ORIGINAL ARTICLE

# Reliability of Visual Temporal Thresholds

*LN Brown, M Eliasziw, LM Metz*

**ABSTRACT:** ***Background:*** Visual processing deficits involving temporal characteristics are typically not captured by the widely used outcome measures (i.e., Expanded Disability Status Scale, Multiple Sclerosis Functional Composite) in multiple sclerosis (MS). Visual temporal thresholds (i.e., measurements of the temporal aspects in visual processing) are typically significantly higher (i.e., prolonged) in MS patients when compared to controls. The test-retest reliability of these thresholds was examined in patients with MS. ***Methods:*** Visual temporal thresholds were measured in 21 stable MS patients during two separate test sessions. Test-retest reliability and the standard error of measurement were calculated. The threshold of change in visual temporal thresholds in MS patients that would correspond to real change beyond measurement error with 95% certainty was also calculated. For comparisons, a control group (n = 10) was included. ***Results:*** The test-retest reliability of this measure of visual temporal thresholds was 0.97. The threshold indicating change beyond chance or measurement error with 95% certainty was 11 ms. Higher thresholds were significantly correlated with longer durations of disease. ***Conclusions:*** This measure of visual temporal thresholds has excellent test-retest reliability and a change of greater than 11 ms is highly likely to represent real change in MS patients. The findings indicate that these measurements may provide useful clinical information about functional changes regarding the temporal aspects of the visual system, which is currently not captured by the Extended Disability Status Scale.

**RÉSUMÉ: Fiabilité des seuils visuels temporaux.** ***Contexte :*** Les déficits du traitement de l'information visuelle ayant des caractéristiques temporales ne sont généralement pas détectés par les mesures communément utilisées (Expanded Disability Status Scale, Multiple Sclerosis Functional Composite) chez les patients atteints de sclérose en plaques (SEP). Les seuils visuels temporaux (c'est-à-dire les mesures des aspects temporaux du traitement de l'information visuelle) sont significativement plus élevés (c'est-à-dire prolongés) chez les patients atteints de SEP par rapport à des témoins. Nous avons évalué la fiabilité du test-retest de ces seuils chez des patients atteints de SEP. ***Méthodes :*** Les seuils visuels temporaux ont été mesurés au cours de deux sessions chez 21 patients atteints de SEP qui étaient stables. La fiabilité test-retest et l'erreur standard de la mesure ont été calculées. Le seuil de changement des seuils visuels temporaux chez les patients atteints de SEP qui correspondrait à un changement réel au-delà de l'erreur de la mesure avec une certitude de 95% a été calculé. À titre de comparaison, un groupe témoin (n = 10) a été inclus. ***Résultats :*** La fiabilité test-retest de cette mesure des seuils visuels temporaux était de 0,97. Le seuil indiquant un changement au-delà du hasard ou d'une erreur de mesure avec une certitude de 95% était de 11 ms. Des seuils plus élevés étaient corrélés significativement à une durée plus longue de la maladie. ***Conclusions :*** Cette mesure des seuils visuels temporaux a une excellente fiabilité test-retest et il existe une forte probabilité qu'un changement de plus de 11 ms représente un changement réel chez les patients atteints de SEP. Selon nos observations, ces mesures peuvent fournir une information utile en clinique sur les changements fonctionnels des aspects temporaux du système visuel, ce qui n'est pas détecté actuellement par le Extended Disability Status Scale.

Can. J. Neurol. Sci. 2007; 34: 433-437

Multiple sclerosis (MS) commonly affects the visual system causing visual dysfunction at some stage in 80% of patients.[1,2] Deficits in visual processing that involve temporal characteristics, contrast sensitivity, and color discrimination are typically not captured by the Expanded Disability Status Scale (EDSS), which is a widely used outcome measure in MS. Moreover, the Multiple Sclerosis Functional Composite (MSFC), which is a new outcome measure in MS, does not include a dimension of visual abilities. Given that visual dysfunction is a common cause of disability in MS,[3] it seems prudent to include measurements of visual dysfunction in routine clinical practice.[3,4] Not surprisingly, it has been suggested that measurements of visual dysfunction in MS may provide relevant information about disease progress.[4]

