SPINE Volume 35, Number 23, pp E1298–E1305
©2010, Lippincott Williams & Wilkins

"The Fly": A New Clinical Assessment and Treatment Method for Deficits of Movement Control in the Cervical Spine

Reliability and Validity

Eythor Kristjansson, PT, PhD,* and Gudny Lilja Oddsdottir, PT, MSc†

Study Design. Test-retest and case-control study designed to detect accuracy of cervical spine movements by comparing 3 incrementally difficult movement patterns. An asymptomatic group, a nontrauma neck pain group, and a group with whiplash-associated disorders, Grade II, were tested (n = 18 in each group).

Objective. To determine the test-retest reliability and the discriminative validity of the new Fly method.

Summary of Background Data. A lack of reliable and valid measures for grading the deficits of movement control in the cervical spine makes it impossible to prescribe treatment appropriate to each patient's respective impairment level.

Methods. Head tracking of a moving fly which appeared on a computer screen. Easy, medium, and difficult patterns, each of which was repeated 3 times in random order, were tested. Amplitude accuracy (deviation of movements), directional accuracy (time on target, undershoots *vs.* overshoots) were compared across patterns and groups on 2 occasions, 1 week apart.

Results. The intraclass correlation coefficient$_{2,1}$ ranged from 0.53 to 0.82 for both variables, except for the subvariable "overshoots" (0.14–0.42). The limits of agreement (LOA) were progressively wider across patterns (easy–medium–difficult) and groups (asymptomatic–nontrauma–whiplash-associated disorder). Analysis of variance with repeated measures revealed significant differences between patterns within each group and between groups respectively for both outcome variables ($P < 0.001$).

Conclusion. The Fly method provides reliable and valid measures for movement control of the cervical spine. Higher means and wider LOA across patterns and subject groups are reasoned to be inherent in the new Fly method and the subject groups tested. The wide LOA in the symptomatic groups supports the development of a normative database. The new Fly method can be used both as an assessment and a treatment method and ensures gradual progression in the treatment for deficits of movement control in patients with neck pain.

Key words: movement control, cervical spine, proprioception, assessment, test-retest reliability, validity. **Spine 2010;35:E1298–E1305**

The clinical implications of the statement, "you can only control what you sense,"[1] have not been researched in detail in patients with neck pain. This is surprising, as it has been known for a long time that the cervical spine is a very delicate sensory organ,[2–6] due to the abundance of mechanoreceptors in the musculoskeletal tissues.[7,8] The dense network of mechanoreceptors in this region is essential for precise neuromuscular control of mobile cervical spine segments.[9–11]

"The Fly" is a recently developed clinical test that measures control of cervical spine movements. The patient is equipped with tracking sensors on the head and asked to follow a cursor (a fly) on the computer screen as accurately as possible.[12] Therefore, the Fly test addresses an important proprioceptive function, which is the regulation of movements, *i.e.*, detection and correction of errors through feedback and reflex mechanisms, while performing active movements.[13,14] This test demonstrated impaired movement control of the cervical spine in a group of patients with whiplash-associated disorders (WAD) of Grades I to II, according to the Québec Task Force classification scheme,[15] when compared with a healthy control group.[12]

The main drawback of the prior Fly test was that no attempt was made to generate incrementally difficult sets of movement patterns to precisely grade the impairment level of each individual patient. The Fly method was developed further in this study to address this shortcoming by creating 3 incrementally difficult classes of movement patterns: easy, medium, and difficult. A new software program was written for that purpose, generating an almost infinite number of movement patterns in each of the 3 classes according to specific criteria. The criteria are based on predefined parameters which form the boundaries for each of the 3 movement classes (patent pending). The movement classes in the new Fly method were created by different relationships between acceler-

From *FORMI, Oslo University Hospital, Ullevaal, Oslo, Norway; and †University of Iceland, Reykjavik, Iceland.
Acknowledgment date: November 3, 2009. First revision date: March 28, 2010. Second revision date: April 14, 2010. Acceptance date: April 26, 2010.
The manuscript submitted does not contain information about medical device(s)/drug(s).
No funds were received in support of this work. One or more of the author(s) has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript: *e.g.*, honoraria, gifts, consultancies, royalties, stocks, stock options, decision making position.
This study was conducted at LSH University Hospital, Reykjavík, Iceland.
A start-up innovation company, NeckCare AS, was recently established by E.K. to make the new methods available for clinicians over the Internet. NeckCare's home address is at Oslo Innovation Center, Gaustadallèen, 21, N-0349 Oslo, Norway.
Financial assistance for this research was received from the Icelandic Research Fund for Graduate Students.
Address correspondence and reprint requests to Eythor Kristjansson, PT, PhD, FORMI, Oslo University Hospital, Ullevaal, Building 37b, NO-0407 Oslo, Norway; E-mail: eythork@tele2.no

