Validation of the "Six Step Spot Test" for gait among patients with multiple sclerosis in Brazil

Karina Pavan¹, Charles Peter Tilbery², Sergio Lianza², Bruna Eriko Matsuda Marangoni¹

ABSTRACT

The clinical manifestations of multiple sclerosis (MS) are variable among patients, and the course of disease is not linear. Different symptoms are presented, with gradual accumulation of disability. These variations are difficult to quantify in clinical practice, and several studies have attempted to create instruments capable of measuring these disabilities. The Six Step Spot Test (SSST) was developed for quantitative evaluation of the lower limbs (LL) over time. Performance in this test reflects the complexity of sensory-motor function, including LL strength, spasticity, coordination and balance, going beyond vision and cognition. The aim of the present study was to validate the SSST in a population of MS patients in Brazil. This prospective study included 75 patients with MS, with EDSS 0 to 6.5 in the study group. Ninety-one healthy subjects were randomly selected for the control group. The results showed that the groups were similar, and that the SSST is a reliable and reproducible test. According to the statistical analysis on the data in this study, the SSST is a valid, reliable and reproducible tool for use in the Brazilian MS patient population.

Key words: gait, multiple sclerosis, disability evaluation, validation studies, reproducibility of results.

Validação do teste de marcha "Six Step Spot Test" em pacientes com esclerose múltipla no Brasil

RESUMO

As manifestações clínicas da esclerose múltipla (EM) são variáveis e o curso da doença não é linear, caracterizado por sintomas, acúmulo progressivo de incapacidades. Estas variações são difíceis de quantificar na prática clínica, motivo pelo qual surgiram diversos estudos para criação de instrumentos capazes de mensurar essas incapacidades. O *Six Step Spot Test* (SSST) foi desenvolvido para avaliar função dos membros inferiores (MMII) de forma quantitativa, mensurado no domínio tempo, avaliando a complexa função sensóriomotora, como força em MMII, espasticidade, coordenação e equilíbrio, além da visão e cognição. Este estudo teve como objetivo validar o SSST para pacientes com EM no Brasil. Estudo prospectivo do qual participaram 75 pacientes com EM e EDSS de 0 a 6,5 no grupo de estudo e 91 indivíduos saudáveis no controle que foram selecionados aleatoriamente. Segundo os dados estatísticos deste estudo, o SSST é válido, confiável e reprodutível para pacientes com EM no Brasil.

Palavras-chave: marcha, esclerose múltipla, avaliação da deficiência, estudos de validação, reprodutibilidade dos testes.

Correspondence

Karina Pavan Av José Giorge 2031 06707-100 Cotia SP - Brasil E-mail: karpav@gmail.com

Received 24 June 2009 Received in final form 7 October 2009 Accepted 20 October 2009 Multiple sclerosis (MS) is a disease that affects the central nervous system (CNS). MS can cause disability, with a consequent

negative influence on the quality of life of patients. Among the disabilities caused by MS, gait abnormalities have been the sub-

Department of Neurology of the Santa Casa Sisters of Mercy Hospital of São Paulo, São Paulo SP, Brazil: ¹Physical Therapist, Santa Casa de Misericórdia de São Paulo, São Paulo SP, Brazil; ²Physician, Santa Casa de Misericórdia de São Paulo, São Paulo SP, Brazil.

ject of several studies, since these are often related to disease progression. Gait assessment in MS is problematic, since most instruments used are not specific for MS, or have not been adapted to the Brazilian population. Therefore, the reliability of these tests is undermined. It is necessary to monitor the progress of patients and to seek scientific evidence based on the various types of therapeutic interventions. Thus, it is important to have gait assessment scales for MS^{1-3} .

In order to obtain reliable measurement of disease progression, as well as a tool for assessment of treatment efficacy with scientific credibility, it is necessary to consider certain psychometric properties using a valid and reliable method⁴.

Nieuwenhuis et al., in Denmark in 2006, developed a test for functional evaluation of the lower limbs that was specific for MS patients. In order to be able to use this method for Brazilian MS patients, the test must be validated⁵.

METHOD

The present prospective study was approved by the Ethics Research Committee of Irmandade Santa Casa de Misericórdia de São Paulo (ISCMSP), and was carried out under protocol number 380/07. The subjects were randomly selected among friends and relatives of patients attending the Neurology Clinic, Department of Rehabilitation. The MS patient group was selected from the MS Care and Treatment Center (*Centro de Atendimento e Tratamento da Esclerose Múltipla*, CATEM).