Several studies have employed psychophysical methods and have demonstrated deficits in MS patients for visual tasks that involve temporal aspects. For instance, in the critical flicker frequency task, MS patients will typically perceive a flickering

From the Department of Clinical Neurosciences (LNB, LMM), Department of Community Health Sciences (ME), University of Calgary, Calgary, AB, Canada.
Received October 12, 2006. Accepted in final form June 11, 2007.
*Reprint requests to:* Lenora N. Brown, C/O Luanne M. Metz, Department of Clinical Neurosciences, Foothills Medical Centre, 1403 29 Street N.W., Calgary, Alberta, T2N 2T9, Canada.

light that increases in its flickering speed as continuous sooner than controls.[5-7] Double-flash resolution tests, where the interval between two flashes of light is decreased until the patient reports one flash, typically demonstrates abnormalities in thresholds of MS patients.[8-11] Impairment in visual temporal tests involving simultaneity vs nonsimultaneity judgments to the onset of pairs of lights have also been reported in MS patients.[12] Visual dysfunction in MS patients has also been demonstrated in color discrimination, low-contrast letter acuity, and contrast sensitivity tasks.[3,4,13] Impaired performance in MS patients for these visual tests has been demonstrated, even in patients with static visual acuities of 20/20.[14] These overall findings suggest that demyelination or neuronal degeneration occurs in both the parvocellular pathway (i.e., color, acuity) and the magnocellular pathway (i.e., temporal characteristics). In fact, Caruana et al,[15] reported a correlation between magnetic resonance imaging (MRI) lesions of the postchiasmal visual pathway and psychophysical thresholds, suggesting that postchiasmal morphological changes have functional effects.

The results of our previous work[12,16,17] suggest that sensory temporal thresholds may reflect the integrity of neural pathways. Sensory temporal thresholds represent the time interval between the onset of two sensory stimuli when they are judged to be simultaneous, and thus may be a functional measurement of central nervous system integrity. We recently reported that relative to controls, tactile and visual temporal thresholds were prolonged in individuals with MS.[12] It was also found that tactile temporal thresholds not only had excellent test-retest reliability, but that a change in threshold of more than 19 ms is highly likely to represent real change with 95% certainty.[17]

Despite the reported impairment in visual temporal processing in MS patients, the test-retest reliability of temporal measurements has yet to be determined. Moreover, it is unknown how much observed change in these visual temporal thresholds is considered to be real change as opposed to measurement error.

The aim of the present study was to assess the test-retest reliability of visual temporal thresholds in a group of relatively stable MS patients. Although reliability is typically expressed in terms of intraclass correlational coefficients (ICCs), this coefficient fails to provide information about whether observed change represents true change or change due to measurement error associated with the instrument. Thus, the present study also calculated the magnitude of change in visual temporal thresholds in MS patients that would correspond to real change with 95% certainty.

## Methods

### *Participants*

A total of 21 MS patients (7 males, 14 females) were tested. There were 17 patients with relapsing-remitting MS (RRMS), 3 with secondary-progressive MS (SPMS), and 1 with primary-progressive MS (PPMS). The ages ranged from 26 to 57 years of age (M = 45.67, SD = 8.41). The range of the Kurtzke EDSS for 19 of the patients (a recent EDSS was unavailable for 2 patients) ranged from 1.0 to 6.5 with a mean score of 3.11 (SD = 1.51). Patients were diagnosed by a neurologist as defined by standard clinical and neuroimaging criteria[18] at a Canadian MS clinic. In order to compare reliability results of visual temporal thresholds for the MS patients, a control group was included. The control group consisted of eight females and two males. The age of the control group ranged from 22 to 52 years of age (M = 38.20, SD = 9.98).

