

Reproducibility of the cervical range of motion (CROM) device for individuals with sub-acute whiplash associated disorders

Mark A. Williams · Esther Williamson ·
Simon Gates · Matthew W. Cooke

Received: 3 February 2011 / Revised: 17 November 2011 / Accepted: 19 November 2011 / Published online: 3 December 2011
© Springer-Verlag 2011

Abstract

Purpose The objective of this study was to evaluate the reproducibility of the cervical range of motion device when measuring both active and passive range of motion in a group of individuals with sub-acute Whiplash Associated Disorders (WAD).

Methods Participants were recruited as part of a large multi-centre Randomised Controlled Trial from UK emergency departments. Experienced research physiotherapists measured active and passive cervical spine movements in all directions. Both intra- and inter-observer reliability and agreement were assessed using the intra-class correlation coefficient, standard error of measurement and limits of agreement methods.

Results Different groups of 39 and 19 subjects were included in the intra and inter-observer studies, respectively. The CROM device demonstrated substantial intra- and inter-observer reliability and agreement for all the active and passive half-cycle movements (ICC range 0.82–0.99) with the exception of one (passive right lateral flexion for inter-observer; ICC 0.77).

Conclusions The CROM device has proven to be a reproducible measurement method for a symptomatic WAD population using the measurement protocol described and can be used with confidence to differentiate individuals according to a single measurement.

Keywords MeSH term ‘Reproducibility of results’ · Cervical spine · CROM

Introduction

Evaluation of cervical spine range of motion (ROM) is commonly undertaken by health care professionals to assist with diagnosis, prognosis and treatment evaluation. Active (patient controlled) and passive (clinician controlled) movements are measured and are purported to provide distinct and valuable information about contractile and non-contractile structures [1]. Cervical spine dysfunction and resulting reduced ROM may occur as a result of insidious or traumatic causes. An example is Whiplash Associated Disorders (WAD), the myriad of signs and symptoms that result from a whiplash mechanism of injury [2], where patients are found to have ROM reduced by approximately 25% compared to asymptomatic populations [3, 4]. ROM measurements have also demonstrated the potential for categorising patients who are at risk of a poor outcome following a whiplash injury [5, 6].

It is important that measurements of ROM are reproducible (stable over time for unchanged individuals) in order to be confident when differentiating between individuals and/or evaluating progress. It is challenging to measure cervical ROM accurately because of the lack of bony landmarks and natural coupling of movements [7], with measurements being influenced by intra-observer or inter-observer variation and clarity of end point due to pain. Previous systematic reviews of the literature [8–10] have concluded that of the numerous tools available, the cervical range of motion (CROM) device has most frequently been evaluated for validity and reproducibility and shown to be worthy of merit.

M. A. Williams (✉) · E. Williamson · S. Gates
Clinical Trials Unit, Warwick Medical School,
University of Warwick, Coventry CV4 7AL, UK
e-mail: m.a.williams.3@warwick.ac.uk

M. W. Cooke
Health Sciences Research Institute, Warwick Medical School,
University of Warwick, Coventry CV4 7AL, UK

There have been 11 published reliability studies (see [10] for references), all concluding that the device is reliable within and between testers. The majority ($n = 8$) of these studies have measured active rather than passive ROM and no studies have utilised both the methods in the same cohort. Only a minority of studies ($n = 3$) have investigated symptomatic participants. Findings from four concurrent validity studies [11–14] have indicated that the CROM has good concurrent validity for active ROM.

The reproducibility of the CROM device has yet to be evaluated in a WAD population, and therefore it is unknown whether the previous promising clinimetric results are applicable to this group.

The aim of this study was to evaluate the intra- and inter-observer reproducibility for the CROM device in a sub-acute WAD population, when measuring both active and passive ROM.

