Quantifying motor control of the lumbar spine by a dynamic sitting device; a pilot reliability study

M.P.E. Gijssel
Student number: 3336352
Final Thesis: 14/08/2011
School of Clinical Health Sciences, Physiotherapy science, Utrecht University, Utrecht, The Netherlands

First supervisor: W. Dankaerts, PhD, MT, PT
Trainee post: Department of Rehabilitation Sciences, University of Leuven, Leuven, Belgium
Second supervisor: T. Takken, PhD

“ONDERGETEKENDE

M.P.E. Gijssel

---

bevestigt hierbij dat de onderhavige verhandeling mag worden geraadpleegd en vrij mag worden gefotokopieerd. Bij het citeren moeten steeds de titel en de auteur van de verhandeling worden vermeld.”

Samenvatting

Introductie: Vrijwel iedereen krijgt tijdens zijn leven een periode van rugpijn te verduren. Door aanwezigheid van pijn kunnen adaptieve, beschermende bewegingen en beperkingen in de motorische controle veroorzaakt worden. Aanhoudende foutieve adaptatie van de motorische controle kan leiden tot een staat van chronische pijn. Een belangrijke oorzaak van chronische aspecifieke lage rugklachten is lumbale segmentale instabiliteit. Bij patiënten met lumbale segmentale instabiliteit wordt een onvermogen om de lumbale wervelkolom te verplaatsen waargenomen. Een nieuw ontwikkelde dynamische stoel meet hoe accuraat een patiënt, gebruikmakend van de lumbaal stabiliserende musculatuur, bepaalde parcours aflegt. Hierdoor wordt inzicht verkregen in de mogelijkheid de lage rug musculatuur te coördineren. Deze studie is opgezet om de congruentie en betrouwbaarheid te bepalen van de scores gegeven door dit meetinstrument.

Methode: Vier therapeuten onderzochten zes patiënten (gemiddelde leeftijd: 43.8 SD ±7.7 jaar) met rugklachten, op twee meetdagen met drie weken tussentijd. Zittend op een dynamische stoel diende de patiënt een op beeldscherm getoond parcours af te leggen. De bewegingen in de onderrug werden gemeten door een computer, de nauwkeurigheid waarmee de bewegingen werden gemaakt werd gescoord. De interbeoordelaarsbetrouwbaarheid werd berekend met behulp van intraclass correlatie (ICC), test hertest data werden geanalyseerd met een gepaarde T-toets voor vier verschillende parkoersen.

Resultaten: Aantal volbrachte cycli toonde een goede tot uitstekende betrouwbaarheid (ICC: 0.87-0.90). Verschillen in doorgebrachte tijd tussen tegenovergestelde zijden van het parkoers gaf een goede betrouwbaarheidsscore (ICC: 0.72-0.97). Na drie weken werd een significante ($p<0.05$) vooruitgang in het aantal afgelegde cycli gevonden.

Conclusie: Het is legitiem scores van verschillende beoordelaars te vergelijken. Dit meetinstrument lijkt nuttig om objectief de mogelijkheid om de lage rug te verplaatsen te kwantificeren. Toekomstig onderzoek zal zich moeten richten op grotere populaties van patiënten met lage rugklachten om de kennis van dit instrument als diagnostisch middel te versterken.

Kernwoorden: motorische controle lumbale wervelkolom meetinstrument

Abstract

Introduction: Nearly everyone is affected by low back pain at some point in time. The presence of pain can cause adaptive or protective altered movement and motor control impairments. Ongoing mal-adaptive motor control behaviour can lead to a chronic state of pain. A significant cause of non specific chronic low back pain is lumbar segmental instability. Patients with lumbar segmental instability demonstrate an inability to reposition the lumbar spine accurately. A newly introduced dynamic sitting device measures how accurately a subject completes different tracks by using lumbar stabilizing muscles, and thereby provides insight into the ability to coordinate lower back muscles. The present study was set up to estimate the reliability and agreement parameters over the test scores given by the TBT.

Methods: Four therapists tested six patients (mean age: 43,8 SD ±7,7 years) with low back pain, on two days with at three-week intervals. Whilst sitting on a dynamic sitting device the patients had to cover a presented track. The movements in the lower back were measured by a computer, the precision with which the movements were made was scored. The inter-rater reliability was calculated using intraclass correlation (ICC), test retest data were analyzed using a paired t-test for four different tracks.

Results: The number of completed cycles showed a good to excellent reliability (ICC: 0.87-0.90). Differences in time spent between opposite sides of the track gave good reliability scores (ICC: 0.72-0.97). Over three weeks a significant ($p<0.05$) improvement was found in the number of cycles covered.