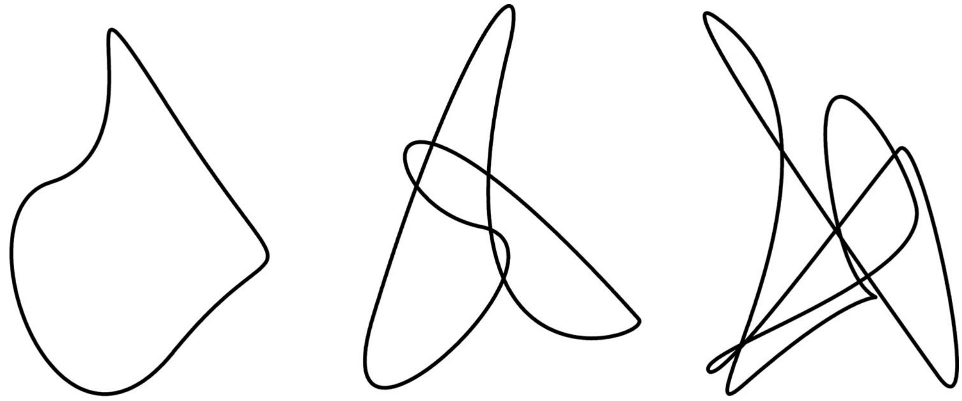

Figure 1. From left to right: Easy, medium, and difficult movement patterns traced by the Fly method, which the participants were required to follow by moving their heads. The duration of the patterns was 25, 40, and 50 seconds, respectively.

ation and deceleration *versus* number and acuity of the angles, as well as shape and length of the trajectories (Figure 1). The novel motion tracking method was also developed for treatment purposes, making it possible to start training at a suitable level of difficulty for a given subject based on the results of the new Fly test.

The aims were to ascertain the test-retest reliability and the discriminative validity of the 3 incrementally difficult movement classes generated by the new Fly method, comparing control, nontrauma neck pain, and whiplash subjects, Grade II. A test-retest and a case-control study were conducted. Three preselected patterns in each class were used to assess the patient's performance in each of the 3 movement classes (Figure 1). Two outcome measures were used, comprising the assessment part of the new Fly method. Comparison of the WAD group with the nontrauma group was performed to ascertain whether patients with WAD had more deficits of movement control in the cervical spine.

Materials and Methods

Participants

Table 1 shows the demographics of the 54 individuals who participated, equally divided into asymptomatic, nontraumatic neck pain, and WAD (Grade II) groups. Samples of convenience were used in all groups. To be included, the asymptomatic group was to have no history of musculoskeletal pain or injury to the neck. Furthermore, the subjects in the WAD group were required to have a history of symptoms after 1 motor vehicle collision with no prior symptoms in the head or neck. Both symptomatic groups had to have had neck and headache symptoms for more than 6 months and less than 2 years and score more than 3 on the visual analogue scale (VAS) during the last week before testing. Individuals were excluded from the WAD group if their symptoms corresponded to Grades III or IV, as classified by the Québec Task Force on WAD.[15] Systematic diseases or psychological disorders of any kind excluded participants from all groups. All participants completed questionnaires recording demographic data and general health. The symptomatic groups also completed the Neck Disability Index (NDI),[16] the Tampa Scale of Kinesiophobia,[17] and graded their minimum and maximum pain on a VAS scale. Ethical clearance for the study was obtained from the National Bioethics Committee.

Table 1. Comparison of Demographics and Results of the Questionnaires

	Asymptomatic (n = 18)	Nontrauma (n = 18)	Whiplash (n = 18)
Age (mean yr [SD])	32.2 (10.9)	38.0 (8.3)	35.5 (11.9)
Gender (male/female)	10/8	7/11	2/16
NDI (SD)*	—	22.3 (7.9)	35.8 (16.4)
Tampa (SD)	—	30.2 (6.2)	35.1 (8.5)
VAS maximum	—	6.7 (2.6)	8.0 (1.4)
VAS minimum*	—	1.9 (1.4)	3.2 (1.5)

*Significant differences by independent *t* test.
SD indicates standard deviation; NDI, neck disability index; Tampa, tampa scale for kinesiophobia; VAS, visual analogue scale.

Measurements

A 3-space Fastrak system was used in the study. The method is briefly summarized in this study; for details, the reader is referred to Kristjansson *et al*[12]; Figure 2 shows the experimental set-up.

Two cursors, 1 blue and 1 black, were seen on the computer screen. The blue cursor (derived from the Fastrak system) indicated movements of the head. The black cursor (derived from the Fly software program) traced the new generated movement patterns. Only the cursors were visible on the computer screen, not their trajectories, which made prediction of movements difficult. A new software program was used to format and process the data for analysis. The horizontal and vertical differences between the positions of the 2 sensors were calculated and used to determine the position of a cursor on a computer screen.

The subjects were asked to use the blue cursor to follow, as accurately as possible, the black cursor of the Fly. Three different movement patterns—easy, medium, and difficult—representing each movement class, were selected from the software program and used for the test procedure. The outcome measures were amplitude accuracy (AA) and directional accuracy (DA). AA was recorded by continuously calculating the absolute distance (radius) between the 2 cursors during the test sequence. For DA, an invisible free zone was created surrounding and moving with the target (the fly). During a trial, the percentage of the total time spent within, ahead of, and behind the free zone represented time on target, overshoots, and undershoots, respectively. The shape and size of the free zone was determined by plotting the coordinates (x, y) of 10 healthy individuals, and this determined a zone representing the differ-

Figure 2. Experimental set-up. The placement of the 2 sensors (arrows) was on the forehead and the back of the head, respectively. The participants were seated in front of a computer and were asked to find their natural head posture. The distance from the participants' earlobe to the computer monitor was 100 cm.

ence between the actual points and the predefined points. By using vectors ($A_t(P_t - Q_t)$), a scatter plot was obtained which was distributed around the center of the coordinate system. The circle around the center was divided into 24 equally large sectors. Then for each of the 24 sectors, quartiles were found and standard deviations (SDs) of the distance from the center were measured. The size of the free zone was set at 2 SDs from the center.