Methods

Participants previously recruited into a large multi-centre RCT [15] evaluating conservative treatment for WAD were invited to participate in this study. Potential participants were those that attended a participating UK Emergency Department (ED) following a whiplash injury. They were asked to contact researchers if they were continuing to have cervical spine symptoms approximately 2 weeks after their ED visit. If they appeared eligible according to the selection criteria below, they were provided with verbal information, and a hospital clinic appointment was arranged. An information sheet was sent to patients at least 24 h prior to the appointment. At the appointment, eligibility was checked, written informed consent was obtained, a questionnaire completed (providing awareness of general clinical status) and ROM measurements were taken for all the patients wishing to participate. The studies were approved by the Trent Multicentre Research Ethics Committee and by the Local Research Ethics Committee, and the Research and Development Committee of each participating centre.

Selection criteria

The following criteria were used to select participants;

- aged 18 years or over,
- experienced a whiplash injury less than 6 weeks before the initial recruitment in the Emergency Department,
- WAD grades I–III reported in the last 24 h,
- no fractures of spine or other bones,
- able to provide written informed consent.

Device

The CROM device (see Fig. 1) consists of two gravity-dependent goniometers and one compass dial on a head-mounted frame allowing measurement of ROM in three planes. A magnetic yoke consisting of two bar magnets held anteriorly and posteriorly is supplied to reduce the influence of thorax rotation. ROM is measured in 2° increments. The CROM device is manufactured by Performance Attainments AssociatesTM with a retail price of \$395.

Measurement protocol

The measurement protocol was developed using CROM device manual (Performance Attainments AssociatesTM) and recommendations from studies and systematic reviews of measurement reliability. Researchers were provided with a manual which included a step-by-step guide to the protocol.

The participant was asked to sit with hips and knees at 90° , feet flat on the floor, arms resting in lap and as far as possible a neutral pelvic position. The measurement process was explained to the participant and then the CROM device was placed on their head. For active movements participants were asked to move their head at a steady pace as far as they felt able, whilst keeping back and shoulders still. The researcher demonstrated the movements to the participant beforehand and explained how participants should attempt to minimise the coupling of rotation and side flexion movements. They were asked to pause at the end of each movement so a reading could be taken and then return to 'neutral' for a 5–10 s rest. Researchers recorded ROM in degrees after each movement. There were no warm-up movements performed with a single movement performed in each direction.



Fig. 1 CROM device with magnetic yoke

For passive movements researchers moved the participant's head as far as the participant would allow indicating this by raising an arm. This was performed in the same upright sitting position. For both active and passive movements, the sequence of the movements was flexion, extension, right rotation, left rotation, right lateral flexion and left lateral flexion.

Intra-observer (IaO) reliability was evaluated by a single observer (MAW) repeating the measurement of the sequence of movements after a 2 min rest period. Inter-observer (IeO) reliability was evaluated by two observers (MAW and EMW) in an order determined randomly by a computer-generated sequence table. Measurements of the first sequence of movements for both studies were blinded using a separate recording sheet, and only one researcher was present in the examination area at any one time for the inter-observer study.

Both the observers were research physiotherapists (MAW and EW) with 5 and >10 years clinical experience. They undertook 2 h training and practise on healthy volunteers using the CROM device and had subsequently used the protocol on patients in a research clinic setting for at least 6 months prior to the study commencing. It was decided that two separate groups of participants would be recruited for the intra- and inter-observer studies. This was due to participants' symptoms being potentially too irritable to complete the multiple repetitions required to conduct the two studies in one appointment.

Statistical analysis and sample size

All statistical analyses were performed using the Statistical Package for Social Sciences (SPSS 17.0, Chicago, IL, USA). Parameters for both reliability and agreement were calculated.

Intra Class Correlation Coefficient (ICC) (2, 1) with 95% confidence intervals (95% CI), Standard Error of Measurements (SEM) and Limits of Agreements (LoA) [16] were calculated for all the active and passive

movements. LoA estimate agreement by presenting values 2 SD above and below the mean difference of the measurements. A value for total active and passive ROM was calculated as the sum of the six half-cycle movements [5] and reliability and agreement parameters calculated for these also.

In order to demonstrate a 'substantial' level of reliability, it was decided that an ICC value of >0.8 would be required [17]. Using calculations devised by Walter et al. [18] for 90% power of testing, a 5% significance level a sample size of 39 subjects was required for both the IaO and IeO studies.