Conclusion: It is legitimate to compare scores measured by different raters. This instrument seems useful to objectively quantify the ability to reposition the lumbar spine. Future research needs to focus on a larger sample size of patients with LBP to enforce knowledge about this instrument as a diagnostic device.

Keywords: motor control lumbar spine measurement

Introduction

Low back pain (LBP) is the most prevalent of musculoskeletal conditions; nearly everyone is affected by it at some point in time.(1) No pathoanatomic/radiologic abnormality is found in 85% of patients with LBP, which entitles them as having 'nonspecific' chronic low back pain (NS-CLBP).(1,2) The presence of pain can cause adaptive or protective altered motor behaviour or movement, and motor control impairments.(3-7) Ongoing mal-adaptive motor control behaviour provides a basis for ongoing peripherally driven nociceptor sensitisation, leading to a chronic state of pain.(8) A number of studies show that mal-adaptive movement and motor control impairments appear to result in ongoing abnormal tissue loading and mechanically provoked pain in subjects with NS-CLBP.(9-13) A significant cause of NS-CLBP is lumbar segmental instability,(14) defined as laxity around the neutral position of a spinal segment.(15) The ligamentous, musculotendenous and neural system interact to provide the necessary stability to the spine.(16) When low neuromuscular control occurs this leads to lumbar segmental instability.(15) Increased body sway and differences in sitting posture are observed in patients with LBP due to altered control of the stabilizing muscles.(12,17-19) Subsequently, the loss of proprioceptive control and consequently the neuromuscular imbalance has a direct relationship with LBP populations.(18,20,21) Proprioception is crucial to stimulate specific neural recruitment patterns of the lumbar stabilizing musculature to meet task demands.(22) Two muscle systems were proposed to act in the maintenance of spinal stability; these are *local* stabilisers (mostly; M. Multifidus, M. Transversus Abdominus and M. Obliquus Internus) and *global* stabilisers (mostly; M. Erector Spinae and M. Rectus Abdominis).(23,24)

Differences in proprioception do exist between individuals with and free from LBP.(20) Due to this, patients with lumbar segmental instability demonstrate an inability to reposition the lumbar spine accurately.(18) Different clinical tests are used to examine muscular coordination of the lumbar spine.(25-28) A newly introduced dynamic sitting device, called the Trunk Balance Trainer (TBT), measures how accurately a subject completes different tracks by using lumbar stabilizing muscles, thereby providing insight into the ability to coordinate lower back muscles.(29) The use of objective parameters in daily practice is encouraged. So far no objective measurement method, to evaluate the ability to reposition the lumbar spine, in patients with NS-CLBP is known. The present study was set up to estimate the reliability and agreement parameters over the test scores given by the TBT. In consideration of a nationwide study, involving twenty private practices in the Netherlands, the primary objective is to determine the inter-rater reliability.

Methods

Subjects

Thirteen potential candidates living within a 30 km radius from Roosendaal, The Netherlands were derived from a national database for TBT users. Four responders agreed to participate, for this reason the research population consisted of two physiotherapists and two orthopaedic manual therapists (raters). Their mean age was 34,5 years (standard deviation [SD]±12,7), median years experience working as a physical therapist 10,6 (SD±10) and their mean months experience with the TBT was 20 months (SD±9,5). All four raters had undergone a two day (16h) training course to become certified in the use of the TBT.

Six patients, including two males and four females, consulted the research setting, a private health clinic in Roosendaal, due to NS-CLBP. Their mean age was 43,8 years (SD±7,7), mean height 170 cm (SD±9), and mean weight 73 kg (SD±11). Their NS-CLBP had a history of a minimum of three months, located in the lower lumbar spine without radiation. Subjects were excluded if they had radiating pain, previous back surgery or specific diagnoses for LBP disorders, consequently radiographic imaging had to be available before inclusion of the study.(12) All data was collected during usual care, so no extra effort was needed from the patients and according to the Dutch law, no approval from a medical ethics committee was needed. All participants provided written informed consent prior to participation and were free to withdraw from this study at any time.

Instrument

Measurements will be conducted by means of the VelDon® Flexchair/TBT (VelDon®, Gildeweg 18, 2632 BA Nootdorp, The Netherlands). The TBT is a dynamic sitting device with a double cant mechanism, mostly used for training lower back muscles or instructing sitting posture. However, since movements can be recorded it is feasible to use the TBT as a testing device as well. A three dimensional accelerometer (sensor), which is placed under the seat, sends 13 signals per second to a receiver connected to a computer. On a screen a dot is displayed in a selected track, which represents the movement of the TBT and thereby of the lumbar spine. Flamaing found that there is a strong correlation between the registration of the TBT performing a dynamic sitting task in the sagittal plane and the actual low back alignment.(30) Each track is divided into octants to measure time spent in these areas. Measurements take 60 seconds and can be performed during training. Multiple outcomes are given by the TBT: firstly the number of cycles covered within a 60 second period, secondly the percentage of time spent in each octant shown, thirdly the percentage at which the subject locates the dot; this can be in the green 'good' area, the yellow 'acceptable' or the red 'wrong' area (Fig. 1). The second and third parameter are expressed in the percentage of time spent in the different areas or octants. VelDon® software is necessary to register and store data in a protected online database, this freeware is available on the website of the company www.veldon.nl, for this study version 23.0 was used.

