěpánková et al., 2015).

Jako jednu z alternativ či doplněk pro MMSE může představovat test hodin. Často se v diagnostice neurodegenerativních onemocnění využívá pro testování vizuospeciálních schopností, zejména jako součást neuropsychologických baterií (Nikolai et al., 2016), o něco méně lze s jeho pomocí identifikovat deficit exekutivních funkcí (Juby,

Tench, & Baker, 2002; Mittal, Gorthi, & Rohatgi, 2010), nicméně bývá často označován i přímo za screeningový test demence (Mittal et al., 2010; Palsetia, Rao, Tiwari, Lodha, & De Sousa, 2018; Rakusa et al., 2018). Tedy funkce, které MMSE pokrývá minimálně, test hodin postihuje primárně.

2. O testu hodin

Test hodin se od doby svého vzniku stal oblíbeným nástrojem pro kognitivní screening, primárně pro kognitivní deficit u Alzheimerovy nemoci a jiných neurodegenerací, nicméně jeho využitelnost přesahuje i do dalších oblastí diagnostiky (Hazan, Frankenburg, Brenkel, & Shulman, 2018). Ačkoliv se na první pohled může zdát, že se jedná o poměrně banální úkol, jedná o komplexní cílem řízenou činnost, test hodin zahrnuje pochopení, percepci, paměť, hrubou motoriku, vizuospaciální funkce, soustředění, znalost čísel, chápání konceptu času a inhibici rozptylujících stimulů (Mittal et al., 2010).

Mittal et al. (2010) upozorňují, že je však třeba brát ohled na limity tohoto testu, selhávání v CDT automaticky neznamena diagnózu demence, ale indikuje potřebu dalšího testování. Stejně tak normální MMSE skóre nevyklučuje diagnózu demence či MCI. Nicméně kombinace těchto dvou testů může být přínosná.

2.1 Administrace testu

Test hodin je sice poměrně rozšířený a známý, avšak neexistuje jednotný přístup k jeho administraci ani ke skórování. Co se týče administrace, v zásadě se buď požádá pacient, aby na čistý list zakreslil celé hodiny, anebo je rovněž četná varianta s již předkresleným kruhem hodin, do nějž pacient vepíše číslice a hodinové ručičky. Nicméně existují i jiné méně užívané formy testu hodin. (Hazan et al., 2018).

Například v české jazykové mutaci screeningového testu MoCA je využívá varianta testu hodin bez překresleného kruhu s instrukcemi: Zde namalujte obrys hodin, umístěte čísla označující hodiny a ručičkami vyznačte 11 hodin a 10 minut (“MoCA Montreal - Cognitive Assessment,” 2018).

Co se týče požadovaného času, teoreticky se může zvolit libovolný čas, nicméně v praxi existují určité časy, které se zpravidla používají. Asi nejoblíbenější variantou je čas 11 hodin 10 minut. A to nejenom proto, že pacient musí využít dva různé kvadranty hodin, ale zároveň pro správné provedení se pacient musí inhibovat tak, aby nedošlo k fenoménu známému jako „frontal pull“, kdy pacient chybně minutovou ručičku nastaví na číslici 10 místo na číslici 2. (Hazan et al., 2018) Dle Ressnera a Ressnerové (2005) je test hodin schopen zachytit perseveraci a poruchy plánování, tedy je možné usuzovat právě na stav frontálních oblastí.

2.2 Skórovací systémy

Přes svoji popularitu Test hodin stále nemá široce přijímaný způsob administrace¹ a již vůbec není možné nalézt shodu na kritériích, podle kterých bychom měli výkony pacientů hodnotit a kvantifikovat (Hazan, Frankenburg, Brenkel, & Shulman, 2017). Je tedy zavádějící hovořit o validitě či reliabilitě metody jako takové – ve skutečnosti vždy posuzujeme konkrétní skórovací systém a s ním spojenou kvalitu administrátora, jehož úkolem je pacientův výstup klasifikovat v souladu s navrženými kritérii. Vůbec nejčastějšími měřítky, kterými autoři níže diskutovaných výzkumů „metodu“ popisují, jsou tedy *senzitivita, specificita a inter-rater reliabilita* právě použitého skórovacího systému.