Results

Intra-observer study

39 participants were recruited between April 2006 and February 2007. One subject was unable to complete the ROM assessment after consenting due to symptoms being too severe. Characteristics for the remaining 38 participants are displayed in Table 1. This sample was made up of a young working population who generally had experienced a whiplash injury as a result of a motor vehicle collision. On average participants were assessed 1 month following injury, and had moderate pain and disability levels.

Mean ROM, ICC (95% CI), SEM and LoA ranges are presented in Table 2. Both active and passive ROMs were 25% less than the expected values for the accepted normal values [19].

ICC values ranged from 0.98 to 0.99 for both active and passive movements indicating a high level of reliability. The SEM ranged from 1.3° to 2.1°. Figures 2 and 3 display intra-observer limit of agreement plots for total active and passive ROM, respectively. There were slightly wider limits of agreement for total passive ROM (−16.0° to 10.1°) compared to active ROM measurements (−15.5° to 9.3°).

Table 1 Characteristics of intra-observer (IaO) and inter-observer (IeO) study participants

	Mean (SD) unless stated	
	IaO study (<i>n</i> = 38)	IeO study (<i>n</i> = 19)
Sex (F:M)	19:19	13:6
Age	38 (11.3)	41 (14.8)
WAD grade, <i>n</i> (%)	I = 2 (5), II = 34 (90), III = 2 (5)	I = 0, II = 18 (95), III = 1 (5)
Injury due to MVC, <i>n</i> (%)	37 (97)	17 (90)
Days between injury and Ax	27 (8.4)	35 (9.2)
Pain VAS 0–10	6 (2.3)	5 (2.4)
NDI score	22 (9.1)	21 (9.7)

Table 2 ROM summary and reliability statistics for intra-observer and inter-observer studies

Movement direction	Normative mean (SD) ROM (°) [19]	Results for IaO study				Results for IeO study			
		Mean (SD) ROM (°)	ICC (95% CI)	SEM/ (°)	LoA/(°)	Mean (SD) ROM (°)	ICC (95% CI)	SEM (°)	LoA (°)
Active flexion	52 (7)	38 (14.3)	0.99 (0.98–0.99)	1.4	−4.3 to 2.9	33 (15.8)	0.83 (0.61–0.93)	6.5	−21.6 to 16.1
Active extension	71 (5)	41 (16.4)	0.99 (0.98–0.99)	1.6	−5.0 to 5.0	41 (18.1)	0.88 (0.72–0.96)	6.3	−18.9 to 18.0
Active Right Rotation	73 (11)	51 (14.1)	0.98 (0.96–0.99)	2	−6.4 to 4.5	45 (17.5)	0.92 (0.80–0.97)	4.9	−13.5 to 15.5
Active left rotation	71 (11)	51 (15.0)	0.99 (0.98–0.99)	1.5	−4.6 to 3.4	45 (17.9)	0.87 (0.68–0.95)	6.5	−13.8 to 21.5
Active right lateral flexion	44 (0)	26 (9.4)	0.98 (0.96–0.99)	1.3	−4.2 to 3.0	25 (8.7)	0.82 (0.59–0.92)	3.7	−12.0 to 10.0
Active left lateral flexion	42 (2)	34 (9.8)	0.98 (0.96–0.99)	1.4	−4.3 to 3.3	32 (10.6)	0.88 (0.70–0.95)	3.7	−7.9 to 12.4
Total active ROM		241 (66.3)	0.99 (0.99–0.99)	6.6	−15.5 to 9.3	222 (79.0)	0.95 (0.86–0.98)	17.7	−50.7 to 56.4
Passive flexion		35 (14.8)	0.98 (0.96–0.99)	2.1	−6.5 to 4.6	32 (17.0)	0.90 (0.76–0.96)	5.4	−16.9 to 14.0
Passive extension		42 (17.8)	0.99 (0.99–1.00)	1.8	−4.3 to 4.4	40 (19.4)	0.96 (0.89–0.98)	3.9	−9.5 to 12.8
Passive right rotation		50 (18.8)	0.99 (0.97–0.99)	1.9	−6.9 to 5.1	44 (18.4)	0.89 (0.67–0.96)	6.1	−10.3 to 19.9
Passive left rotation		55 (20.3)	0.99 (0.98–0.99)	2	−6.6 to 6.3	45 (22.0)	0.85 (0.60–0.94)	8.5	−17.4 to 29.4
Passive right lateral flexion		26 (11.0)	0.98 (0.97–0.99)	1.6	−4.4 to 4.0	23 (10.1)	0.77 (0.36–0.92)	4.8	−16.6 to 7.9
Passive left lateral flexion		32 (9.3)	0.98 (0.95–0.99)	1.3	−4.3 to 2.6	30 (10.5)	0.88 (0.70–0.95)	3.6	−8.4 to 12.2
Total passive ROM		240 (80.1)	0.99 (0.99–0.99)	8	−16.0 to 10.1	213 (90.4)	0.96 (0.90–0.99)	18.1	−42.1 to 59.3