Procedure

When using the TBT a strict protocol needs to be applied. To improve concordance two raters (JvB, EV) evaluated and pilot tested the protocol. Teaching the research protocol thoroughly to the raters diminished carryover effects. Furthermore, assessing trials followed each other with short intervals of one minute testing, one minute resting.(31) All data was collected within two days, at the same circumstances and at the same time of day. The first two tracks where measured by four raters, the third and fourth track by two raters. A three week test-retest interval was chosen on learning and memory effects between the two test days. The patients were examined by an independent orthopaedic manual therapist (EV) one week before the first measurement day. Patients registered their average experienced pain over the last three days ahead of the test days on an ungraduated 100 mm visual analogue scale (VAS).(32) Before each test the starting conditions were checked by the rater and the patient was informed about the test procedure while sitting on the TBT. After this, the sensitivity of the sensor speed was adjusted to the patient, so each track could be made comfortably, and mobility of the spine was not a limitation for covering the tracks. One extra minute of training was recorded and stored in the database in view of a learning effect, as recommended by different studies for muscle testing.(33-35) All measurements were gathered during a training program, which the patients received weekly as part of a cognitive-behavioural approach, according to the European guideline for NS-CLBP.(36) During the 60 seconds of testing per track per rater, the patient was not verbally supported. Since movements made on the TBT are low load contractions (< 20% maximum voluntary isometric contraction) fatigue was not expected to occur during measurement.(29,37) The raters were blinded to the subject scores until the completion of the study to decrease rater bias. The order in which the raters conducted the measurements was randomly allocated. The tests were performed in the same order for all patients.

Statistical analysis

To determine inter-rater reliability twelve measurements from six patients taken on two test days were used. The intraclass correlation coefficient (ICC) (absolute agreement) was used for the inter-rater reliability analysis, where acceptable values for interpretation where proposed: 0.0≥0.4: poor, 0.4≤0.75: fair to good, and 0.75≤1.00: good to excellent agreement with α=0.05.(38,39) To express the response stability both the standard error of measurement (SEM) and minimal detectable difference (MDD) where used. Since the standard deviation of the measurement errors reflects the reliability of the response, the SEM is calculated using the formula: SEM = $\sqrt{(\sigma^2_{pt} + \sigma^2_{residual})}$ where the σ^2_{pt} stands for the variance of the rater and the $\sigma^2_{residual}$ for the interaction between rater and patient.(40) Following this, the MDD was calculated according to the formula: MDD = SEM × $\sqrt{2}$ × 1.96. ANOVA was used to assess a difference in the number of years of experience using the TBT. To assess agreement between the first two raters Bland-Altman plots were used and the 95% limits of agreement (LoA) were calculated.

Difference in scores between the two measuring days where compared using a paired-samples T test for all three instrument parameters and the VAS score. The second parameter ‘difference in time spent per octant’ was split in two scores. The first score was the difference in time spent between the opposite octants demanded. The ideal score for this parameter would be 0 so both sides where level. The second was the deviation in time spent in the other six octants. Scores for contrary preference-side per patient where adjusted so the scores would not cancel each other out. Pearson’s coefficient of correlation for the strength of association between the mean number of cycles covered over all tracks and the change in VAS score was calculated, Portney and Watkins’ criteria for interpretation were used.(31) SPSS for Windows (v.16.0, Chicago: SPSS Inc) and MedCalc (v.11.4, Mariakerke: MedCalc Software bvba) were used for data management and analyses.

Results

All raters and patients completed the tests. The ICC for inter-rater reliability on track 1 to 4 was good to excellent for the number of cycles that were covered: ICCs between 0.87 and 0.90. The reliability for the difference in percentage of time spent in two (A and B) of the eight octants was good to excellent on track 1,3 and 4 (ICC: 0.86-0.97) and fair to good on track 2 (ICC: 0.72). Patients reported that they had more difficulty moving the dot correctly on that side of the track where they spent most time, both during the outward and return movement. The agreement on percentage of time spent in six of eight octants showed ICCs between 0.23 and 0.99 for track 1 and 3 respectively. A poor agreement was shown for the third parameter; the deviation between the different areas that the subject located the dot (table 1). ANOVA showed no significant difference between raters (p ranged from 0.14 to 0.89). In the Bland-Altman plots the scores around the zero point are spread evenly and randomly indicating an acceptable observed error for all parameters (figure 3 and 4). LoA varied from -4.15 to 2.98 and -1.48 to 1.98 for track 1 and 4 respectively. Smaller ranges were found on the second and third parameter (figure 4).