One hundred and sixty-six participants in this study were divided into two groups, after signing the consent form: a study group (75) and a control group (91). On average, the patients in the study group were aged 40.6 years (SD 11.1), while the individuals in the control group had an average age of 38 years (SD 13.3). The comparison of age between groups showed no statistically significant difference (p=0.180).

Inclusion criteria for the study were:

Control group: healthy subjects who were accompanying CATEM patients or were employees of the ISC-

MSP. They had no history of lower limb (LL) surgery, no alcohol consumption, no vertigo, no sensory abnormalities in the LL, no major visual deficit and no use of hypnotic drugs, and they were all literate individuals.

– **Study group:** individuals with confirmed MS diagnosis according to the McDonald criteria, without any restriction on disease development, but with a grading between zero and 6.5 in the Expanded Disability Status Scale (EDSS). They had no history of LL surgery, no alcohol consumption, no vertigo and no visual impairment that prevented them from viewing the field test. They were not undergoing any relapse, they were not using any psychotropic drugs and they were all literate individuals.

The SSST, a test that quantitatively assesses LL function measured over time, was conducted between July and September 2008, always in the morning, in a closed environment, without influence from noise or visual stimulus. Non-slip paint was used to mark out the field. The research subjects were instructed to wear comfortable shoes.

The test field measured one by five meters (width and length, respectively). In the middle of the width line (0.50 m), a circle of 20 cm diameter was placed in the proximal and distal areas of the line. On each lateral line, two circles were placed at distances of one and three meters on the left. Another two circles were placed at two and four meters. Five wooden spheres with a diameter of 8 cm and a height of 4 cm, weighing 125 grams each were placed in the center of each circle, except in one of them, which was considered to be the starting point (Figure).

The subjects were instructed to walk on alternating sides and to push the sphere outside of the circle, alternating between the medial and lateral sides of the feet. The subjects were brought to the test as quickly as possible, without running, and did the route four times. For the first and third movements, moving the ball was done with the right lower limb (RLL), while for the second and fourth movements, it was done with the left lower limb (LLL). The test was timed and the result was the average of the four pathways. Before the test itself, the examiner showed the subjects how to perform the test, both verbal-

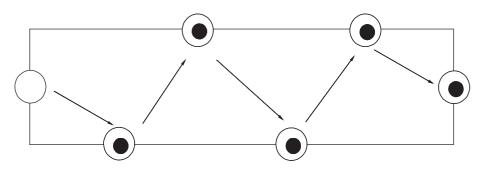


Figure. Field of Six Step Spot Test.

Table 1. Summary measurements of average final (test and retest) results in the groups, considering gender.

	Group	Gender	N	Average	SD	Median
Test	Study	F M	59 16	15.4 13.0	10.9 4.7	11.3
	Control	F	16 55	7.2	1.3	12.6 6.9
		Μ	36	6.5	0.8	6.4
Retest	Study	F M	59 16	14.5 12.6	9.5 4.2	11.1 12.7
	Control	F M	28 19	7.3 6.7	1.0 0.7	7.2 6.8

F: female; M: male; N: number of subjects; SD: standard deviation.

Table 2. Pearson's correlation coefficients for time for the test and retest, and the delta for the two groups.

Variance 1	Variance 2	Study	Control
SSST test	Age EDSS	0.27 0.62	0.43
SSST retest	Age EDSS	0.24 0.64	0.38
Delta	Age EDSS	0.24 0.27	0.18

SSST: Six Spot Step Test; EDSS: Expanded Disability Status Scale.

ly and in practice. The subjects practiced the SSST three times for proper assimilation and learning, and only after five minutes of rest did the subject really take the test. Two hours after performing it, a retest was carried out.

SSST qualitative values were compared using analysis of variance (ANOVA), Mann-Whitney test or Kruskal-Wallis test. The comparison of SSST between the test and the retest was conducted using the Wilcoxon test.

To assess the reliability of the instrument, the intraclass correlation coefficient (ICC) was calculated, with assessment of confidence intervals.

The significance level was considered to be 5% (0.05). The SPSS statistical package was used for evaluations (Statistical Package for Social Sciences), version 13.0.