Procedure

The participants were measured during the day on 2 occasions 1 week apart. A research assistant provided the participants with written and verbal information about the test procedure. Demographic data were recorded and the symptomatic groups completed the questionnaires. The examiner, who was blinded to the group allocation, explained the intention and nature of the task required of the participants. To familiarize them with the task, all participants executed one movement pattern twice. This pattern was not used for measurements. The participants were then required to repeat each of the 3 movement patterns (Figure 1) 3 times, with a 10-second interval between each pattern. The test was performed in random order across patterns and trials. The participants had no knowledge about the different difficulty grades of the patterns.

Data Analysis

Analysis of variance (ANOVA) with repeated measures was used for comparison between the 2 independent variables, patterns, and groups. The model used to describe the data includes the main effects for the pattern, within-week trial, and group. Week effects, as well as the subject, were the random effects. Also included are the 2-way group-pattern, trial-group, and group-week interactions. The absolute error in millimeters ± SD was used to indicate AA. DA or time on target, overshoots *versus* undershoots, were each indicated as the percentage of the total time used to perform the trial. The mean of 3 trials for each movement pattern was calculated for each subject for both dependent variables. The statistical methods used for assessing test-retest reliability were the following: the intraclass correlation coefficients (ICCs) using an absolute agreement definition, the 2-way random effects model, single measures ICC and 95% confidence interval (CI) ($ICC_{2,1}$),[18] the mean difference between test and retest together with the 95% CI, and the method of assessing agreement between 2 measurements of the same subject. This included a scatter plot of the differences between Test 1 and 2 against their mean with 95% limits of agreement (LOA).[19]

Two-way ANOVA with the questionnaires as covariates and Pearson correlation were used to ascertain the association between the test results *versus* the results of the questionnaires and VAS.

Number, subjects, means, and SDs were used for description of data. Analyses were performed with the procedures implemented by the R statistical software (R version 2.6.2, copyright 2008, The R Foundation for Statistical Computing). The significant level was set at $P < 0.05$ for all tests.

Results

Table 1 shows the demographics and group comparison of the questionnaires and VAS.

Amplitude Accuracy

The mean difference and 95% CI revealed no statistically significant difference between test sessions 1 and 2 in any of the movement patterns, except for the easy pattern in the asymptomatic group (Table 2). Figures 3A to C show the differences in the easy pattern between test sessions 1 and 2, plotted against their mean for each subject in the control, nontrauma, and WAD groups, respectively.

A repeated-measures ANOVA revealed significant differences between patterns within each group and between groups, respectively ($df = 2$; $F = 341.0$; $P < 0.001$) (Figure 4). There was a significant group-pattern interaction, that is, the patterns affect the groups differently ($df = 4$; $F = 8.10$; $P < 0.001$). The test-retest calculations were marginally significant ($df = 1$; $F = 3.19$; $P = 0.074$). However, looking at individual paired t test for each group-pattern combination, a significant difference was only revealed in the easy pattern for the asymptomatic group ($t = 2.09$; $df = 53$; $P = 0.041$).

Directional Accuracy

Table 3 depicts test-retest reliability for test sessions 1 and 2 for each group and each subvariable, time on target, undershoots, and overshoots, respectively.

Time on Target. Repeated-measures ANOVA revealed significant differences between patterns within each group and between groups, respectively ($df = 2$; $F = 498.84$; $P < 0.001$). There was a significant group-pattern interaction, that is, the patterns affect the groups differently ($df = 4$; $F = 5.24$; $P < 0.001$) and no significant difference was revealed between test sessions 1 and 2 ($df = 1$; $F = 0.02$; $P = 0.874$).

Table 2. Test-Retest Reliability for Amplitude Accuracy (mm) in Each of the 3 Movement Patterns and Each of the 3 Subject Groups (n = 18 in Each Group)