V přehledové práci Hazana et al. (2017) se dozvídáme, že existuje více než 20 postupů, jak výsledné provedení hodin ohodnotit; některé jsou volné, spíše kvalitativní a vycházejí z „klinického dojmu“, některé naopak udávají přesná kritéria včetně exaktně definovaných odchylek, rozestupů mezi jednotlivými elementy, jejich velikosti, poměry apod. Stranou pak stojí automatické zpracování počítačem pomocí strojového učení, které recentně navrhli Souillard-Mandar et al. (2015). Mezi nejčastěji užívané skórovací systémy, které se objevují i v meta-analýzách popsaných níže, se však řadí zejména hodnocení Shulmanovo (Shulman, Shedletsy, & Silver, 1986), které je nejstarší, a Sunderlandovo (Sunderland et al., 1989). V českém prostředí se pak vedle jednoduchých kritérií stanovených v rámci screeningových metod, jakou je např. *Montrealský kognitivní test* (MoCA; Nasreddine, 2015) či *Addenbrookský kognitivní test* (ACE-R; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006) ujal novější a podrobnější systém Cohenův (Cohen, Ricci, Kibby, & Edmonds, 2000), který byl v původním článku – poněkud překvapivě – využit pro mapování vývoje různých kognitivních a exekutivních schopností u dětí. Na rozdíl od ostatních systémů vychází z odlišného teoretického rámce a hodnotí

¹ Jak zmiňují Hazan et al. (2017), někteří praktici nechají pacienta nakreslit celé hodiny na čistý papír, jiní předkládají list s předtištěnou kružnicí přesně definované velikosti. Jiný způsob dokonce vyžaduje pouhé překreslení předlohy hodin, což je sice velmi vzdálené původní úloze, avšak někdy se sleduje, zda pacient po tomto odlišném zadání svůj předchozí patologický výkon „nenormalizuje“. Podle Hazana et al. (2017) se lze setkat také se zadáními, které vyžadují nastavení času na modelu hodin, čtení času apod. Velmi diskutovaný je vlastní čas, který v klasickém zadání po pacientovi požadujeme pomocí ručiček vyznačit – zdaleka nejčastější je však 11:10, na kterém lze snáze pozorovat některé specifické chyby.

separátně formu hodin a schopnost nastavení času (Mazancová, 2014). Bohužel se objevuje v naprostém minimu studií a pokud je nám známo, není v žádné meta-analýze.

3. Psychometrické vlastnosti testu

Vzhledem k nejednotnosti rozmanitých skórovacích technik testu hodin nelze stanovit psychometrické vlastnosti testu hodin jako takového, pouze jeho konkrétních variant a skórovacích systémů (Park, Jeong, & Seomun, 2018). Každý skórovací systém má svou vlastní metodologii a instrukce a ze srovnání jednotlivých systémů vychází, že žádný z nich není konzistentně nejlepší z hlediska prediktivní validity pro demence (Mainland, Amodeo, & Shulman, 2014).

3.1 Meta-analýzy skórovacích systémů

Dostáváme se tedy k problematice samotných důkazů o validitě a dalších vlastnostech metody. Jak vyplývá z předchozích odstavců, není snadné je systematicky zhodnotit. Pátrání po informacích jsme zahájili v několika databázích (PubMed, MEDLINE, EBSCO a ve vyhledávači Google Scholar), kam jsme zadávali klíčová slova „clock“, „drawing“, „test“, „meta-analysis“, „metaanalysis“, „meta“, „analysis“, „dementia“, „cognitive“, „impairment“ a „screening“. Již z uvedených klíčových slov vyplývá, že naším cílem bylo vyhledat primárně meta-analýzy, které vnímáme jako jeden z nejvyšších úrovní důkazů. Soustředili jsme se zejména na recentní práce.

Překvapilo nás, že k takto široce využívanému testu s relativně dlouhou historií existují pouze dvě studie splňující výše uvedené požadavky. První pochází od Parka, Jeonga a Seomuna (2018) a autoři v ní shrnují všechny systematicky vyhledané články o Testu hodin nezávisle na skórovacích kritériích. Druhá, starší meta-analýza od týmu Tsoi, Chan, Hirai, Wong a Kwok (2015) pojednává obecně o screeningových zkouškách určených k časnému záchytu demence různé etiologie, včetně Testu hodin.

3.1.1 Park et al. (2018)

Jedná se o nejnovější meta-analýzu týkající se Testu hodin (podle autorů jde o první práci svého druhu), která si dala přímo za cíl zhodnotit „diagnostickou přesnost“ metody (resp. celkem 7 skórovacích systémů). Autorský tým postupoval podle několika uznávaných „guidelines“ a metodik standardně používaných u meta-analýz diagnostických metod (např. protokolu QUADAS-2). Při výpočtech „meta-analytické“ („pooled“, „sdružené“) senzitivity, specificity, věrohodnostních poměrů a dalších ukazatelů schopností metody pak autoři užívali i doporučené postupy postihující rozdíly mezi jednotlivými studii a další zkrácení. Do studie bylo zahrnuto celkem 18 prací, přičemž

„sdružené“ odhady parametrů výzkumníci spočítali na 15 z nich. Celkem se tak meta-analýzy „zúčastnilo“ 5 531 jedinců (pacientů s demencí různé, neurodegenerativní i vaskulární etiologie, participantů s mírnou kognitivní poruchou (MCI) a zdravých kontrol).