Exclusion criteria for all patients included substance abuse, use of sedating medications, occurrence of an MS exacerbation within the previous 30 days and visual impairment (cataracts, glaucoma, current episode of optic neuritis). Three patients had previous optic neuritis. Control participants were excluded if they had ever been under a doctor's care for a neurological or psychiatric illness. All control participants reported to be in good health, which included self-reporting of visual disturbances. The study was approved by the local Institutional Review Board and informed consent was obtained from all patients before testing.

### *Visual Apparatus*

The visual stimuli consisted of six red light emitting diodes (LEDs) which were placed on a horizontal plane against a white background. As shown in Figure 1, the LEDs were equally spaced at 33 mm (center to center) and symmetrically arranged with respect to the center of the visual display. A black cross in the center of the screen served as the central fixation point. Three LEDs were positioned on each side of the black cross for a total of five paired locations. The LO (left outer) pair consisted of LEDs 1 and 2, the LI (left inner) pair consisted of LEDs 2 and 3, the BIL (bilateral; involving both visual half-fields) pair consisted of LEDs 3 and 4, the RI (right inner) pair consisted of LEDs 4 and 5, and the RO (right outer) pair consisted of LEDs 5 and 6.

A chin rest was fixed at a viewing distance of 57 cm so that each cm of the display subtended 1° at the participant's eyes. The mean light intensity emitted by the LEDs was 122.6 cd/m$^2$ (range 115 to 150 cd/m$^2$), while the mean luminance of the visual display was approximately 45 cd/m$^2$. A video camera was mounted at the rear of the display facing the participant. The lens of the video camera was aligned with a small hole in the visual display that was just above the central fixation point. A monitor connected to the camera was used to observe the participant's eyes in order to ensure that central fixation was maintained.

### *Parameter Estimation Procedures*

A modified version of the Parameter Estimation by Sequential Testing (PEST) algorithm, which allows one to determine the

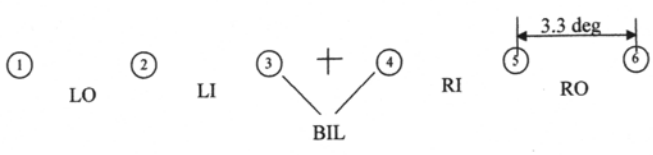


***Figure 1:*** *Illustration of the six locations of LEDs and the six locations of the pairs of spatially separated LEDs. Deg = degree; LO = left outer; LI = left inner; BIL = bilateral first pair; RI = right inner; RO = right outer.*



https://doi.org/10.1017/S0317167100007319 Published online by Cambridge University Press

value of thresholds in as few trials as possible, was used to adjust the millisecond intervals separating the onset of the pairs of stimuli.[19,20] Using a binary forced-choice (YES/NO) paradigm, the interval between the onset of the first stimulus to the onset of the second stimulus (Stimulus Onset Asynchrony, SOA) was adjusted (increased or decreased) according to the participant's previous response. Participants responded as to whether they perceived the onset of the stimulus pairs as simultaneous (YES response) or non-simultaneous (NO response). The interval between the onset of stimulus pairs increased for "YES" responses and decreased for "NO" responses. The program terminated when the SOA step size reached 1 ms. The initial SOA step size was set at 10 ms, while the termination step size was set 1 ms. The temporal threshold was the SOA value upon termination for each sequence.

### *Procedure*

Each test session consisted of a questionnaire to acquire general background information, the Landolt C Eyechart to measure near binocular visual acuity (corrected vision), a short visual screening test, a practice block of trials, and three blocks of test trials. The questionnaire consisted of questions regarding overall health, handedness, and medication use.

Participants were instructed to focus on the fixation point (cross) with their chin resting on the chin rest. The visual screening test consisted of 20 trials that required participants to correctly identify the LED that was illuminated (i.e., LEDs were labeled with numbers for identification). For the visual temporal task, participants were advised that two of the six lights would be illuminated on each trial and to respond "YES" if they judged the illumination onsets to be simultaneous and "NO" if they judged the illumination onsets to be different. The order in which the five pairs of LEDs (i.e., LO, LI, BIL, RI, and RO) and the direction of illumination (left before right, right before left) were randomly interwoven within each block of trials. Onsets differed but the offsets for each LED pair were simultaneous. The pair of LEDs remained illuminated for a duration of 1000 ms once both LEDs were illuminated.