Movement Pattern	Test 1 Mean (SD)	Test 2 Mean (SD)	Mean Difference, (95% CI)	$ICC_{2,1}$ (95% CI)	95% LOA
Asymptomatic					
Easy	1.78 (0.33)	1.70 (0.35)	0.09 (0.01 to 0.17)	0.59 (0.38–0.74)	−0.42 to 0.60
Medium	2.17 (0.44)	2.08 (0.47)	0.09 (−0.02 to 0.21)	0.53 (0.32–0.70)	−0.58 to 0.78
Difficult	2.64 (0.52)	2.54 (0.59)	0.10 (−0.03 to 0.25)	0.56 (0.34–0.72)	−0.66 to 0.88
Nontrauma					
Easy	2.06 (0.52)	2.01 (0.53)	0.05 (−0.05 to 0.16)	0.71 (0.55–0.82)	−0.54 to 0.65
Medium	2.70 (0.88)	2.67 (1.01)	0.02 (−0.17 to 0.23)	0.67 (0.49–0.79)	−1.29 to 1.34
Difficult	3.42 (1.30)	3.26 (1.26)	0.17 (−0.09 to 0.43)	0.72 (0.56–0.82)	−1.02 to 1.36
Whiplash					
Easy	2.52 (0.78)	2.66 (1.05)	−0.15 (−0.35 to 0.05)	0.68 (0.51–0.80)	−1.47 to 1.68
Medium	3.45 (1.45)	3.36 (1.31)	0.08 (−0.22 to 0.39)	0.67 (0.49–0.79)	−1.46 to 1.63
Difficult	4.09 (1.51)	3.95 (1.52)	0.14 (−0.21 to 0.48)	0.65 (0.47–0.78)	−1.93 to 2.21
All participants					
Easy	2.12 (0.64)	2.12 (0.81)	0.01 (−0.08 to 0.07)	0.75 (0.67–0.81)	−0.90 to 0.89
Medium	2.77 (1.13)	2.70 (1.11)	0.07 (−0.06 to 0.19)	0.73 (0.65–0.79)	−1.15 to 1.28
Difficult	3.38 (1.31)	3.25 (1.31)	0.14 (−0.01 to 0.28)	0.73 (0.65–0.79)	−0.70 to 0.97
Across patterns	2.76 (1.18)	2.69 (1.19)	0.07 (−0.01 to 0.14)	0.78 (0.74–0.81)	−0.89 to 1.03

Mean (SD) for test session 1 and 2; mean difference, 95% CI (test 2 − test 1); Intraclass correlation coefficient ($ICC_{2,1}$) (95% CI); and 95% LOA.
SD indicates standard deviation; CI, confidence interval; ICC, intraclass correlation coefficients; LOA, limits of agreement.

Undershoots. Repeated-measures ANOVA revealed significant differences between patterns within groups ($df = 2$; $F = 551.23$; $P < 0.001$) and a significant group-pattern interaction ($df = 4$; $F = 4.33$; $P = 0.001$). Significant differences were found between test occasions 1 and 2 ($df = 1$; $F = 9.23$; $P = 0.002$). However, looking at individual paired *t* test for each group-pattern combination, a significant difference was only revealed in the difficult pattern in the asymptomatic group ($t = -2.36$; $df = 53$; $P = 0.022$).

Overshoots. Repeated-measures ANOVA revealed significant differences between patterns within groups ($df = 2$; $F = 86.47$; $P < 0.001$) but there was not a significant group-pattern interaction ($df = 4$; $F = 1.11$; $P = 0.348$). Significant differences were revealed between test occasions 1 and 2 ($df = 1$; $F = 29.60$; $P < 0.001$). The individual paired *t* test for each group-pattern combination showed significant differences in all patterns in the asymptomatic group, in the difficult pattern in the nontrauma group, and in the medium pattern in the WAD group between test sessions.

Relationship Between the Test Results and the Questionnaires

No significant differences were found. For example, a 2-way ANOVA including NDI as a covariate when modeling AA in the patient groups, taking into account pattern type differences, revealed no significance ($df = 1$; $F = 0.75$; $P = 0.390$). Figure 5 shows a scatter plot comparing the AA results *versus* the NDI scores.

Discussion

The results of this study indicate that the new Fly method is both reliable and valid for measuring movement control of the cervical spine. Several reliability statistics were used to provide information regarding whether some outcome measures were more reliable than others, and whether the new method was reliable for different groups of subjects as well as for individual subjects. The $ICC_{2,1}$ results (Tables 2 and 3) indicate that the outcome measures were sufficiently reliable to observe change in each of the 3 subject groups, except for "overshoots" in the asymptomatic and the nontrauma groups. In AA, the ICC values ranged from 0.53 to 0.72 (95% CI, 0.32%–0.82%). In DA, the ICC values ranged from 0.14 to 0.82 (95% CI, −0.03% to 0.89%). The lower range (0.14–0.42) comes from the overshoots measure. Larger variations between subjects (as indicated by bigger SDs of the means) result in higher ICC values in the 2 patient groups, compared with the more homogeneous data for the healthy control group (Table 2).[20] The LOA for AA reflect this, as progressively wider limits were observed across the easy, medium, and difficult patterns and the control, nontrauma, and WAD groups, respectively (Table 2 and Figures 3A–C). Contrary to the ICCs, the LOA indicates the results in the unit of measurement. LOA can therefore be used as a "reference range," because when the difference between 2 measurements, as a result of, for example, treatment intervention, is outside LOA, it represents a real change.[21] It implies that relatively larger differences in an individual performance would be required in the symptomatic groups to confidently state that a real change had taken place.