Výše zmíněný Shulmanův systém použilo celkem 7 analyzovaných prací, z nichž byla odvozena „sdružená“ senzitivita 82 % [73,1 %; 88,4 %]², specificita 75,7 % [54,3 %; 89,1 %] a plocha pod tzv. *hierarchical summary receiver operating characteristic* (HSROC) křivkou (*area under curve*; AUC) dosáhla uspokojivé hodnoty 0,857; Shulmanovo skórování tak autoři považují za nejlepší a nejpřesnější. Zvýrazňují jeho senzitivitu a doporučují ho i v samotném závěru k použití v každodenní klinické praxi.

Druhý zmíněný, Sunderlandův systém se naopak v jejich meta-analýze vyznačoval nejvyšší „sdruženou“ specificitou: 87,9 % [74,0 %; 94,8 %], senzitivitou 72,6 % [62,9 %; 80,6 %] a AUC = 0,835. Ostatní analyzované systémy (Rouleau, Watson, Manos, Freedman a Wolf-Klein) pak byly celkově horší – vyskytovala se u nich např. vysoká specificita, ale velmi nízká senzitivita.

V diskusi autoři zmiňují běžná omezení podobných prací, mezi něž se řadí zejména publikační zkreslení, které nadhodnocuje skutečné efekty. Dále si stěžují na heterogenitu použitých skórovacích systémů a obecný nedostatek studií vyhodnocujících validitu metody.

V uvedené meta-analýze však osobně spatřujeme ještě mnohem závažnější skutečnosti, které by se daly chápat jako nedostatky či chyby. Předně nás zarazí, že se autoři nijak nepozastavují nad složením „diagnostické skupiny“ – ve většině analyzovaných prací tuto skupinu tvoří osoby s blíže nedefinovanou demencí, u pouhých tří studií je přímo uvedena Alzheimerova nemoc ve stádiu demence, dvakrát je diagnózou MCI a u jedné práce se dokonce uvádí „cognitive impairment“, což není žádná existující diagnóza, ale poněkud vágně definovaný symptom. V jedné studii (Riedel, Klotsche, Förstl, & Wittchen, 2013) jsou např. zahrnuti pacienti s Parkinsonovou nemocí a přítomným, nebo absentujícím kognitivním deficitem; studie se nadto zabývá především efektem komorbidní deprese³. V reálném prostředí běžné ordinace se na druhou stranu také

² Všechny konfidenční intervaly v naší práci jsou 95 %.

³ Pro zajímavost uveďme, že autoři zmíněného článku zjistili u pacientů s komorbiditou deprese doslova *zhroucení specificity* k 55,8 % [47,6 %; 63,7 %], senzitivita pak vyšla 72,3 % [64,7 %; 79,1 %]. Test hodin je tak u depresivních pacientů prakticky nepoužitelný – pravděpodobnost demence u pozitivního

poměrně často setkáváme s pacienty s nekonkluzivní symptomatikou zahrnující četné komorbidity (velice často je to právě porucha nálady), jejichž „pravá“ diagnóza zůstává dlouho neodhalena. Horší je podle nás zahrnutí prací s pacienty s MCI, k jejichž zachycení je Test hodin pravděpodobně nedostatečný (viz další části našeho textu) a může tak „znehodnocovat“ některé skórovací systémy (studie skórující dle autory vyzdvihovaného Shulmanova systému obsahovaly zrovna pacienty výhradně jen s demencí).

3.1.2 Tsoi et al. (2015)

Meta-analýza od Tsoi et al. (2015) není zaměřena pouze na test hodin, ale obecně na různé známé kognitivní testy využívané pro detekci demence. Studie splňovala standartní směrnice PRISMA a Cochrane Diagnostic Test Accuracy Working Group. Nicméně vzhledem k tomu, že je studie zaměřena na široké spektrum testů kognice, Testu hodin je věnováno v rámci článku minimum prostoru, více se zde autoři věnují známějším metodám typu MMSE a další. Je to však pochopitelné s ohledem na to, kolik průřezových studií vhodných pro zařazení do meta-analýzy autoři našli pro MMSE (108) a pro Test hodin (18).

Průřezové studie byly do analýz zařazeny, pokud splňovaly následující kritéria.

1. Pacienti v rámci daných studií byli testováni za účelem detekce demence spojené s Alzheimerovou nemocí, vaskulární demencí či Parkinsonovou nemocí, a to v klinickém prostředí.
2. Testování pacientů probíhalo v rámci osobního kontaktu (face-to-face).
3. V rámci studií byla použita standartní diagnostická kritéria (DSM, ICD, NINCDS-ADRDA a další) pro definování demence, popřípadě klinický úsudek na základě komplexní série vyšetření.
4. Studie uváděly počet zahrnutých pacientů s demencí a hodnotily přesnost použitých testů (senzitivitu, specificitu) nebo poskytovaly data, ze kterých tyto informace bylo možno odvodit.