Each test session began with a practice block of trials during which the smallest SOA value (the value at which a "NO" response could confidently be established) was estimated (nonsimultaneity). Fifteen ms were added to this estimated value which served as an initial SOA value for the four test blocks of trials. Therefore, the initial SOA value for the commencement of the test blocks of trials was individualized for each participant. Verbal responses were entered into the computer which triggered the delivery of the next pair of stimuli.

To facilitate the assessment of test-retest reliability, each participant underwent the complete test session within a two week period with the exception of five patients who were tested within three to five weeks. Although one patient completed the test session within 20 weeks, the mean threshold for each test session was comparable (i.e., 55.20 ms and 54.23 ms for tests 1 and 2 respectively). All patients reported no MS-related change during the second test session. All control participants were tested within a two week period with the exception of one participant who was tested within six weeks.

### *Statistical Method*

For each test session, the mean visual temporal threshold was calculated for each patient from the three blocks of trials of stimulus presentations for the five paired presentations. The data were analyzed by calculating the ICC and the standard error of measurement (SEM) for the two repeated sessions of visual temporal thresholds. The SEM is defined as measurement error and refers to inherent variability of measurements of the same quantity within the same individual.[21,22] A simple equation was used to calculate the SEM from the two sessions per subject.

The threshold value (δ) that represents real change beyond measurement error (i.e., repeatability[9]) with 95% certainty was calculated as $\sqrt{2}$ ? 1.96 x SEM.[22] Observed changes within the range of –δ to δ represents change that is within the realm of measurement error. In contrast, an observed change between two test periods that is greater than δ or less than –δ represents change exceeding the variability of the measurement instrument itself. In this case, one can be 95% certain that real change has occurred. As an additional step to assessing the variability of the measurements, a rank correlation coefficient was calculated to determine whether the variability in the test-retest measures increased with the magnitude (mean values) of visual temporal thresholds.

## RESULTS

Figure 2 displays the overall mean visual temporal threshold for the MS patients and the controls for each test session. Figure 3 is a Bland-Altman plot, showing the magnitude of the differences between Test 1 and Test 2 versus the mean of Test 1 and Test 2 for the MS patients. The differences appear as a random scatter about the mean difference of -2.30, shown as the solid middle horizontal line. The calculated rank correlation coefficient indicated that the variability in the repeat measures was independent of the magnitude of the temporal threshold ($r$ = .27, $p$ = .09). The ICC was .97 for the MS group data, indicating almost perfect test-retest reliability. The SEM of 3.97 was small

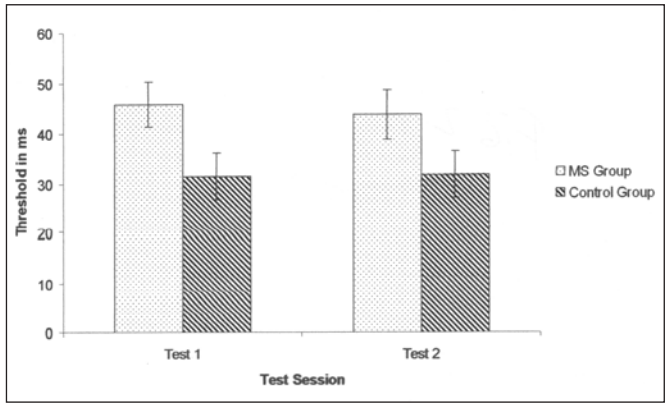


***Figure 2:*** *Mean overall temporal threshold and standard error bars for the MS and Control Group as a function of test session.*



https://doi.org/10.1017/S0317167100007319 Published online by Cambridge University Press