No systematic changes between the test-retest trials were noted for the variables in the AA measures, as zero was included in the 95% CI, except for the easy pattern in the asymptomatic group, which possibly indicates a learning effect. For DA, the "time on target" variable revealed no systematic changes between the test-retest trial. The same applied for the "undershoots," except in the difficult pattern for the asymptomatic group. All subjects spent less time "ahead of target," and the greatest fluctuations were observed in the "overshoots" variable, revealing systematic changes between test-retest in all

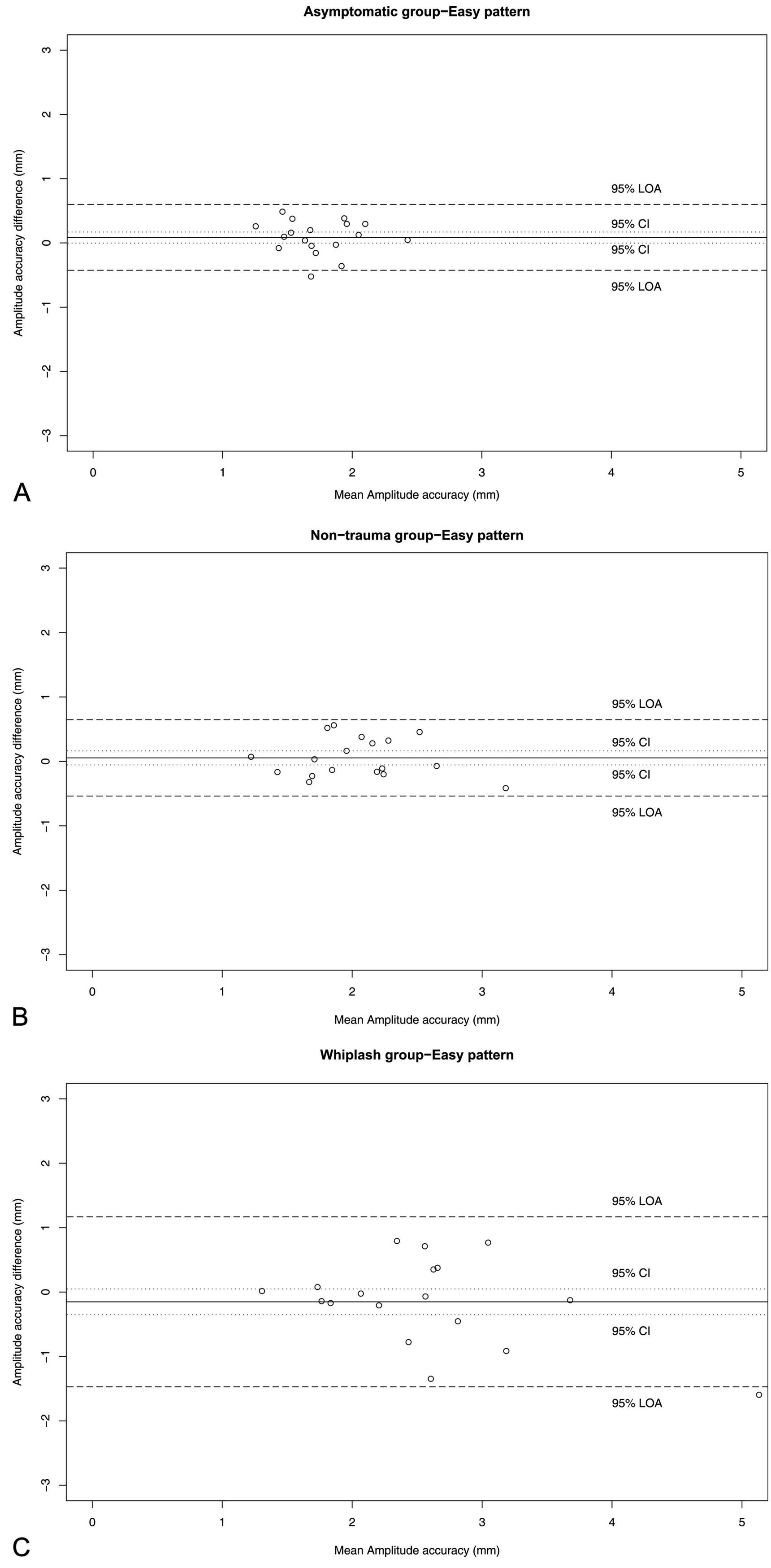

Figure 3. **A,** Bland and Altman graph with LOA. The differences between test sessions 1 and 2 (Test 2 − Test 1) in the easy pattern, plotted against their mean for each subject in the AA (mm) in 18 asymptomatic subjects, together with the 95% CI and the 95% LOA. **B,** Bland and Altman graph with LOA. The differences between test sessions 1 and 2 (Test 2 − Test 1) in the easy pattern, plotted against their mean for each subject in the AA (mm) in 18 nontrauma subjects, together with the 95% CI and the 95% LOA. **C,** Bland and Altman graph with LOA. The differences between test sessions 1 and 2 (Test 2 − Test 1) in the easy pattern, plotted against their mean for each subject in the AA (mm) in 18 WAD subjects, together with the 95% CI and the 95% LOA.