RESULTS

Comparison between groups

There were no differences between the two groups regarding race, education, marital status and dominance (p values equal to 0.220, 0.172, 0.921 and 0.171 respectively). When the groups were compared regarding gender, it was observed that there was a significant difference between them (p=0.012). In order to verify whether this difference could affect the reliability of the test, analysis of variance was used to compare the two groups in relation to gender, and for test and retest.

When gender was compared within each group, the final average trend was similar for men and women.

Comparison of average times in the test, controlled for gender and group, indicated that the difference between the times was due only to the group (p<0.001), as confirmed by the analysis of variance. Gender did not interfere in test or retest time (Table 1).

Factors that may influence the SSST

To study the factors that could influence the time taken in the SSST, the Pearson correlation coefficient was calculated between the times for the test and retest and the delta of age and EDSS, in order to verify whether there was a difference in time for gender and education, as well as for EDSS. The study group was thus divided into two groups, according to the EDSS score. The first group of individuals had EDSS varying from zero to 3 (48 subjects -48%) and the second group had EDSS varying from 3.5 to 6.5 (27 subjects -36%).

Table 2 shows the Pearson's correlation coefficients for times and the delta for age, which were relatively low values. Since the times for the test and retest seemed to increase, the EDSS seemed to show the same tendency (0.619 and 0.635, respectively).

Table 3 shows a summary of the final average measurements (test and retest), according to the group and schooling. The analysis of variance showed a significant difference (p<0.001) between the groups (study and control). This difference was not observed when the schooling variable was assessed in the test and retest, respectively (p=0.271, p=0.437).

Reliability test

To evaluate the reliability of the SSST, the test and retest were applied in both groups. The results are presented in Tables 4 and 5.

Table 4 presents the summary of time measurements for SSST performance in the test and in the retest, according to the group. Although this table might suggest that the average and median times were very similar in each group, the values are significantly different in the

Table 3. Summary measurements of average final (test and retest) results in the groups, considering schooling.

	Group	Schooling	N	Average	SD	Median
Test	Study	Basic	6	11.4	2.3	11.1
		High School	35	13.5	6.5	11.0
		Superior	34	17.0	13.0	11.5
	Control	Basic	16	7.7	1.7	7.1
		High School	35	7.0	1.0	6.8
		Superior	40	6.6	0.8	6.3
Retest	Study	Basic	6	11.0	2.3	10.7
		High School	35	12.9	5.5	11.1
		Superior	34	15.9	11.3	12.2
	Control	Basic	11	7.1	1.3	6.9
		High School	15	7.3	0.9	7.4
		Superior	21	6.9	0.7	6.8

N: number of subjects; SD: standard deviation.

Table 4. Summary of SSST values according to EDSS categories for test and retest.

				Study	Study		
Variance			N	Average	SD	Р	
EDSS	0.0 to 3.0	Test Retest	48 48	10.5 10.2	2.5 2.4	<0.001*	
	3.5 to 6.5	Test Retest	27 27	22.8 21.1	13.0 11.0	0.019*	

^{*}Wilcoxon; SSST: Six Spot Step Test; EDSS: Expanded Disability Status Scale; N: number of subjects; SD: standard deviation.

Table 5. Summary measurements of time (in seconds) in SSST test and retest, for the study and control groups.

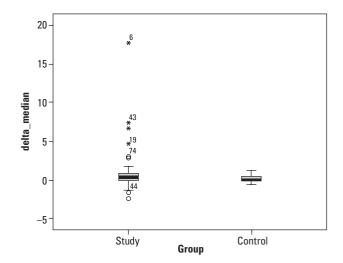
and retest, for the study and control groups.						
	Group					
Summary	Stu	ıdy	Con	Control		
measurements	Test	retest	Test	retest		
N	75	75	91	47		
Average	14.91	14.12	7.22	7.04		
Median	11.37	11.20	7.17	6.87		
SD	9.94	8.64	1.00	0.94		
Least	5.89	6.20	5.59	4.90		
Maximum	63.59	64.10	10.18	9.70		
p*	< 0.001		0.021			

^{*}Wilcoxon's test; N: number of subjects; SD: standard deviation.

comparison between test and re-test (p<0.001 for the study group and p=0.021 for the control group).