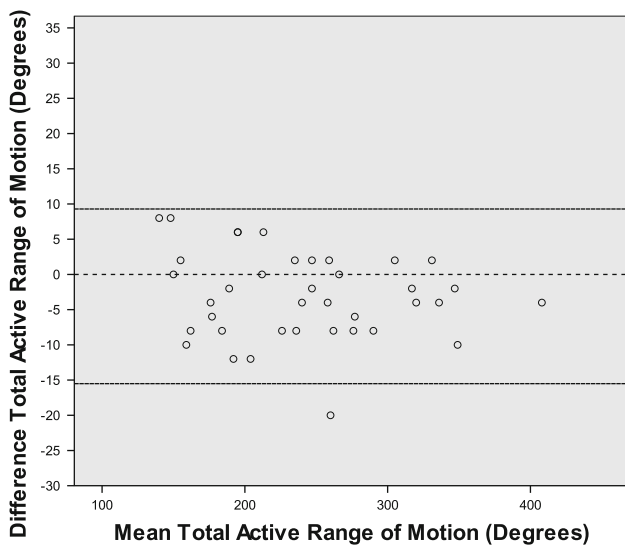


Fig. 2 Intra-observer study limit of agreement plot for total active cervical ROM ($n = 38$)

Inter-observer study

19 participants were recruited between March and November 2007, when the supplying RCT completed recruitment. Characteristics for the 19 participants are displayed in Table 1. A similar demographic to the intra-observer study was recruited in that participants were, on

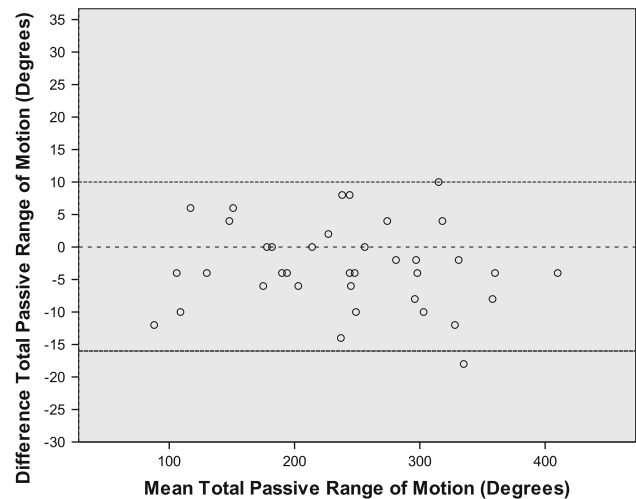


Fig. 3 Intra-observer study limit of agreement plot for total passive cervical ROM ($n = 37$)

average, of working age, had sub-acute WAD as a result of a motor vehicle collision and reported moderate pain and disability.