Despite the fact that the sensitivity of the sensor movement speed was adjusted, large differences between patients were seen in the number of cycles covered with a maximum range of twelve cycles. Furthermore, a significant ($p<0.05$) improvement in the number of completed cycles was found between the two measuring days (table 2). The mean rated VAS score on the first test day was 62 mm (SD±18) however, three weeks later it was 25 mm (SD±22). Pair wise comparison revealed that the mean VAS score was significantly decreased over three weeks ($p<0.002$). Due to the fact that genuine changes occurred in the measured variable, the test-retest reliability could not be calculated. Moreover, patient characteristics on the two measurement days altered and were incomparable for this reason. No analysis could be performed on a stabile subgroup.

The calculations for the correlation coefficient yield r=0.59. A moderate to good relationship is considered,(31) which suggests that there is an association between the number of cycles covered and the VAS score. However, no statistical significance (p<0.2) was identified.

Discussion

The primary objective was to determine the inter-rater reliability of the test scores given by the TBT. The present study showed that the parameters 'number of cycles covered' and 'difference in percentage time spent between the opposite octants' showed good to excellent inter-rater reliability. The scores for deviation from the A and B octant seem very unstable as they provide different outcomes without structural error or pattern. Evaluating these scores seems of no importance. A small error variance of the repeated measures ANOVA was shown which reflects no interaction between raters and subjects. However, due to a low variability among subjects' scores in the third parameter (88.2% of the cases located the dot 99-100% in the green area) the ICC shows a negative value and cannot be considered valid in this sample. Hence, the Bland-Altman plot and LoA reflect a low variability in figure 4c. In spite of the poor reliability found in this study on the deviation between different areas the subject locates the dot, patients do have more difficulty to reposition the lumbar spine as instructed when moving out of the green 'good' area. Concluding from the present study, the parameters 'number of cycles covered' and 'difference in percentage of time spent between the opposite octants' seem useful in evaluating the ability to reposition the lumbar spine in patients with NS-CLBP. The scores of the TBT might help to gain insight into the ability to reposition the lumbar spine and consequently express neuromuscular balance in an objective measure.(16,21,41,42)

Using data from two different measurement days is undesirable when calculating the inter-rater reliability. However, when calculating the ICCs independently for the two measurement days, commensurable scores were found. Due to a small sample size that was used, the present study should be considered as a pilot study. Moreover a wide variability in the patients' scores where observed which resulted in high SEM and MDD scores and wide LoA. Due to the fact that the Pearson's correlation was calculated over a small sample there was no statistical significance. An approximation by the method of Walter et. al.(43), recommended for reliability studies, showed that eleven raters were needed to satisfy the required sample size. Consequently, prudence is called for regarding data from the present study.

The parameter number of cycles covered consequently improved in all patients and seems thereby useful to evaluate their progression over time, furthermore this parameter is easy to understand for raters and patients and therefore relevant in practice. Flamaing et. al.(30) also advised using the number of cycles covered to register progression over time. Yet, even when a change beyond the MDD is found, values for important clinical improvement must be found to determine the benefit of the intervention.(44) An improvement in balance between the opposite octants over time was expected, however no significant (<0.05) improvement was found, except for track 3. Following these study findings three weeks of training does not decrease the difference in time spent between the opposite sides of the tracks.

The test-retest reliability could not be calculated, due to the fact that the intervals were chosen too widely, genuine changes in the measured variable were resulted. Future research should choose smaller intervals to gain insight into learning effects and possibly the appearance of fatigue.

Tidstrand and Horneij(28) stated that differences between the right and left side have an important role in diagnosing dysfunctional muscular coordination. Thereby, the parameter measuring difference in time spent between the opposite octants seems convenient for objective registration of a possible deficient to reposition the lumbar spine on one side. When low neuromuscular control leads to a state of pain, this pain might also influence the ability to reposition the lumbar spine.(18,20) For all the aspects mentioned above more research, on subgroups of patients with LBP in large numbers, is paramount to reveal differences between groups.(12,45) Through this the TBT might have surplus value in diagnostics regarding LBP in clinical practice.