According to the Pearson correlation between the times of testing and retesting, considering the lowest or the highest degree of EDSS, the value for EDSS 0 to 3 was 0.979, while for EDSS 3.5 to 6.5, the value was 0.960.

In order to understand the change of behavior for the test and retest, the difference between the two times was calculated and called "delta". The closer to zero the value of delta was, the closer were the times between test and re-



Graph. Delta of the study and control groups presented in Boxplot.

test. Considering the mean and median values for the control group, it seems that a lower delta is obtained in comparison with the other group. This result was independent of the larger variability of the study group, as shown in the Graph. No statistically significant difference was observed when both groups were compared (p=0.063).

The intraclass correlation coefficient (ICC) was calculated in order to verify the reliability of the instrument.

These values could vary from zero to one, and the closer to one they were, the greater the reliability of the instrument was.

The value of 0.983 was obtained for the ICC of the MS patients group, with a 95% confidence interval (values ranging from 0.972 to 0.989). The ICC value for the control group was 0.945 (95% CI range=0.902 to 0.970), thus demonstrating a high degree of reliability in both groups.

The Pearson correlation coefficient between the times for the test and retest was calculated, and the value of 0.975 was obtained for the study group; while the value 0.898 was obtained for the control group. This result further strengthened the ICC results.

DISCUSSION

The relevance of MS studies focusing on gait lies in the characteristics of this progressive disease that affects young adults. After an evolution period of 10 years, 50% of these patients may be unfit to perform their professional and personal activities, thus resulting in individual and family suffering, associated with a great economic burden on society. These patients may be lost to the workforce, at the same time that they generate a high cost with their prolonged course of treatment⁶⁻⁹.

With better-defined diagnostic criteria due to imaging advances, and with better understanding of the disease, there have been increasing numbers of diagnoses. Thus, a greater need for disability evaluations among MS patients has been created^{7,10-12}.

Disability evaluation among MS patients has given rise to the challenge of constructing instruments that are able to translate this clinical parameter of MS¹³.

Although the EDSS is the most widely used scale for MS evaluation¹⁵, and is considered to be the gold standard, its application is limited only to doctors, since there is an evaluation of several functional systems. The authors agree with Nieuwenhuis et al., 2006⁵ who said that this is not a linear scale, and its psychometric properties are not reliable, since it is not entirely objective and incorporates some subjective variables. Parameters such as functional bladder and bowel scores (self-reported by the patient) and the ability to walk (100 to 500 meters) may be unreliable since the interview is carried out in an office¹⁴.

Although the American Committee considers the Multiple Sclerosis Functional Composite Measure (MSFC) to be a standard test because it evaluates the motor performance of the upper limbs (UL) and LL, along with cognition ^{15,16}, there are points that are not fully covered by this form of assessment, when gait is considered. A more complex test such as SSST may increase the sensitivity of this evaluation, but before it is fully used in Brazil, it needs to be validated.

The ability to walk is commonly used as a measure-

ment of the impact of disease among patients with MS^{17,18}. On average, two thirds of these patients retain the ability to walk for up to 20 years after the diagnoses. Due to other associated disabilities such as weakness, spasticity, lack of coordination, fatigue and vestibular disorders, MS patients may walk more slowly, for shorter distances. When they reach the disease phase that requires double support, there may also be a decreased range of motion in the ankle, knee and hip^{1,19,20}. Tests that can identify the influence of these factors could help detect changes in disease evolution, thereby allowing earlier interventions.

An ideal tool for use in clinical practice must be sensitive enough to assess what it proposes to evaluate. It also needs to be reproducible and easy to apply, understand and interpret, with low cost, taking into consideration the particular features of the population under evaluation. The authors believe that SSST has all of these characteristics for evaluating MS patients.

The time taken to walk is a common clinical measurement of mobility and function. Among patients with neurological disorders, the slower they move, the higher are the sensory-motor and cognitive dysfunctions^{3,20,21}.

The SSST was the subject of this study because it is extremely important to have assessment tools that can be sensitive to the particular features of the disease being studied. This is exactly the case with this test, which takes into account the various changes that may be present in MS patients. The application of test and retest in order to verify its reliability showed that the time between them was different. However, despite being mathematically significant, both for average and median values, this difference did not affect the clinical characteristics, since it was of the order of hundredths of a second (Table 5). Gait-related spatial-temporal parameters in MS are highly variable between and among the days. However, there are no reports on the variability of gait parameters, which may explain this difference between test and retest².