Mean ROM, ICC (95% CI) and SEM are presented in Table 2. As with the intra-observer study, ROM in all the planes of motion was consistently less than expected normal values. Intra class correlation coefficients ranged from 0.82 to 0.95 and 0.77 to 0.96 for active and passive half-cycle

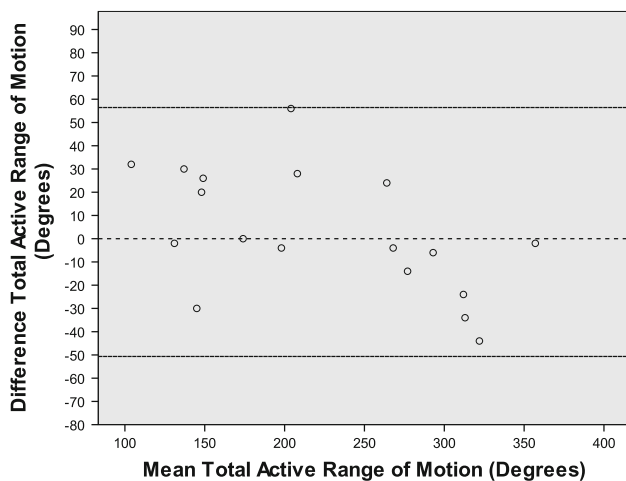


Fig. 4 Inter-observer study limit of agreement plot for total active cervical ROM ($n = 18$)

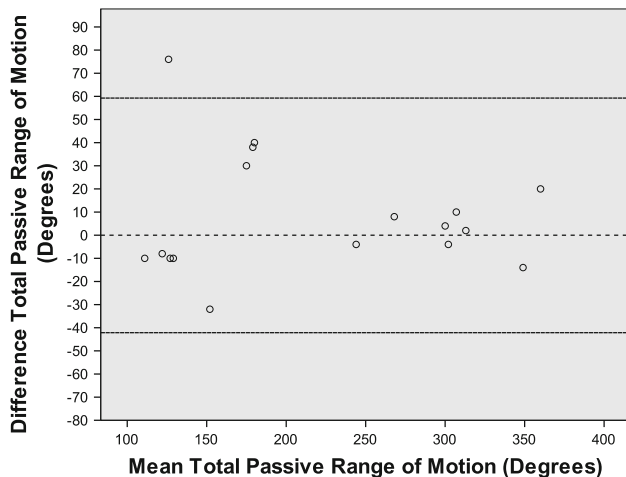


Fig. 5 Inter-observer study limit of agreement plot for total passive cervical ROM ($n = 18$)

ROM measurements. Standard error of measurement ranged from 3.6° to 8.5° for half-cycle ROM measurements. Limits of agreement ranged from -6.4 to 5.0 and -6.9 to 6.3 for active and passive half-cycle measurements respectively. Figures 4 and 5 display inter-observer limit of agreement plots for total active and passive ROM respectively. Alternative to the intra-observer study, total passive ROM had slightly narrower limits of agreement (-42.1 to 59.3) than the active ROM measurements (-50.7° to 56.4°). Limits of agreement were wider for the inter-observer study when compared to the results of the intra-observer study.

Discussion

Of the numerous tools available for measuring cervical spine ROM the CROM device has already undergone the most

clinimetric evaluation. This has shown the CROM to be a valid and reproducible method. The current study adds to this evidence base by evaluating both active and passive methods of assessment in a single cohort and in a pre-specified symptomatic population. Evaluation in symptomatic clinical populations is vital to provide clinicians with confidence in the results of their clinical measurements.

The results presented indicate “substantial” intra- and inter-tester reliability according to Shrout [17] over the short-term measurement periods. One inter-observer measurement just failed to prove substantial reliability (Passive RLF-ICC 0.77) although the confidence interval is wide and this finding may be due to the small sample size. As expected, intra-observer was greater than the inter-observer reliability as demonstrated by both the ICC and LoA results. It has been reported that reliability is reduced when evaluating symptomatic populations compared to healthy controls [20]; however the findings of this study do not support this—the results were comparable with the previous investigations of CROM device reliability using asymptomatic participants [21, 22]. The substantial reliability demonstrated may have in part have been due to the robust training and standardised measurement protocol.

There did not appear to be any great difference in reliability between active and passive methods of measurement. Only a very small number ($\sim 10\%$) of previous reliability studies of any devices have evaluated both at the same time, none of which have used the CROM device. The findings of this study compare favourably to other studies evaluating the reliability of measuring passive ROM [23, 24].