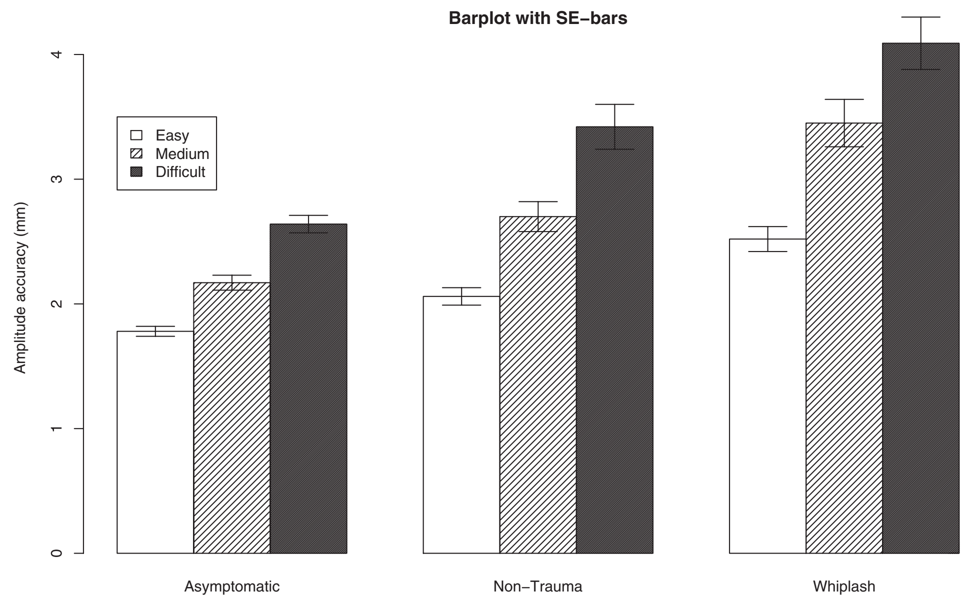

Figure 4. Within and between group comparison for AA. The results revealed significant differences between all 3 movement patterns within each group and between each of the 3 groups, respectively.

patterns for the asymptomatic group, in the difficult and medium patterns for the nontrauma group and the WAD group, respectively (Table 3).

According to Rankin and Stokes,[20] at least 50 subjects are needed in reliability studies; otherwise, the 95% LOA would be too wide. Therefore, one reason for the wider LOA in the symptomatic groups in this study may be the small sample size. However, the wider LOA in the WAD group compared with the nontrauma group cannot be explained by small sample sizes alone. Wider LOA in the WAD group may parallel the clinical observation that symptoms in patients with WAD tend to change more from one time or day to another, when compared to patients with nontraumatic neck pain. To be 95% confident of detecting a real change in the easy pattern, the wider LOA in the WAD group suggests that AA, on average, would need to decrease (improve) by about 1.47 mm, and increase (deteriorate) by about 1.68 mm. Individual improvement in patients with WAD would therefore need to exceed the scores of the asymptomatic group. A patient with WAD and an AA of 2.52 mm would have to show improvement of $2.52 - 1.47 = 1.05$, to reveal a real change, which is a better performance than the mean outcome measure for the asymptomatic group (1.78 mm) (Table 2). This example shows that the clinometric properties of the measurements, described by the reliability measures, may differ from what clinicians judge as minimally and clinically important changes. Instead of using LOA as a "reference range," we suggest that a reference database of normal values for each age group and gender would be more useful when assessing improvements in the symptomatic groups. Development of a normative database for the new Fly method is already in progress.

The 3 incrementally difficult movement patterns generated by the new Fly method were significantly different

Table 3. Test-Retest Reliability for Directional Accuracy (%) in Each of the 3 Movement Patterns and Each of the 3 Subject Groups (n = 18 in Each Group)

	Test Session 1, Mean (SD)			Test Session 2, Mean (SD)			$ICC_{2,1}$ (95% CI)		
Test Pattern	Time on Target	Overshoots	Undershoots	Time on Target	Overshoots	Undershoots	Time on Target	Overshoots	Undershoots
Control									
Easy	63.3 (12.9)	13.3* (7.4)	23.4 (11.2)	64.9 (13.5)	10.9 (5.9)	24.2 (13.6)	0.62 (0.43 to 0.76)	0.21 (−0.03 to 0.45)	0.56 (0.35 to 0.72)
Medium	54.8 (13.5)	7.9* (5.3)	37.3 (12.6)	54.4 (13.2)	5.7 (4.0)	39.9 (13.4)	0.58 (0.37 to 0.74)	0.32 (0.07 to 0.54)	0.57 (0.36 to 0.73)
Difficult	40.1 (9.3)	12.1* (7.1)	47.8* (11.3)	40.7 (9.5)	8.1 (5.3)	51.2 (10.7)	0.56 (0.35 to 0.72)	0.38 (0.09 to 0.60)	0.50 (0.27 to 0.68)
Nontrauma									
Easy	54.7 (19.1)	14.7 (9.2)	30.6 (19.8)	53.9 (17.9)	14.5 (9.5)	31.5 (18.8)	0.78 (0.65 to 0.86)	0.42 (0.17 to 0.62)	0.69 (0.53 to 0.81)
Medium	44.0 (16.1)	8.1 (5.4)	47.9 (16.0)	41.7 (16.4)	7.8 (4.8)	50.5 (16.8)	0.64 (0.45 to 0.77)	0.14 (−0.14 to 0.39)	0.58 (0.38 to 0.73)
Difficult	30.6 (12.9)	11.9* (6.8)	57.4 (14.4)	32.7 (14.2)	9.5 (6.2)	57.8 (16.1)	0.76 (0.61 to 0.85)	0.34 (0.09 to 0.55)	0.68 (0.51 to 0.80)
WAD									
Easy	42.7 (20.5)	14.9 (11.5)	42.4 (21.8)	42.1 (20.3)	12.6 (9.7)	45.3 (23.8)	0.72 (0.56 to 0.82)	0.61 (0.41 to 0.75)	0.82 (0.71 to 0.89)
Medium	34.1 (16.5)	9.8* (5.6)	56.1 (17.1)	33.3 (16.9)	7.6 (5.2)	59.1 (18.1)	0.81 (0.70 to 0.89)	0.57 (0.33 to 0.74)	0.79 (0.65 to 0.87)
Difficult	24.4 (12.9)	11.3 (6.5)	64.3 (13.7)	25.7 (16.0)	10.2 (7.5)	64.1 (18.9)	0.74 (0.59 to 0.84)	0.62 (0.42 to 0.76)	0.70 (0.53 to 0.81)