In general, the patients with MS had significantly different gait parameters, compared with healthy subjects. MS patients walked more slowly, as was confirmed by the SSST, showing that the study group had a median time of 11.37 seconds, while the control group had 7.17 seconds^{1,19,20}.

In the original study using SSST, which was carried out in Denmark, the study group showed lower median values. Patients did the test in 10.7 seconds, while controls did it in 6.4 seconds, meaning that the population in Denmark performed the test a bit faster than the Brazilian subjects did. However, they can be considered to be similar. This difference may be due to the colder climate of Denmark, since heat decreases the nerve conduction, thereby slowing the movements.

In the study by Morris et al.¹, gait and fatigue were compared among MS patients in the morning and in the

afternoon. These authors found that there was little variability in the gait pattern, despite the significant increment in fatigue. Despite this result, the present study was carried out always in the morning, in order to avoid any influence from another variable.

In analyzing our sample in order to verify the psychometric properties of the test, it was found that there was little difference between the test and the retest. In analyzing this difference by calculating a delta value, it was observed that in the control group, this difference was closer to zero, meaning that the SSST time did not vary much. Although the control group showed similar test and retest results, the study group showed a higher difference, due to greater variability relating to different levels of disabilities.

Temporal-spatial variability often occurs among individuals with higher degree of disability measured by EDSS¹⁴. Individuals with EDSS scores of up to three may walk without support or auxiliary devices for distances of more than 500 meters. This variability increases with the severity of MS, in parallel with an increase in the EDSS values. Individuals with EDSS over six need to use a cane, crutches or Zimmer frame to walk 100 meters^{5,18,20}.

The median values of both groups were very close to zero, thus indicating that the SSST is reproducible for the two groups. Therefore, in comparing the delta between the groups, there was no statistical difference. This was reinforced and assured by the ICC (0.983). This result indicates a high degree of reliability, with similar values to those obtained by Nieuwenhuis et al. in 2006⁵, when they developed the SSST (0.95).

Amato and Ponziane¹³ reported that individuals with EDSS between 4 and 7.5 presented greater involvement of LL, relating to the maximum distance they could walk. In our study, it was found that individuals with EDSS between 3.5 and 6.5 needed longer times to perform the SSST due to these disabilities, thus indicating that the greater the disability was, the greater the time taken for the test.

In the sample assessed in this study, age did not have any strong correlation with time taken to do the SSST, although Nieuwenhuis et al.⁵ observed that the group of patients between 20-39 years of age was the fastest, scoring 9.1 seconds. In their study, the group aged between 40-80 years performed the test in 13 seconds. These values are closer to the average age of our sample and, as reported in the literature, the higher the age is, the slower the gait is. The age-related speed reduction may be due to prolonged duration of double support use. The findings of gait time reduction among individuals of different ages in the control group indicated an impairment of gait stability in these patients.

Regarding the variable of gender, men performed the test more quickly than women only in the control group.

The same result was reported by Tilbery et al. ¹⁶ when examining 91 normal subjects for the Brazilian standardization of the Timed 25 Foot Walk (T25FW) for the MSFC. That study investigated the relationship between the level of education and time taken to complete the T25FW, and found that the higher the schooling was, the lower the time taken for the test was. This was also the case in the present study regarding the control group, since it showed that individuals with higher schooling did the SSST in a shorter time (p=0.017). However, this did not occur in the study group, since their disability (measured by the EDSS) was the factor that influenced test time.

The SSST is a reproducible, reliable and valid test for MS patients in Brazil. The authors agree with Nieuwenhuis et al.⁵ that this test is a good alternative for gait assessment during rehabilitation. It is also worth remembering that this test is easy to apply, favored by easy understanding, representation and low cost. This test can be applied by doctors and physiotherapists, thereby enabling its widespread use among MS patients in Brazil, for observing the course of disease and managing pharmacological or rehabilitation interventions.

The results from the present study have led the authors to propose a short study to be carried out soon, to modify the MSFC, since the LL motor item of the latter does not take into account the sensory-motor characteristics of the disease and the use of auxiliary resources. Therefore, the authors believe that this modified test may be a replacement based on the SSST, perhaps allowing a more complete assessment of broader functional capacity. This new, modified instrument could increase the knowledge of MS in this field.

