

Abstrakt

Úvod: U většiny funkčních testů, které se používají v rehabilitaci a mezi něž řadíme i testy posturální stabilizace (TPS), není k dispozici zlatý standard ani jiná objektivizační kritéria. V souladu s požadavky Evidence Based Practice/Medicine je tedy žádoucí pokusit se stanovit shodu mezi hodnotiteli a shodu hodnotitele se sebou samým jako první krok v procesu objektivizace.

Cíle: Prostřednictvím dvojice zkušených a dvojice nezkušených hodnotitelů stanovit inter- a intra-rater reliabilitu u tří baterií TPS, vypracovat podrobnou popisnou statistiku a porovnat naše výsledky s literaturou.

Materiál a metody: Byly vybrány tři baterie aspekčních TPS zaměřených na bederní páteř (Kolář 2006, Luomajoki et al. 2007, Tidstrand a Horneij 2009). Provedení testů bylo nezávisle hodnoceno současně dvěma zkušenými a dvěma nezkušenými hodnotiteli na smíšeném souboru zdravých osob ($n = 15$, 10 žen, průměrný věk $44,5 \pm 11,0$ let)

a pacientů s vertebrogenním algickým syndromem bederní páteře bez neurologického deficitu a neurochirurgického výkonu na páteři v posledních pěti letech ($n = 16$,

8 žen, průměrný věk $45,9 \pm 13,6$ let). Hodnocení testů předcházela jejich formalizace (pozitivita jednotlivých znaků, binární hodnocení celkové pozitivity testu). Ke stanovení reliability jsme použili procentní shodu ($a\%$) a koeficient kappa podle Cohena (κ).

Výsledky: Zkušení hodnotitelé častěji hodnotí testy jako pozitivní. Rozložení relativních četností se v některých případech vylučuje s použitím koeficientu κ . U všech testových baterií konstatujeme velmi dobré ($\kappa \geq 0,80$), podstatné ($0,60 \leq \kappa < 0,80$) a dostatečné ($0,40 \leq \kappa < 0,60$) shody mezi hodnotiteli. U zkušených hodnotitelů nacházíme vyšší počet významných shod hodnotitele se sebou samým ($a\%$) než u nezkušených.

Závěr: Potvrdilo se naše očekávání, že shoda významně závisí na zkušenosti hodnotitele a že shoda hodnotitele se sebou samým dosahuje vyšší spolehlivosti než shoda mezi hodnotiteli. Předpoklad, že výsledky testů v dané formalizaci a standardizaci budou rozlišovat mezi pacienty a kontrolami, se nepotvrdil. Srovnání našich výsledků s literaturou je uspokojivé.

Klíčová slova

posturální stabilizace, funkční testy, shoda hodnotitele se sebou samým, shoda mezi hodnotiteli, reliabilita

Abstract

Introduction: In most functional tests used in physical therapy, including the tests of postural stabilization (TPS), there is no gold standard. In compliance with the Evidence Based Practice/Medicine requirements, an attempt at setting inter-rater and intra-rater reliability as the first step towards objectivization is rather desirable.

Objectives: To evaluate inter-rater and intra-rater reliability in three different TPS through a couple of experienced and inexperienced examiners, carry out detailed descriptive statistics and compare our results with the literature.

Material and methods: Twelve different TPS have been selected, all focused on lumbar spine (Kolář 2006, Luomajoki et al. 2007, Tidstrand and Horneij 2009). The test performance was independently rated by two experienced and two inexperienced therapists on a mixed sample of healthy persons ($n = 15$, 10 women, average age 44.5 ± 11.0 years) and patients suffering from chronic low back pain without neurologic deficit and without neurosurgical operation on the spine in the last 5 years ($n = 16$, 8 women, average age 45.9 ± 13.6 years). The evaluation of the tests was preceded by their formalization. Cohen's kappa coefficients (κ) and percentage agreements were calculated.

Results: Experienced raters score more often tests as positive. In some cases the prevalence of positive findings is incompatible with the κ coefficient. In all the tests we found good ($\kappa \geq 0.80$), substantial ($0.60 \leq \kappa < 0.80$) and moderate ($0.40 \leq \kappa < 0.60$) inter-rater reliability. There is a higher number of substantial percentage agreement in more experienced evaluators.

Conclusion: Our expectation that agreement depends strongly on the experience of the evaluator and that intra-rater reliability is higher than inter-rater reliability has been confirmed. Our expectation that the test results under given formalization and standardization would differ between patients and controls has not been confirmed. The comparison of our results with literature is satisfactory.

Keywords

postural stabilization, functional tests, intra-rater reliability, inter-rater reliability