Conclusion

The TBT has a good inter-rater reliability, due to the results that were obtained it is legitimate to compare scores measured in twenty clinical practices in a future national study. Due to genuine changes in the measured variables test-retest reliability could not be calculated. This instrument seems useful to objectively quantify the ability to reposition the lumbar spine. Values for important clinical improvement must be found to determine the benefit when using the TBT as a training intervention. Future research needs to focus on a larger sample size and more homogeneous subgroups of patients with LBP to enforce knowledge about the TBT as a diagnostic and training device.

Acknowledgments

The first author thanks prof. dr. Wim Dankaerts (first supervisor) and dr. Tim Takken (second supervisor) for their kind assistance throughout this study.

Tables:

Table 1 **Inter-rater reliability n=12**

	Track 1 r=4		Track 2 r=4		Track 3 r=2		Track 4 r=2	
	ICC (CI)	SEM MDD	ICC (CI)	SEM MDD	ICC (CI)	SEM MDD	ICC (CI)	SEM MDD
Number of cycles	0.87 (0.72-0.95)	3.12 8.65	0.90 (0.77-0.96)	1.10 3.06	0.89 (0.66-0.97)	2.15 5.95	0.90 (0.69-0.97)	0.87 2.40
Difference in %time spent A vs. B†	0.86 (0.70-0.95)	2.61 5.98	0.72 (0.48-0.89)	1.67 5.05	0.94 (0.81- 0.98)	1.95 5.32	0.97 (89-0.99)	1.22 4.59
Deviation in %time from A and B†	0.23	6.26 17.35	0.43	5.72 15.85	0.99 (0.96-1.00)	6.58 18.23	0.95 (0.84-0.99)	6.33 17.56
%time spent in green 'good' area†	-0.05	4.31 11.96	-0.05	1.26 3.48	0.37	4.29 11.9	0.10	2.98 8.25

n=number of patients, r=number of raters, ICC = intraclass correlation coefficient (for single measures), CI=95% confidence interval, SEM = standard error of measurement, MDD = minimal detectable difference, †SEM and MDD are expressed in percentage of time, A and B are two opposite octants per track