Autorům se podařilo identifikovat celkem 18 studií zabývajících se testem hodin, které splňovaly daná kritéria. 9 studií používalo Sunderlandův systém, 9 studií Shulmanův

výsledku (positive predictive value) je téměř shodná s pravděpodobností u výsledku negativního (negative predictive value): 62,5 % [55,1 %; 69,5 %], resp. 66,4 % [57,6 %; 74,4 %]. Dodejme však, že velmi záleží na prevalenci obou chorob (Riedel et al., 2013). Tento moment dobře postihují i samotní autoři meta-analýzy v přehledných grafech (Park et al., 2018).

systém. Získaná souhrnná senzitivita pro Sunderlandův systém byla 76 % [69 %; 83 %] a souhrnná specifická 85 % [76 %; 91 %]. Získaná souhrnná senzitivita pro Shulmanův systém byla 83 % [75 %; 89 %] a souhrnná specifická 84 % [69 %; 92 %]. O něco lépe tedy z tohoto srovnání vychází Shulmanův systém s velmi uspokojivou senzitivitou i specifitou. AUC pro testy hodin autoři bohužel neuvádí.

Domníváme se, že tato meta-analýza si lépe poradila s výběrem vhodných studií, a tedy získala více homogenní vzorek, co se týče diagnostických skupin. Studie zabývající se pouze MCI či studie s nejasnými diagnostickými skupinami nebyly do analýz zařazeny. Nicméně jako nedostačující vnímáme prostor věnovaný Testu hodin v rámci této studie, kdy sice v tabulkách s výsledky analýz test je přítomen ale chybí jakýkoliv verbální komentář v rámci článku. Autoři si dali práci s hledáním, výběrem a analýzou studií s Testem hodin, tedy je politováníhodné, že jej pak v článku jinak zcela opomenuli a neudělali více doprovodných analýz získaných dat.

3.2 Validita testu hodin pro diagnostiku MCI

Obě dvě meta-analýzy rozvedené výše se shodují na tom, že Test hodin je při použití vybraných (a v anglofonní literatuře zdaleka nejčastějších) skórovacích systémů Shulmana a Sunderlanda vhodný pro screening demence různé etiologie. Méně zkoumanou otázkou, která se pochopitelně mohla objevit až po publikování příslušných kritérií a jejich zavedení do praxe na přelomu tisíciletí (viz např. Petersen, 2004), je validita testu hodin pro screening a diagnostiku MCI. K takto specifikovanému tématu se nám nepodařilo dohledat jedinou meta-analýzu, ale identifikovali jsme jednu systematickou přehledovou studii (Ehreke, Luppá, König, & Riedel-Heller, 2009) a několik novějších studií, které podle nás také přinášejí cenné informace. Kromě výše zmíněné systematické přehledové studie jsme našli recentnější články, které podle nás přinášejí kvalitní důkazy relevantní i v České republice (velká část překvapivě pochází od tuzemských týmů). Ze studia desítek prací o metodě se domníváme, že čím je práce aktuálnější, tím jsou její závěry validnější, a tak je zde pro úplnost zmíníme.

3.2.1 Ehreke et al. (2009)

Autoři této přehledové studie systematicky vyhledali práce týkající se využitelnosti Testu hodin při diagnostice MCI (bez ohledu na použité skórovací systémy a uchopení konceptu MCI). Bylo nalezeno celkem 9 studií z nichž pouze ve 4 případech se skupiny kognitivně normálních starších osob ve skóru získaném z Testu hodin *statisticky lišily*

od pacientů s MCI (!), což podle autorů samo o sobě poukazuje na nízkou validitu metody k diagnostice MCI. Pakliže vyhledané studie uváděly senzitivitu a specifitu, byly tyto údaje vzhledem k zamýšlenému použití „většinou neuspokojivé“ (s. 56) a pohybovaly se mezi 50–80 %, resp. 65–90 % (Ehreke et al., 2009). Zajímavé je, že pouze 2 autorské týmy z celkových 9 doporučují test k popsání účelu používat.

Limity přehledové studie autoři uvádějí již v abstraktu, kde upozorňují na různou operacionalizaci a teoretická východiska MCI, která ještě v nedávné minulosti nebyla mezinárodně uznávaná a široce používaná. Většina studií ovšem respektovala zmíněná Petersonova (2004) kritéria a zakládala se na výsledcích komplexního neuropsychologického vyšetření. Některé práce dokonce dělily pacienty s MCI na běžně rozeznávané subtypy, ale autoři v přehledové studii reportovali jen celkové výsledky za celou skupinu MCI. Dalším problémem pak byla nám již dobře známá nejednotnost skórovacích systémů, jejichž výsledky se, jak jsme viděli v předchozích oddílech naší práce, podstatně liší i v rámci syndromu demence. V diskutované přehledové studii však byla fragmentace skórování ještě výraznější a patrně je to i jeden z důvodů, proč se autoři nepokusili o meta-analytické shrnutí.