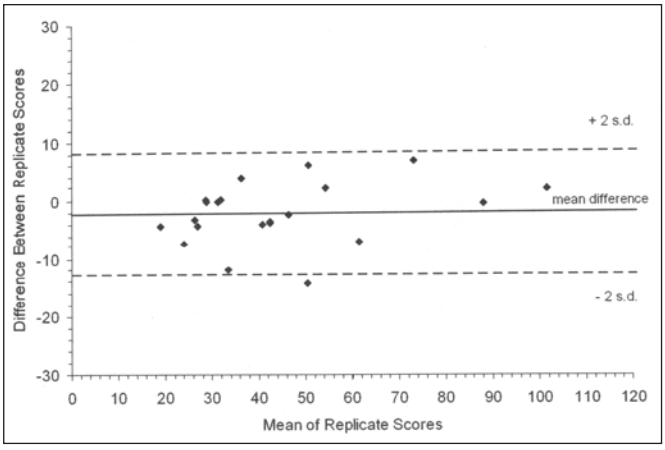


***Figure 3:*** *Test-retest data for patients. Bland-Altman plot showing the difference in mean thresholds for tests 1 and 2 as a function of the mean visual temporal thresholds for tests 1 and 2.*

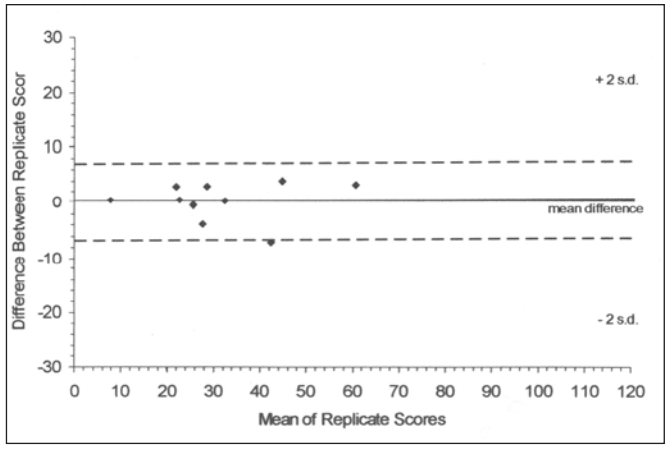


***Figure 4:*** *Test-retest data for controls. Bland-Altman plot showing the difference in mean thresholds for tests 1 and 2 as a function of the mean visual temporal thresholds for tests 1 and 2.*

when compared to the mean threshold of 44.64 ms, calculated from the 42 tests. The threshold value of real change (δ) was 11.00 ms. Thus, one can be 95% certain that an observed change on two occasions is real beyond chance if it is greater than 11 ms.

For comparison, the ICC for the control group data was .97. The SEM of 2.32 was small when compared to a mean threshold of 31.42 ms, for the 20 tests. The threshold value of real change (δ) was 6.44 ms. Figure 4 shows the magnitude of the differences between Test 1 and Test 2 versus the mean of Test 1 and Test 2 for the controls.

For the MS group, correlational analyses were undertaken to determine whether binocular visual acuity scores (Tests 1 and 2), EDSS, duration of disease (months), and age were related to visual temporal thresholds of Test 1 and Test 2. The duration of disease was significantly correlated with both Test 1 ($r = .53$, $p = .01$), and Test 2 ($r = .47$, $p = .03$). Acuity scores, EDSS, and age did not correlate with overall visual temporal thresholds for either Test 1 or Test 2.

## Discussion

This study demonstrated almost perfect test-retest reliability of visual temporal thresholds in a group of relatively stable MS patients, as assessed by an ICC. In addition, if the observed change in visual temporal thresholds in MS patients from one occasion to another exceeds 11 ms, one can be 95% certain that real change has occurred beyond measurement error. Finally, the results confirmed that the variability in threshold values was statistically independent of the magnitude of the thresholds. The results indicated a small 2 ms improvement in visual thresholds for the MS patients on test two when compared to test one.