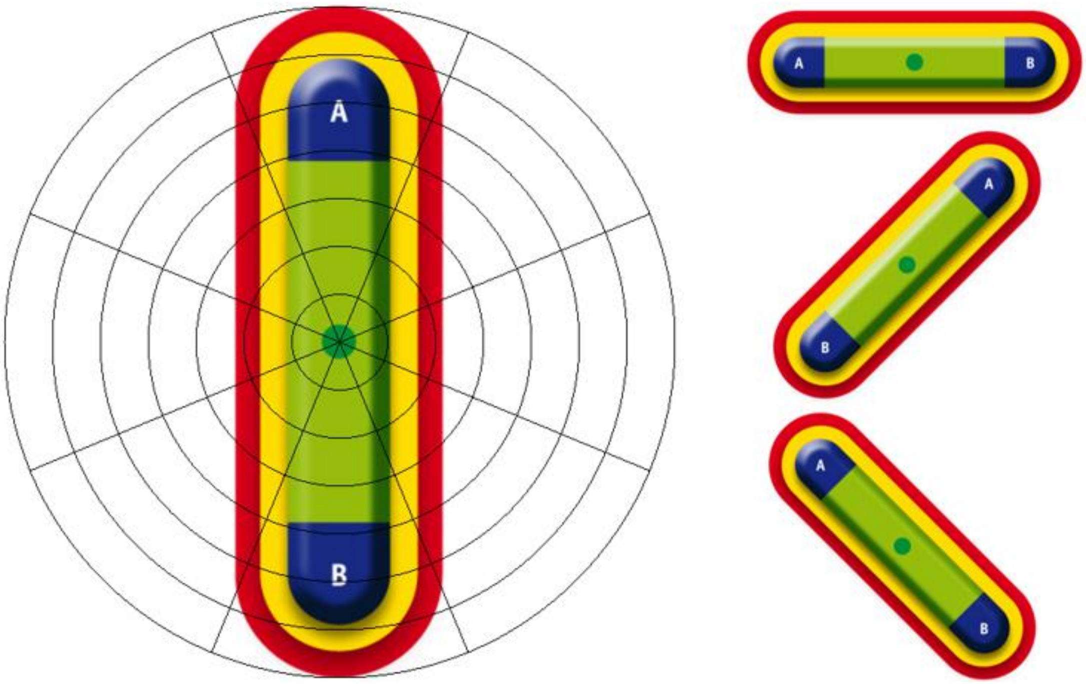
Table 2 **Test-Retest Data n=6**

	Mean Cycles ± SD [range]		Difference in % time spent A vs. B† ± SD		Deviation from A and B† ± SD		%time spent in green 'good' area† ± SD	
T	0	1	0	1	0	1	0	1
Track 1	6 ± 3.5 [10]	9 ± 4.2 [12] *	7.0 ± 6.8	8.5 ± 5.3	12,3 ± 3.5	11,5 ± 2.8	98,4 ± 1.8	99,6 ± 0.4
Track 2	6 ± 3.3 [8]	7 ± 2.2 [6]	4.9 ± 3.2	4.2 ± 1.5	10,3 ± 2.9	10,6 ± 2.7	99,8 ± 0.2	99,5 ± 0.5
Track 3	5 ± 0.7 [2]	6 ± 2.5 [7] *	9.1 ± 2.8	4.4 ± 3.8 *	14,4 ± 7.8	11,4 ± 5.8 *	98,5 ± 2.2	99,3 ± 1.2
Track 4	5 ± 1.8 [5]	6 ± 2.0 [5] *	6.4 ± 6.1	8.2 ± 7.6	12,1 ± 4.3	13,3 ± 6.4	99,3 ± 1.1	99,5 ± 1.0

T= test day, between test day 0 and 1 was a 3 week interval, SD = standard deviation, n=number of patients, * significantly improved $p < 0.05$
†expressed in percentage of time

Figures:

Figure 1

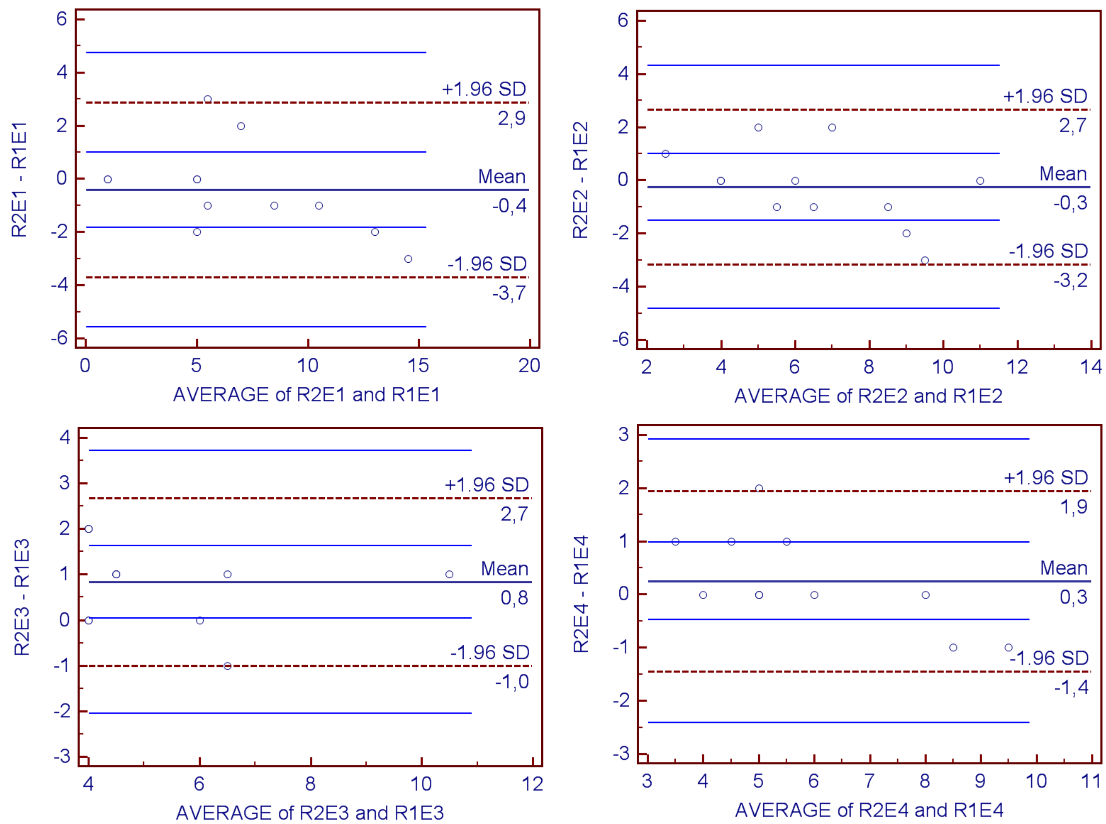


Track 1 divided into octants and tracks 2 to 4

Figure 2

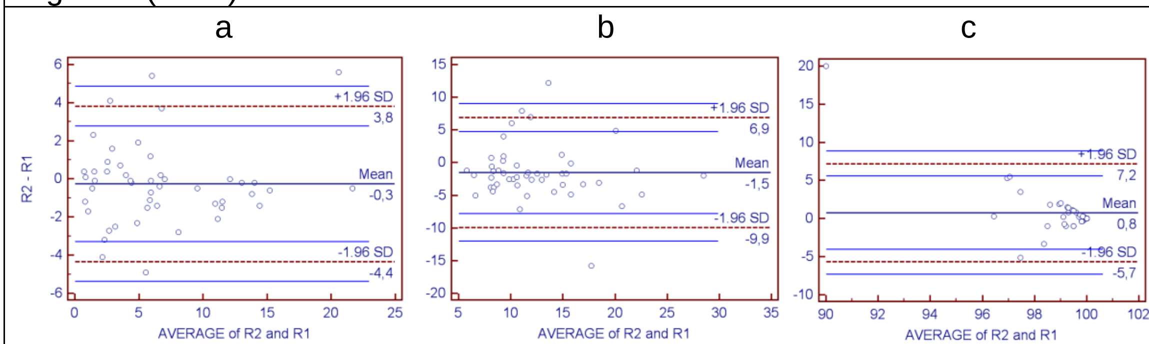
Live set-up in action

Figure 3 (n=6)



Bland-Altman plots for number of cycles covered 95 % limits of agreement: -4.15 to 2.98 (E1); -3.22 to 2.72 (E2); -1.04 to 2.71(E3); -1.48 to 1.98 (E4), R=rater, E=track

Figure 4 (n=48)



Bland-Altman plots for the percentage of time spent: a.between the opposite octants LoA (-3.78 to 3.26), b.in the other six octants LoA (-4.97 to 1.89), c.in the green area LoA (-3.15 to 4.67), R=rater

References:

(1) Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81(9):646-656.

(2) Dillingham T. Evaluation and management of low back pain: an overview. State of the Art Reviews 1995;9(3):559-574.

(3) Taimela S, Kankaanpaa M, Luoto S. The effect of lumbar fatigue on the ability to sense a change in lumbar position. A controlled study. Spine (Phila Pa 1976) 1999 Jul 1;24(13):1322-1327.

(4) Hall TM, Elvey RL. Nerve trunk pain: physical diagnosis and treatment. Man Ther 1999 May;4(2):63-73.

(5) Hodges PW, Moseley GL. Pain and motor control of the lumbopelvic region: effect and possible mechanisms. J Electromyogr Kinesiol 2003 Aug;13(4):361-370.

(6) van Dieen JH, Selen LP, Cholewicki J. Trunk muscle activation in low-back pain patients, an analysis of the literature. J Electromyogr Kinesiol 2003 Aug;13(4):333-351.

(7) Panjabi MM. Clinical spinal instability and low back pain. J Electromyogr Kinesiol 2003 Aug;13(4):371-379.

(8) O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. Man Ther 2005 Nov;10(4):242-255.

(9) O'Sullivan PB. Lumbar segmental 'instability': clinical presentation and specific stabilizing exercise management. Man Ther 2000 Feb;5(1):2-12.

(10) Hodges PW, Richardson CA. Relationship between limb movement speed and associated contraction of the trunk muscles. Ergonomics 1997 Nov;40(11):1220-1230.

(11) O'Sullivan PB, Mitchell T, Bulich P, Waller R, Holte J. The relationship beween posture and back muscle endurance in industrial workers with flexion-related low back pain. Man Ther 2006 Nov;11(4):264-271.

(12) Dankaerts W, O'Sullivan P, Burnett A, Straker L. Differences in sitting postures are associated with nonspecific chronic low back pain disorders when patients are subclassified. Spine (Phila Pa 1976) 2006 Mar 15;31(6):698-704.

(13) Shirazi-Adl A, Drouin G. Load-bearing role of facets in a lumbar segment under sagittal plane loadings. J Biomech 1987;20(6):601-613.

(14) Long DM, BenDebba M, Torgerson WS, Boyd RJ, Dawson EG, Hardy RW, et al. Persistent back pain and sciatica in the United States: patient characteristics. J Spinal Disord 1996 Feb;9(1):40-58.

(15) Panjabi MM. The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. J Spinal Disord 1992 Dec;5(4):390-6; discussion 397.

(16) Panjabi MM. The stabilizing system of the spine. Part I. Function, dysfunction, adaptation, and enhancement. J Spinal Disord 1992 Dec;5(4):383-9; discussion 397.

(17) Nies N, Sinnott PL. Variations in balance and body sway in middle-aged adults. Subjects with healthy backs compared with subjects with low-back dysfunction. Spine (Phila Pa 1976) 1991 Mar;16(3):325-330.

(18) O'Sullivan PB, Burnett A, Floyd AN, Gadsdon K, Logiudice J, Miller D, et al. Lumbar repositioning deficit in a specific low back pain population. Spine (Phila Pa 1976) 2003 May 15;28(10):1074-1079.