3.2.2 Vyhnálek et al. (2016)

Autoři této studie se zaměřili na zajímavé porovnání kvantitativního Cohenova (2000) systému a subjektivního „dojmu“ třech kognitivních neurologů, třech neuropsychologů a šesti neurologů, kteří zaslepeně k diagnóze (pouze se znalostí věku pacienta) řadili hodiny do kategorií „abnormální“, „suspektní“ a „normální“. Výběrový soubor obsahoval pacienty s mírnou demencí při Alzheimerově nemoci (AD), jedince s amnestickou variantou MCI (aMCI) a zdravé kontroly srovnatelného věku a vzdělání. Výsledky poukazují na překvapivě vysokou senzitivitu „subjektivního“ hodnocení („suspektní“ hodiny byly zařazeny spolu s „abnormálními“ do jedné, pozitivní kategorie) – nejlepšího výsledku dosáhli pro AD neuropsychologové (prům. 89,3 %), následovaní kognitivními neurology (79,3 %) a neurology (80,3 %). Pro aMCI pak senzitivita klesla (uvedeno ve stejném pořadí) na 84,0 %; 63,9 % a 61,8 %, specifita vyšla 46,7 %; 61,7 % a 62,8 %. Cohenův komplexní, sedmnáctibodový systém vykázal pro aMCI senzitivitu 69,4 % a specifitu 62,1 % (pro AD byla senzitivita 92,0 % a specifita totožná s tou u aMCI); při přísnějším cut-off skóru byla senzitivita uspokojivých 93,9 %, ale specifita klesla na nepoužitelných 18,2 %. Pro aMCI tedy kvantitativní skórovací systém nepřináší oproti jednoduchému subjektivnímu hodnocení do tří kategorií klinicky cennější výsledky.

3.2.3 Rubínová et al. (2014)

Od stejného týmu pochází i další studie, která se v kontextu aMCI zaměřila na tři skórovací systémy: Shulmanův, Cohenův a Babinsův (dva poslední systémy autoři nazývají komplexními). Rubínová et al. (2014) aplikovali zaslepené hodnocení hodin a kromě výpočtu senzitivity a specifity použili k hodnocení metody i logistické regresní modely. Ty označily jako signifikantní prediktory pouze Cohenův a Babinsův systém (spolu se screeningovou zkouškou MMSE); Shulmanův systém, věk, vzdělání ani pohlaví signifikantní nebyly. Výsledky se ukazují být pro diagnostiku aMCI opět neuspokojivé – Cohenův systém sice zachytil 87,5 % pacientů, ale správně označil pouze 12,5 % zdravých kontrol (!). Podobně systém Babinsův, který byl vyvinut speciálně pro MCI, dosáhl relativně vysoké senzitivity 81,3 %, ale specifitu měl pouze 27,1 %.

3.2.4 Ricci et al. (2016)

Další relativně recentní práce, která se nedostala do systematické přehledové studie z roku 2009, sestává ze tří fází, v jejichž rámci byl vyvinut nový skórovací systém hodnotící specifické chyby v oblasti kružnice, čísel a ručiček. Autoři svůj systém podrobili i faktorové analýze, která prokázala šestici faktorů vysvětlujících 72,8 % variance. Vzniklý systém má pak tři cluster: „ommission“, „position“ a „code“, které se jednotlivě hodnotí pětibodovou Likertovou škálou a poukazují na chyby vynecháním elementů, jejich umístěním a zaznamenáním (např. zakroužkováním čísla místo vyznačení času ručičkami, vepsáním času v digitálním formátu apod.). Senzitivita vůči aMCI vyšla 76 %, specifita 84 %. Pro mírnou demenci při Alzheimerově nemoci pak autoři zjistili velmi uspokojivou senzitivitu 91 % a specifitu 90 %. Na studii jsme nenalezli žádné velké metodologické nedostatky, skupiny autoři definují podle mezinárodně uznávaných kritérií a vývoj testu probíhal postupně, ve třech nezávislých fázích. Jedná se však o zatím nereplikovanou práci.