Although the sample size in the current study is small, the lack of correlation between visual temporal thresholds and EDSS scores are congruent with our previous findings such that no correlation between these two measures was found.[12] It should be noted however, that the MS patients in the present study were relatively stable and they did not have appreciable visual dysfunction, which may possibly have contributed to the lack of correlation between visual temporal thresholds and EDSS measurements. Taken together, these results suggest that MS patients demonstrate deficits in the ability to time visual information, which is currently not captured by the EDSS. These findings also demonstrate the need to evaluate the value of including a visual functioning component, especially one that captures the temporal characteristics of visual processing, to the MSFC.

The amount of change in overall visual temporal thresholds between tests for the MS patients ranged from .10 ms to 14.37 ms, while the amount of change in thresholds between tests for the control group was .20 ms to 7.20 ms. Regardless of presentation location, MS patients demonstrated excellent test-retest reliability. This finding indicates that visual assessment of temporal and nasal retinal regions has good test-retest reliability.

We recently reported that tactile temporal thresholds had excellent test-retest reliability, and that a change in threshold of more than 19 ms was highly likely to represent real change (i.e., 95% confidence).[17] In the present study we demonstrated that visual temporal thresholds also have excellent test-retest reliability and that a change of more than 11 ms is likely to represent real change (i.e. 95% confidence). We have also previously shown a strong correlation between tactile and visual temporal thresholds in MS patients correlation ($r = .85$, $p = .008$).[12] Given that visual temporal thresholds are typically lower than tactile temporal thresholds,[12] it is not surprising that the amount of change that is likely to represent real change is lower for visual thresholds relative to tactile thresholds.

Future investigation should include optical coherence tomography (OCT) to determine the relationship between the temporal characteristics of visual function and retinal nerve fiber



https://doi.org/10.1017/S0317167100007319 Published online by Cambridge University Press

layer (RNFL) thickness. Optical coherence tomography is a relatively new technique that measures RNFL thickness and is able to assess axonal preservation *in vivo*.[23] Although OCT may be a potential structural biomarker in MS, the relationship between structural retinal changes and visual functioning has not been established in humans. For instance, low-contrast letter acuity, contrast sensitivity, and pattern electroretinogram recordings have been shown to correlate well with RNFL thickness in MS patients,[24,25] while abnormal visual evoked potentials (VEPs) failed to correlate with RNFL thickness, even though there was a significant reduction in RNFL thickness of MS patients who were previously affected by optic neuritis.[25] A possible explanation proposed by the authors for the absence of correlation between VEPs and RNFL thickness may be attributed to both impaired retinal function and to delayed neural conduction in the postretinal pathways.[25] Thus, it is important to establish the relationships between the functioning of the visual afferent pathways (i.e., parvocellular and magnocellular pathways) and retinal structural alterations. With respect to the present study, future research will need to determine the degree to which the delayed neural conduction that is associated with MS, is attributed to retinal structural alterations, impaired retinal functional, afferent pathways to the visual cortex, and/or to the pathways from visual cortex to the language centers as required for verbal responses.

We have previously demonstrated that visual temporal thresholds are significantly higher in MS patients when compared to age-matched controls.[12] These increased thresholds may be due to demyelination, axonal injury and loss, which altered the functioning of the visual system in MS. Taken together with the present results showing near perfect test-retest reliability, these findings indicate that visual temporal thresholds should be considered for further study as a measure of outcome in MS.

## Acknowledgements

This research was supported by the Multiple Sclerosis Society of Canada and by the Natural Sciences and Engineering Research Council of Canada.