Mean (SD) for test session 1 and 2, and intraclass correlation coefficient ($ICC_{2,1}$) (95% CI).
*Significant differences by paired t test between test session 1 and test session 2.
SD indicates standard deviation; CI, confidence interval; WAD, whiplash-associated disorders; ICC, intraclass correlation coefficients.

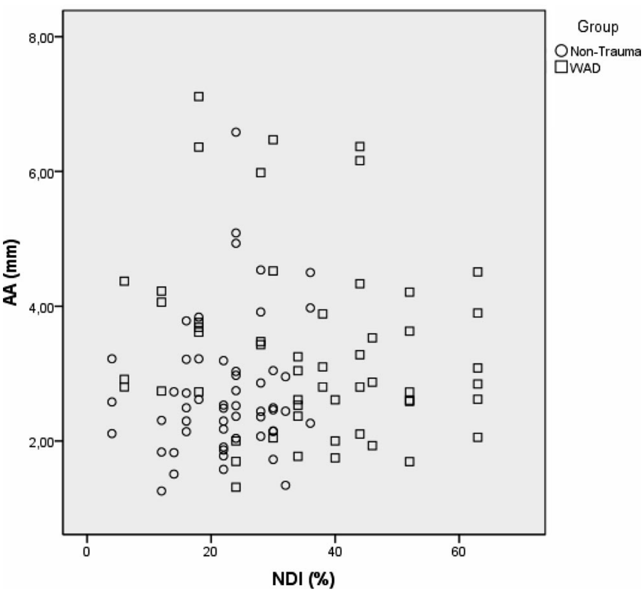

Figure 5. A scatter plot of the AA test results and the NDI scores for the WAD group and the nontrauma neck pain group. No linear relation was revealed, ($r = 0.057$).

within and between each of the 3 subjects groups (Figure 4). The results of this study fulfilled the expected clinical requirements of the new Fly method that the group with whiplash would perform significantly worse than the nontrauma neck pain group. The method can be used both as an assessment and a treatment method.

The treatment can consequently be started at each patient's respective impairment level, enhancing gradual progression in the treatment for deficits of movement control in the cervical spine. In motor skill learning it is important that the level of task difficulty is adjustable to the individual patient's skill level, to ensure that the exercise is neither too difficult nor too trivial to perform. This aspect has been highlighted in the challenge point framework theory.[22] The optimal challenge point represents the degree of task difficulty needed for an individual of a specific skill level to optimize learning. By adjusting the task difficulty to the change in ability, the optimal challenge point is maintained.[22] This can be accomplished by moving on to the next difficulty level in the new Fly method (Supplemental Digital Content 1, Video, online only, available at: http://links.lww.com/BRS/A476, which demonstrates both the assessment and the treatment part of the Fly method in a clinical setting).

We hypothesize that the differences between the symptomatic groups may reflect sudden *versus* gradual onset. Higher scores in the WAD group may be due to direct damage to the mechanoreceptors and their axons, because they have lower tensile strength than the surrounding collagen fibers.[23] Therefore, this may lead to effects similar to those of "deafferentation," which causes diminished accuracy and adaptability when complex tasks (reflected by the incremental movement patterns in this study) are tested.[24] Muscle fatigue may be the cardinal sign in the nontrauma group[25] and may be caused by the Cinderella syndrome, which has been well described elsewhere.[26] Arthrogenic pain inhibition may be more prominent in the WAD group.[27,28] Thus, higher test scores and variability for the WAD group compared with the nontrauma group in the new Fly test (Tables 2 and 3) and the NDI (Table 1) may be inherent in the new test patterns and the onset or etiology of the symptoms. Moreover, no relationship was revealed between the test results, on the one hand, and pain, disability or fear of movement, on the other, which strengthens the test results obtained in this study.