3.2.5 Rakusa, Jensterle, & Mlakar (2018)

Nejnovější práci hodnotící validitu Testu hodin pro screening MCI přinesli letos slovinští výzkumníci, kteří zahrnuli skupinu aMCI a AD podle platných mezinárodních kritérií, a navíc přidali skupinu pacientů se smíšenou demencí (Alzheimerova nemoc a vaskulární etiologie). Bohužel však k hodnocení výkonu pacientů v Testu hodin použili svůj vlastní, pouze 4bodový systém, který vzhledem k aMCI prokázal nepříliš uspokojivou senzitivitu 69 % [54,58 %; 81,75 %] a specifitu 91 % [77,86 %; 97,41 %]. Mnohem

lepších výsledků – plně v souladu se všemi výše uvedenými poznatky – dosáhli u obou zbývajících skupin, kde byly obě uváděné hodnoty velmi blízko kolem 90 %.

3.3 Reliabilita testu

Překvapilo nás, že jen malá část studií prezentuje informace o reliabilitě skórovacích systémů; některé týmy dokonce využívají pouze jednoho hodnotitele, soustředí se na senzitivitu, specifickou k diagnóze. Lezaková (2012) hovoří o stabilně vysoké shodě hodnotitelů, a to nezávisle na skórovací kritériích či cílových populacích.

3.3.1 Savaş, Yerlikaya, & Yener (2018)

Recentní zhodnocení reliability (a to včetně test-retestové, kterou reportuje jen mizivé procento analyzovaných prací) provedli turečtí výzkumníci. Výhodou studie je fakt, že se jedná o normativní projekt, a tak je vzorek reprezentativní a pečlivě vyšetřený, na druhou stranu je těžištěm práce soubor zdravých starších osob. Test-retestová reliabilita byla proto spočítána na 50 zdravých jedincích, kde (po průměrně 12 měsících) vyšla (vyjádřená vnitřním korelačním koeficientem – ICC) pro Shulmanův systém 72,0 % [55,4 %; 83,1 %] pro méně známý Manosův-Wuův systém pak 80,6 % [68,1 %; 88,5 %]. Jiným typem reliability je shoda hodnotitelů, která u podobných metod nabývá velkého významu. Shoda dvou hodnotitelů pro Shulmanův systém vyšla 79,0 % [62,5 %; 88,8 %] pro zdravé starší osoby 95,0 % [89,7 %; 97,6 %] pro MCI a 94,8 % [81,4 %; 98,0 %] pro AD, což jsou velmi uspokojivé hodnoty.

3.3.2 Mazancova, Nikolai, Stepankova, Kopecek, & Bezdicek (2016)

Autoři zkoumali shodu tří hodnotitelů napříč velkou skupinou zdravých kontrol a pacientů s *neamnestickou* MCI (naMCI). Výkony participantů hodnotili pomocí jednoduššího Shulmanova systému a dvou komplexních skórovacích kritérií – Babinsova a Cohenova. Výsledky hovoří pro velmi dobrou shodu hodnotitelů u všech tří systémů, které spolu navzájem korelují: ve skupině starších zdravých osob dosáhl Shulmanův systém ICC = 0,812, pro Babinsův 0,897 a Cohenův 0,857. Ve skupině naMCI vyšly koeficienty ICC (uvedeno ve stejném pořadí) 0,733; 0,845 a 0,920. Zajímavostí, kterou jiné týmy vůbec nezmiňují, je zhodnocení vnitřní konzistence, jež pro oba komplexní systémy (Shulmanův systém nemá více položek) vychází velmi nízká (v rozmezí Cronbachovy α od 0,069 do 0,446). Osobně se však ptáme, nakolik má u této metody hodnocení vnitřní konzistence smysl (hodiny mohou mít pouze izolovaný deficit).

4. Závěr

Výsledné ohodnocení metody: Pro screening demence 2, pro screening MCI 4.

Test hodin je velmi snadný pro administraci a zabere přibližně 2 minuty, zároveň je pacienti dobře přijímán. Test hodin vidíme jako vhodný doplněk neuropsychologických baterií typu UDS (Uniform Data Set) nebo screeningového testu MMSE (Mini Mental State Examination) a v takovém případě bychom jej ohodnotili i známkou 1, nedoporučujeme jej však používat samostatně jako jedinou metodu, proto tedy známka pro test samotný jakožto screening demence je 2. Nicméně jako rychlé orientační zhodnocení mentálního stavu pacienta může posloužit.

Jako hlavní nedostatek testu vidíme jeho nejednotnost a existenci mnoha skórovacích systémů, což výrazně znesnadňuje jakékoliv snahy o ustanovení psychometrických vlastností tohoto testu. Vždy se dá pouze hovořit o validitě a reliabilitě konkrétního skórovacího systému. Navíc ani podobnější skórovací systémy nejsou s to zachytit všechny typy chyb a selhávání v testu, mnohé z nich lze zachytit pouze kvalitativním hodnocením či při pozorování pacienta při výkonu testu. Avšak pro dobrou shodu posuzovatelů jsou skórovací systémy nutné.