Mean active and passive lateral flexion ROM was reduced to the right compared to the left. We could speculate that this may be the result of the small sample size and therefore a small number of low readings for right lateral flexion had a considerable influence on the mean score. However, it is not uncommon for studies of symptomatic cervical spine populations to find asymmetries in single plane ROM measurements (e.g. [25]). We have recorded a similar finding from a much larger cohort of whiplash-injured patients ($n = 599$, data currently unpublished). There are numerous potential reasons for the asymmetry, not least mechanism of injury in this population.

We calculated ICC, SEM and LoA in order to provide representation of different aspects of reproducibility (reliability and agreement). Failing to do so has been a common criticism of previous studies [9, 19].

It is acknowledged that there are limitations to the studies reported. The significant concern is the small sample size of the inter-observer study ($n = 19$). The target sample size of 39 participants was not achieved. At the time of commencing both studies, it was believed that there would be enough participants from the supplying RCT to

reach the target sample size for both studies. However, the final phase of recruitment to the supplying RCT tended to occur at sites other than those where the reliability studies were being conducted. Failing to achieve the target sample size reduces the confidence in the results of the study and may lead to a greater chance of a type I or II errors. However, the 95% confidence intervals around the ICC estimates are relatively narrow, especially for the intra-observer study, indicating the variability wasn't excessive. For the inter-observer study, the lower 95% confidence interval scores still indicate a moderate level of reliability [17] for the majority of individual movements and substantial reliability for the sum total ROM score. Passive right lateral flexion appears to be an exception with a lower 95% CI ICC equivalent to slight reliability and therefore changes in this measurement when different observers perform measurements may need to be interpreted with more caution.

A small number of patients withdrew from the study ($n = 1$), and therefore although noted for completeness of reporting, the authors feel this is unlikely to significantly bias the findings.

We evaluated reproducibility over short time periods (approximately 2 min for IaO and 5 min for IeO). There were a number of reasons for doing this. First, from a practical perspective it allowed us to measure participants during one hospital visit. Second, with the symptomatic population studied, we wanted to minimise the variation due to biological and symptomatic factors. With the short re-testing period of the intra-observer study there was potential for the introduction of recall bias. The assessor (and author) was aware of this possibility and sought to minimise this by using separate recording sheets and the distraction of performing all the movements in one cycle, and then repeating after the rest period. It would be advantageous to conduct further studies using longer test-retest periods, which are more commonly used in a clinical setting (e.g. 1 week).

The researchers found the CROM easy to fit (no need for reliance on anatomical landmarks) and read following a small amount of training which is an important consideration for implementation into clinical practice. The main barriers to widespread clinical use may be the current cost of the CROM device and its use being limited to measuring cervical spine ROM, although the CROM is as reliable as more expensive and complex cervical ROM measurement systems [26].

Conclusions

It is concluded from this study that the CROM device produces substantially reliable ROM measurements for

both active and passive methods in a WAD population. Reproducibility has been demonstrated for both repeated use by one observer and use by two different observers. The confidence in the results of the inter-observer study is reduced due to the small sample size.

Acknowledgments The authors are grateful to all the participants of the study.

Conflict of interest None.