## References

1. McDonald WI, Barnes D. The ocular manifestations of multiple sclerosis: abnormalities of the afferent visual system. J Neurol Neurosurg Psychiatry. 1992;55:747–52.
2. Newman NJ. Multiple sclerosis and related demyelinating diseases. In: Miller NB, Newman NJ, editors. Walsh and Hoyt's clinical neuro-ophthalmology, 5th ed. V5, p. 5539-676. Baltimore: Williams & Wilkins; 1998.
3. Balcer LJ, Baier ML, Cohen JA, Kooijmans MF, Sandrock AW, Nano-Schiavi ML, et al. Contrast letter acuity as a visual component for the Multiple Sclerosis Functional Composite. Neurol. 2003;61:1367-73.
4. Baier ML. Low-contrast letter acuity testing captures visual dysfunction in patients with multiple sclerosis. Neurol. 2005;64:992-5.
5. Daley ML, Swank RL, Ellison CM. Flicker fusion thresholds in multiple sclerosis: a functional measure of neurological damage. Arch Neurol. 1979;36:292–5.
6. Mason RJ, Snelgar RS, Foster DH, Heron JR, Jones RE. Abnormalities of chromatic and luminance critical flicker frequency in multiple sclerosis. Invest Ophthalmol Vis Sci. 1982;23:246–52.
7. Titcombe AF, Willison RG. Flicker fusion in multiple sclerosis. J Neurol Neurosurg Psychiatry. 1961;24:260–6.
8. Galvin RJ, Heron JR, Regan D. Subclinical optic neuropathy in multiple sclerosis. Arch Neurol. 1977;34:666–70.
9. Patterson VH, Foster DH, Heron JR, Mason RJ. Multiple sclerosis: luminance threshold and measurements of temporal characteristics of vision. Arch Neurol. 1981;38:687–9.
10. Vleugels L, Charlier A, van Nunen A, Lafosse C, Vogels R, Ketelaer P, et al. Temporal resolution deficits in the visual fields of MS patients. Vis Res. 1999;39:2429–38.
11. Vleugels L, van Nunen A, Lafosse C, Ketelaer P, Vandenbussche E. Temporal and spatial resolution in foveal vision of multiple sclerosis patients. Vis Res. 1998;38:2987–97.
12. Brown LN, Metz LM, Sainsbury RS. Sensory temporal thresholds and interhemispheric transfer times in multiple sclerosis: a preliminary study of a new outcome measure. J Clin Exp Neuropsych. 2003;25(6):783-92.
13. Flanagan P, Markulev C. Spatio-temporal selectivity of loss of colour and luminance contrast sensitivity with multiple sclerosis and optic neuritis. Ophthamol Physiol Opt. 2004;25:57-65.
14. Balcer LJ. Optic Neuritis. N Engl J Med. 2006; 354:1273-80.
15. Caruana PA, Davies MB, Weatherby SJM, Williams R, Haq N, Foster DH, et al. Correlation of MRI lesions with visual psychophysical deficit in secondary progressive multiple sclerosis. Brain. 2000;123:1471–80.
16. Brown LN, Metz LM. Tactile temporal thresholds detect relapse-related changes in multiple sclerosis: a preliminary study. Mult Scler. 2005;11:725-7.
17. Brown LN, Metz LM, Eliasziw M. Identifying reliable change in tactile temporal thresholds in multiple sclerosis: test-retest reliability. Mult Scler. 2006;12:573-7.
18. McDonald WI, Compston A, Edan G, McDonald, WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. Ann Neurol. 2001;50(1):121-7.
19. Findlay JM. Estimates on probability functions: a more virulent PEST. Percept Psychophys. 1978;23:181-5.
20. Pentland A. Maximum likelihood estimation: the best PEST. Percept Psychophys. 1980;28:304-12.
21. Bland JM, Altman DG. Measurement error. BMJ. 1996;313:744.
22. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. Phys Ther. 1994;74:777-88.
23. Sergott RC. Optical coherence tomography: measuring in-vivo axonal survival and neuroprotection in multiple sclerosis and optic neuritis. Curr Opin Ophthamol. 2005;16:346-50.
24. Fisher JB, Jacobs DA, Markowitz CE, Galetta SL, Volpe NJ, Nano-Schiavi ML, et al. Relation of visual function to retinal nerve fiber layer thickness in multiple sclerosis. Ophthalmol. 2006;113(2):324-32.
25. Parisi V, Manni G, Spadara M, Colacino G, Restuccia R, Marchi S, et al. Correlation between morphological and functional retinal impairment in multiple sclerosis patients. Invest Ophthalmol Vis Sci. 1999;40(11):2520-7.