Zatímco v anglofonní vědecké literatuře se těší velké oblíbenosti skórovací systém Shulmana či Sunderlanda, v české neuropsychologii jsou preferovány méně známé, avšak novější a detailnější skórovací systémy Cohena a Babinse. Z různých studií všechny tyto systémy vychází s uspokojivou senzitivitou a specificitou pro screening demence, avšak meta-analýzy, jakožto vyšší forma důkazů, chybí. Podařilo se nám dohledat pouze dvě takové studie, z nichž jedna měla navíc, dle nás, závažné metodologické nedostatky.

Co se týče reliability testu, studie mají tendenci se jí příliš nevěnovat, popřípadě pouze inter-rater reliabilitě, která bývá u zmíněných systémů uspokojivá. Vzhledem však ke specifickým diagnostickým skupin (zhoršování stavu kognice v čase, fluktuace stavu kognice a podobně), kterým je test určen, je obtížné pro test hodin stanovit test-retestovou reliabilitu, dává smysl jí měřit pouze u kontrolních skupin. O tom, zdali je vhodné pro test hodin zjišťovat vnitřní konzistenci, respektive takovou konzistenci očekávat, máme rovněž pochyby.

Pro MCI se test neprokázal být dostatečně senzitivním, pro pacienty pouze s mírnou poruchou kognice je příliš jednoduchý a patrně zde dochází k efektu stropu. Test tedy není schopen spolehlivě detekovat MCI.

Do budoucna vidíme jako žádoucí více sjednotit užívané systémy pro hodnocení testu hodin, ideálně ve prospěch detailnějších skórovacích systémů, a používat jej samostatně pouze pro screening demence, nikoliv MCI.

5. Seznam použité literatury

- Cohen, M. J., Ricci, C. A., Kibby, M. Y., & Edmonds, J. E. (2000). Developmental Progression of Clock Face Drawing in Children. *Child Neuropsychology*, 6(1), 64–76. [https://doi.org/10.1076/0929-7049\(200003\)6:1;1-b;ft064](https://doi.org/10.1076/0929-7049(200003)6:1;1-b;ft064)
- Ehreke, L., Luppá, M., König, H.-H., & Riedel-Heller, S. G. (2010). Is the Clock Drawing Test a screening tool for the diagnosis of mild cognitive impairment? A systematic review. *International Psychogeriatrics*, 22(1), 56–63. <https://doi.org/10.1017/S1041610209990676>
- Hazan, E., Frankenburg, F., Brenkel, M., & Shulman, K. (2018). The test of time: a history of clock drawing. *International Journal Of Geriatric Psychiatry*, 33(1), e22–e30. <https://doi.org/10.1002/gps.4731>
- Juby, A., Tench, S., & Baker, V. (2002). The value of clock drawing in identifying executive cognitive dysfunction in people with a normal Mini-Mental State Examination score. *CMAJ: Canadian Medical Association Journal*, 167(8), 859–864.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford: Oxford University Press.
- Mainland, B. J., Amodeo, S., & Shulman, K. I. (2014). Multiple clock drawing scoring systems: simpler is better. *International Journal of Geriatric Psychiatry*, 29(2), 127–136. <https://doi.org/10.1002/gps.3992>
- Mazancová, A. (2014). *Test hodin: normativní studie na české populaci ve stáří (dipломová práce)*. Retrieved from: <https://is.cuni.cz/webapps/zzp/detail/136353/>
- Mazancova, A. F., Nikolai, T., Stepankova, H., Kopecek, M., & Bezdicek, O. (2016). The Reliability of Clock Drawing Test Scoring Systems Modeled on the Normative

- Data in Healthy Aging and Nonamnesic Mild Cognitive Impairment. *Assessment*, 24(7), 945–957. <https://doi.org/10.1177/1073191116632586>
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrookes Cognitive Examination Revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21(11), 1078–1085. <https://doi.org/10.1002/gps.1610>
- Mittal, C., Gorthi, S., & Rohatgi, S. (2010). Early Cognitive Impairment: Role of Clock Drawing Test. *Medical Journal Armed Forces India*, 66(1), 25–28. [https://doi.org/10.1016/S0377-1237\(10\)80087-5](https://doi.org/10.1016/S0377-1237(10)80087-5)
- MoCA Montreal - Cognitive Assessment. (2018). Retrieved October 18, 2018, from <https://www.mocatest.org/wp-content/uploads/2015/tests-instructions/MoCA-Instructions-Czech.pdf>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nikolai, Štěpánková, H., Vyhnálek, M., & Kopeček, M. (2016). Neuropsychologická Diagnostika Kognitivního Deficitu Ve Stáří. *Československá Psychologie*, 60(5), 525–541.
- Palsetia, D., Rao, G. P., Tiwari, S. C., Lodha, P., & De Sousa, A. (2018). The Clock Drawing Test versus Mini-mental Status Examination as a Screening Tool for Dementia: A Clinical Comparison. *Indian Journal of Psychological Medicine*, 40(1), 1–10. https://doi.org/10.4103/IJPSYM.IJPSYM_244_17