■ Key Points

- The new Fly method provides reliable and valid measures for movement control of the cervical spine and can be used both as an assessment and treatment method, ensuring gradual progression in the treatment for deficits of movement control in patients with neck pain.
- Three incrementally difficult movement patterns (easy, medium, and difficult), each of which was repeated 3 times in random order, were tested on 2 occasions 1 week apart. An asymptomatic group, a nontrauma group, and a group with WAD, Grade II, participated.
- The method revealed significant differences between patterns within each group and between groups respectively for the 2 outcome variables, AA and DA.
- Higher means and wider LOA across patterns and subject groups are reasoned to be inherent in the new Fly method and the subject groups tested. The wide LOA supports the development of a normative database for reference when assessing improvement in deficits of movement control in the cervical spine in patients with neck pain.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.spinejournal.com).

References

1. McCloskey D, Prochazka A. The role of sensory information in the guidance of voluntary movement. *Somatosens Mot Res* 1994;11:69–76.
2. Magnus R, DeKleijn A. Die Abhängigkeit des Tonus der Extremitätenmuskeln von der Kopfstellung. *Pflugers Arch Gesamte Physiol Menschen Tiere* 1912;145:455–584.
3. Magnus R. Some results of studies in the physiology of posture. Cameron prize lectures. *Lancet* 1926;211:531–6.
4. McCough G, Derring I, Ling T. Location of receptors for tonic neck reflexes. *J Neurophysiol* 1951;14:191–5.
5. Ryan GM, Cope S. Cervical vertigo. *Lancet* 1955;31:1355–8.
6. Lindsey K, Roberts T, Rosenberg J. Asymmetric tonic labyrinth reflexes and their interaction with neck reflexes in the decerebrate cat. *J Physiol* 1976;261:583–601.
7. Kulkarni V, Chandy MJ, Babu KS. Quantitative study of muscle spindles in suboccipital muscles of human foetuses. *Neurol India* 2001;49:355–9.
8. Liu J, Thornell L, Pedrosa-Domellof F. Muscle spindles in the deep muscles

of the human neck: a morphological and immunocytochemical study. *J Histochem Cytochem* 2003;51:175–86.
9. Abrahams VC. The physiology of neck muscles; their role in head movement and maintenance of posture. *Can J Physiol Pharmacol* 1977;55:332–8.
10. Richmond FJ, Bakker DA. Anatomical organization and sensory receptor content of soft tissues surrounding upper cervical vertebrae in the cat. *J Neurophysiol* 1982;48:49–61.
11. Dutia MB. The muscles and joints of the neck: their specialisation and role in head movement. *Prog Neurobiol* 1991;37:165–78.
12. Kristjansson E, Hardardottir L, Asmundardottir M, et al. A new clinical test for cervicocephalic kinesthetic sensibility: "the fly." *Arch Phys Med Rehabil* 2004;85:490–5.
13. Gandevia S, Burke D. Does the nervous system depend on kinesthetic information to control natural limb movements? *Behav Brain Res* 1992; 15:614–32.
14. Prochazka A. Proprioceptive feedback and movement regulation. In: Rowell L, Shepherd JT, eds. *Handbook of Physiology, Sec. 12: Exercise: Regulation and Integration of Multiple Systems*. New York, NY: Oxford University Press; 1996:89–127.
15. Spitzer W, Skovron M, Salmi L, et al. Scientific monograph of the Quebec task force on whiplash-associated disorders: redefining "whiplash" and its management. *Spine* 1995;20:1–73.
16. Vernon H, Mior S. The neck disability index: a study of reliability and validity. *J Manipulative Physiol Ther* 1991;14:409–15.
17. Kori S, Miller R, Todd D. Kinesiophobia: a new view of chronic pain behavior. *Pain Manag* 1990:35–43.
18. Shrout P, Fleiss JL. Intraclass correlations. Uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
19. Bland J, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135–60.
20. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analysis. *Clin Rehabil* 1998;12:187–99.
21. Lexell J, Downham D. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil* 2005;84:719–23.
22. Guadagnoli M, Lee T. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *J Mot Behav* 2004; 36:212–24.
23. McLain RF. Mechanoreceptor endings in human cervical facet joints. *Spine* 1994;19:495–501.
24. Gandevia SC. Kinesthesia: roles for afferent signals and motor commands. In: Rowell L, Shepherd JT, eds. *Handbook of Physiology, Sec. 12: Exercise: Regulation and Integration of Multiple Systems*. New York, NY: Oxford University Press; 1996:128–72.
25. Johansson H, Windhorst U, Djupsjöbacka M, et al. *Chronic Work-Related Myalgia: Neuromuscular Mechanisms Behind Work-Related Chronic Muscle Pain Syndromes*. Umeå, Sweden: Gävle University Press; 2003.
26. Hägg GM. Static work loads and occupational myalgia—a new explanation model. In: Anderson PA, Hobart DJ, Danoff JV, eds. *Electromyographical Kinesiology*. Amsterdam, The Netherlands: Elsevier Science Publishers; 1991:141–4.
27. Thunberg J, Hellström F, Sjölander P, et al. Influences on the fusimotor-muscle spindle system from chemosensitive nerve endings in cervical facet joints in the cat: possible implications for whiplash induced disorders. *Pain* 2001;91:15–22.
28. Windhorst U, Kokkoroyiannis T. Dynamic behavior of alpha-motoneurons subjected to recurrent inhibition and reflex feedback via muscle spindles. *Neuroscience* 1992;47:897–907.