According to the statistical results from the present study, the SSST is a valid, reliable and reproducible tool for Brazilian MS patients.

REFERENCES

- Morris ME, Cantwell C, Vowels L, Dodd K. Changes I gait and fatigue from morning to afternoon in people with multiple sclerosis. J Neurol Neurosurg Psychiatry 2002;72:361-365.
- Albrecht H, Wötzel C, Erasmus LP, Kleinpeter M, König N, Pöllmann W. Dayto-day variability of maximum walking distance in MS patients can mislead to relevant changes in the Expanded Disability Status Scale (EDSS): average walking speed is a more constant parameter. Mult Scler 2001;7:105-109.
- Savci S, Inal-Ince D, Arikan H, et al. Six-minute walk distance as measure of functional exercise capacity in multiple sclerosis. Disabil and Rehabil 2005;27: 1365-1371.
- Goulart F, Pereira LX. Uso de escalas para avaliação na doença de Parkinson em fisioterapia. Fisioter Pesqu 2005;11:49-56.
- Nieuwnhuis MM, Tongeren H Van, Sorensen PS, Ravnborg M. The Six Spot Test: a new measurement for walking ability in multple sclerosis. Mult Scler 2006:12:495-500.
- Moreira MA, Felipe E, Mendes MF, Tilbery CP. Estudo descritivo de suas formas em 302 casos. Arq Neuropsiquiatr 2000;58:460-466.
- Lana-Peixoto MA, Callegaro D, Moreira MA, et al. Consenso expandido do BCTRIMS para o tratamento da esclerose múltipla: diretrizes baseadas em evidências e recomendações. Arq Neuropsquiatr 2002;60:881-886.
- 8. Bennett R, Papshev D. Comparing the cost-effectiveness of the disease-mod-

- ifying drugs utilized in the management of relapsing forms of multiple sclerosis: a model evaluating the clinical and economic impact of current treatment options [abstract]. JMCP 2007;13:170.
- Callegaro D, Goldbaum M, Moraes L, et al. The prevalence of multiple sclerosis in the city of São Paulo, Brazil, 1997. Acta Neurol Scand 2001;104:208-213.
- 10. Luetic G. MS in Latin America. Int Mult Scler J 2008;15:6-11.
- 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;50:121-127.
- Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 Revisions to the "McDonald Criteria". Ann Neurol 2005;58:840-846.
- 13. Amato MP, Ponziani G. Quantification of impairment in MS: discussion of the scales in use. Mult Scler 1999;5:216-219.
- Fauchard-Renard C, Renard J-F, Miret N, Hannequin D, Mihout B, Weber J. Mesure de l'efficacité thérapeutique au cours d'une poussée de sclérose en plaques par l'analyse de la marche: comparasion avec le score EDSS. Rev Neurol 2001;157:649-654.
- 15. Hoogervorst ELJ, Zwemmer JNP, Jelles B, Polman CH, Uitdehaag BMJ. Multi-

- ple Sclerosis Impact Scale (MSIS-29): relation to estabilished measures of impairment and disability. Mult Scler 2004;10:569-574.
- Tilbery CP, Mendes MF, Thomaz RB, et al. Padronização da Multiple Sclerosis Functional Composite Measure (MSFC) na população brasileira. Arq Neuropsiquiatr 2005:63:127-132.
- Martin CL, Phillips BA, Kilpatrick TJ, et al. Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. Mult Scler 2006;12: 620-628.
- Nilsagard Y, Lundholm C, Gunnarsson L-G, Denison E. Clinical relevance using timed walk tests and "timed up and go" testing in persons with multiple sclerosis. Physiother Res Int 2007;12:105-114.
- Sheffler LR, Hennessey MT, Knutson JS, Naples GG, Chae J. Functional effect of an ankle foot orthesis on gait in multiple sclerosis: a pilot study. Am J Phys Med Rehabil 2008;87:26-32.
- 20. Crenshaw SJ, Royer TD, Richards JG, Hudson DJ. Gait variability in people with multiple sclerosis. Mult Scler 2006;12:613-619.
- 21. Cattaneo D, Jonsdottir J, Zocchi, Regola A. clinical of balance exercises on people with multiple sclerosis: a piloty study. Clin Rehabil 2007;21:771-781.