- Park, J., Jeong, E., & Seomun, G. (2018). The clock drawing test: A systematic review and meta-analysis of diagnostic accuracy. *Journal Of Advanced Nursing*. <https://doi.org/10.1111/jan.13810>
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256(3), 183–194. <https://doi.org/10.1111/j.1365-2796.2004.01388.x>
- Rakusa, M., Jensterle, J., & Mlakar, J. (2018). Clock Drawing Test: A Simple Scoring System for the Accurate Screening of Cognitive Impairment in Patients with Mild Cognitive Impairment and Dementia. *Dementia & Geriatric Cognitive Disorders*, 45(5/6), 326–334. <https://doi.org/10.1159/000490381>
- Ressner, P., & Ressnerová, E. (2005). Test hodin, přehledná informace a zhodnocení škal dle Shulmana, Sunderlanda a Hendriksena. *Neurologie pro Praxi*, 3(6), 316–322.
- Riedel, O., Klotsche, J., Förstl, H., & Wittchen, H.-U. (2013). Clock Drawing Test: Is It Useful for Dementia Screening in Patients Having Parkinson Disease With and Without Depression? *Journal Of Geriatric Psychiatry And Neurology*, 26(3), 151–157. <https://doi.org/10.1177/0891988713490994>
- Ricci, M., Pigliautile, M., D'Ambrosio, V., Ercolani, S., Bianchini, C., Ruggiero, C., ... Mecocci, P. (2016). The clock drawing test as a screening tool in mild cognitive impairment and very mild dementia: A new brief method of scoring and normative data in the elderly. *Neurological Sciences*, 37(6), 867–873. <https://doi.org/10.1007/s10072-016-2480-6>
- Rubínová, E., Nikolai, T., Marková, H., Šiffelová, K., Laczó, J., Hort, J., & Vyhnálek, M. (2014). Clock Drawing Test and the diagnosis of amnesic mild cognitive impairment: Can more detailed scoring systems do the work? *Journal of Clinical &*

- Experimental Neuropsychology*, 36(10), 1076–1083.
<https://doi.org/10.1080/13803395.2014.977233>
- Savaş, D. D., Yerlikaya, D., & Yener, G. G. (2018). Validity, Reliability and Turkish Norm Values of the Clock Drawing Test for Two Different Scoring Systems. *Turkish Journal Of Neurology*, 24(2), 143–152. <https://doi.org/10.4274/tnd.26504>
- Shulman, K. I. (2000). Clock-drawing: is it the ideal cognitive screening test? *International Journal of Geriatric Psychiatry*, 15(6), 548–561. [https://doi.org/10.1002/1099-1166\(200006\)15:6<548::AID-GPS242>3.0.CO;2-U](https://doi.org/10.1002/1099-1166(200006)15:6<548::AID-GPS242>3.0.CO;2-U)
- Shulman, K. I., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, 1(2), 135–140. <https://doi.org/10.1002/gps.930010209>
- Souillard-Mandar, W., Davis, R., Rudin, C., Au, R., Libon, D. J., Swenson, R., ... Penney, D. L. (2015). Learning classification models of cognitive conditions from subtle behaviors in the digital Clock Drawing Test. *Machine Learning*, 102(3), 393–441. <https://doi.org/10.1007/s10994-015-5529-5>
- Sunderland, T., Hill, J. L., Mellow, A. M., Lawlor, B. A., Gundersheimer, J., Newhouse, P. A., & Grafman, J. H. (1989). Clock Drawing in Alzheimers Disease. *Journal of the American Geriatrics Society*, 37(8), 725–729. <https://doi.org/10.1111/j.1532-5415.1989.tb02233.x>
- Štěpánková, H., Nikolai, T., Lukavský, J., Bezdíček, O., Vrajobá, M., & Kopeček, M. (2015). Mini-Mental State Examination – česká normativní studie. *Česká a Slovenská Neurologie a Neurochirurgie*, 78(1), 57–63.
- Tsoi, K. K. F., Chan, J. Y. C., Hirai, H. W., Wong, S. Y. S., & Kwok, T. C. Y. (2015). Cognitive Tests to Detect Dementia: A Systematic Review and Meta-analysis.

JAMA Internal Medicine, 175(9), 1450–1458.

<https://doi.org/10.1001/jamainternmed.2015.2152>

Vyhnálek, M., Rubínová, E., Marková, H., Nikolai, T., Laczó, J., Andel, R., & Hort, J. (2016). Clock drawing test in screening for Alzheimers dementia and mild cognitive impairment in clinical practice. *International Journal of Geriatric Psychiatry*, 32(9), 933–939. <https://doi.org/10.1002/gps.4>