(19) Mak JN, Hu Y, Cheng AC, Kwok HY, Chen YH, Luk KD. Flexion-relaxation ratio in sitting: application in low back pain rehabilitation. Spine (Phila Pa 1976) 2010 Jul 15;35(16):1532-1538.

(20) Gill KP, Callaghan MJ. The measurement of lumbar proprioception in individuals with and without low back pain. Spine (Phila Pa 1976) 1998 Feb 1;23(3):371-377.

(21) Renkawitz T, Boluki D, Grifka J. The association of low back pain, neuromuscular imbalance, and trunk extension strength in athletes. Spine J 2006 Nov-Dec;6(6):673-683.

(22) Behm DG, Drinkwater EJ, Willardson JM, Cowley PM. The use of instability to train the core musculature. Appl Physiol Nutr Metab 2010 Feb;35(1):91-108.

(23) Bergmark A. Stability of the lumbar spine. A study in mechanical engineering. Acta Orthop Scand Suppl 1989;230:1-54.

(24) Panjabi MM. A hypothesis of chronic back pain: ligament subfailure injuries lead to muscle control dysfunction. Eur Spine J 2006 May;15(5):668-676.

(25) Dankaerts W, O'Sullivan PB, Straker LM, Burnett AF, Skouen JS. The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. Man Ther 2006 Feb;11(1):28-39.

(26) Fritz JM, Brennan GP, Clifford SN, Hunter SJ, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. Spine (Phila Pa 1976) 2006 Jan 1;31(1):77-82.

(27) Hicks GE, Fritz JM, Delitto A, Mishock J. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch Phys Med Rehabil 2003 Dec;84(12):1858-1864.

(28) Tidstrand J, Horneij E. Inter-rater reliability of three standardized functional tests in patients with low back pain. BMC Musculoskelet Disord 2009 Jun 2;10:58.

(29) Dunne J. To investigate low back posture and trunk muscle activation while exercising on an unstable chair (Flexchair) in a healthy population. in press.

(30) Flamaing PJ, Groenen E, Dankaerts W, Granitzer Ma, Meesen R, van Etten L. Assessment of lumbar spine posture during sitting on a dynamic sitting device (Flexchair®). (unpub) 2008.

(31) Portney LG, Watkins MP. Foundations of Clinical Research. Applications to Practice. 3rd ed. New Jersey, United States of America: Pearson Education, Inc; 2009.

(32) Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983 May;16(1):87-101.

(33) Giles B, Henke P, Edmonds J, McNeil D. Reproducibility of isokinetic muscle strength measurements in normal and arthritic individuals. Scand J Rehabil Med 1990;22(2):93-99.

(34) Graves JE, Pollock ML, Carpenter DM, Leggett SH, Jones A, MacMillan M, et al. Quantitative assessment of full range-of-motion isometric lumbar extension strength. Spine (Phila Pa 1976) 1990 Apr;15(4):289-294.

(35) Newton M, Waddell G. Trunk strength testing with iso-machines. Part 1: Review of a decade of scientific evidence. Spine (Phila Pa 1976) 1993 Jun 1;18(7):801-811.

(36) Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of

chronic nonspecific low back pain. Eur Spine J 2006 Mar;15 Suppl 2:S192-300.

(37) Lee SC, Becker CN, Binder-Macleod SA. Activation of human quadriceps femoris muscle during dynamic contractions: effects of load on fatigue. J Appl Physiol 2000 Sep;89(3):926-936.

(38) Rousson V. Assessing inter-rater reliability when the raters are fixed: Two concepts and two estimates. Biom J 2011 Mar 21.

(39) Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428.

(40) 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 Aug;74(8):777-788.

(41) Kang YM, Choi WS, Pickar JG. Electrophysiologic evidence for an intersegmental reflex pathway between lumbar paraspinal tissues. Spine (Phila Pa 1976) 2002 Feb 1;27(3):E56-63.

(42) Holm S, Indahl A, Solomonow M. Sensorimotor control of the spine. J Electromyogr Kinesiol 2002 Jun;12(3):219-234.

(43) Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998 Jan 15;17(1):101-110.

(44) Wright JG. The minimal important difference: who's to say what is important? J Clin Epidemiol 1996 Nov;49(11):1221-1222.

(45) Dankaerts W, O'Sullivan P, Burnett A, Straker L, Davey P, Gupta R. Discriminating healthy controls and two clinical subgroups of nonspecific chronic low back pain patients using trunk muscle activation and lumbosacral kinematics of postures and movements: a statistical classification model. Spine (Phila Pa 1976) 2009 Jul 1;34(15):1610-1618.