References

- Maitland GD (2001) Maitland's vertebral manipulation. Butterworth Heinemann, Oxford
- Spitzer WO, Skovron ML, Salmi LR, Cassidy JD, Duranceau J, Suissa S, Zeiss E (1995) Scientific monograph of the Quebec Task Force on whiplash-associated disorders: redefining "whiplash" and its management. *Spine* 20:1S–73S [erratum appears in *Spine* 1995 Nov 1;20(21):2372]
- Dall'Alba PT, Sterling MM, Treleaven JM, Edwards SL, Jull GA (2001) Cervical range of motion discriminates between asymptomatic persons and those with whiplash. *Spine* 26:2090–2094
- Cagnie B, Cools A, De Loose V, Cambier D, Danneels L (2007) Reliability and normative database of the Zebris cervical range-of-motion system in healthy controls with preliminary validation in a group of patients with neck pain. *J Manipulative Physiol Ther* 30:450–455
- Kasch H, Bach FW, Jensen TS (2001) Handicap after acute whiplash injury: a 1-year prospective study of risk factors. *Neurology* 56:1637–1643
- Hartling L, Brison RJ, Ardern C, Pickett W (2001) Prognostic value of the quebec classification of whiplash-associated disorders. *Spine* 26:36–41
- Bogduk N, Mercer S (2000) Biomechanics of the cervical spine. I: Normal kinematics. *Clin Biomech (Bristol, Avon)* 15:633–648
- Jordan K (2000) Assessment of published reliability studies for cervical spine range-of-motion measurement tools. *J Manipulative Physiol Ther* 23:180–195
- de Koning C, van den Heuvel S, Staal J, Smits-Engelsman B, Hendriks E (2008) Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: a systematic review. *Eur Spine J* 17:905–921
- Williams M, McCarthy C, Chorti A, Cooke M, Gates S (2010) A systematic review of reliability and validity studies of methods for measuring active and passive cervical range of motion. *J Manipulative Physiol Ther* 33:138–155. doi:10.1016/j.jmpt.2009.12.009
- Hole DE, Cook JM, Bolton JE (1995) Reliability and concurrent validity of two instruments for measuring cervical range of motion: effects of age and gender. *Man ther* 1:36–42
- Tousignant M, de Bellefeuille L, O'Donoghue S, Grahovac S (2000) Criterion validity of the cervical range of motion (CROM) goniometer for cervical flexion and extension. *Spine* 25:324–330
- Tousignant M, Duclos E, Lafleche S, Mayer A, Tousignant-Laflamme Y, Brosseau L, O'Sullivan JP (2002) Validity study for the cervical range of motion device used for lateral flexion in patients with neck pain. *Spine* 27:812–817
- Tousignant M, Smeesters C, Breton AM, Breton E, Corriveau H (2006) Criterion validity study of the cervical range of motion

- (CROM) device for rotational range of motion on healthy adults. *J Orthop Sports Physic Ther* 36:242–248
15. Lamb SE, Gates S, Underwood MR, Cooke MW, Ashby D, Szczepura A, Williams MA, Williamson EM, Withers EJ, Mt Isa S, Gumber A, Team MS (2007) Managing Injuries of the Neck Trial (MINT): design of a randomised controlled trial of treatments for whiplash associated disorders. *BMC Musculoskelet Disord* 8:7
 16. Bland J, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307–310
 17. Shrout PE (1998) Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 7:301–317
 18. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. *Stat Med* 17:101–110
 19. Chen J, Solinger AB, Poncet JF, Lantz CA (1999) Meta-analysis of normative cervical motion. *Spine* 24:1571–1578
 20. Assink N, Bergman GJ, Knoester B, Winters JC, Dijkstra PU, Postema K (2005) Interobserver reliability of neck-mobility measurement by means of the flock-of-birds electromagnetic tracking system. *J Manipulative Physiol Ther* 28:408–413
 21. Youdas JW, Carey JR, Garrett TR (1991) Reliability of measurements of cervical spine range of motion—comparison of three methods. *Phys Ther* 71:98–104
 22. Peolsson A, Hednlund R, Ertzgaard S, Oberg B (2000) Intra- and inter-tester reliability and range of motion of the neck. *Physiother Can Summer* 52:233–242
 23. Love S, Gringmuth RH, Kazemi M, Cornacchia P, Schmolke M (1998) Interexaminer and intraexaminer reliability of cervical passive range of motion using the CROM and Cybex 320 EDI. *J Can Chiropr Assoc* 42:222–228
 24. Nilsson N, Christensen HW, Hartvigsen J (1996) The interexaminer reliability of measuring passive cervical range of motion, revisited. *J Manipulative Physiol Ther* 19:302–305
 25. Dvir Z, Gal-Eshel N, Shamir B, Prushansky T, Pevzner E, Peretz C (2006) Cervical motion in patients with chronic disorders of the cervical spine: a reproducibility study. *Spine* 31:E394–E399
 26. Gelalis I, DeFrate L, Stafilas K, Pakos E, Kang J, Gilbertson L (2009) Three-dimensional analysis of cervical spine motion: reliability of a computer assisted magnetic tracking device compared to inclinometer. *Eur Spine J* 18:276–281