Reliability, usability and safety of a novel inertial sensor-based system to test craniocervical flexion movement control

Elena Bocos-Corredor
Universidad San Pablo-CEU, CEU Universities

Filippo Moggioli (filippo.moggioli@ceu.es)
Universidad San Pablo-CEU, CEU Universities

Tomás Pérez-Fernández
Universidad San Pablo-CEU, CEU Universities

Susan Armijo-Olivo
University of Applied Sciences

Sonia Liébana
Universidad San Pablo-CEU, CEU Universities

Aitor Martín-Pintado-Zugasti
Universidad San Pablo-CEU, CEU Universities

Research Article

Keywords:

Posted Date: January 24th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3671738/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.
Abstract

Background

Neck pain has a significant global impact, ranking as the fourth leading cause of disability. Recurrent neck pain often leads to impaired sensorimotor control, particularly in craniocervical flexion (CFF). The Craniocervical Flexion Test (CCFT) has been widely investigated to assess the performance of deep cervical flexor muscles. However, its use requires skilled assessors who need to subjectively detect compensations, progressive increases in range of motion (ROM) or excessive superficial flexor activation during the test. The aim of this study was to design and develop a novel Craniocervical Flexion Movement Control Test (CFMCT) based on inertial sensor technology and real-time computer feedback and to evaluate its safety and usability, as well as inter and intra-rater reliability in both healthy individuals and patients with neck pain.

Methods

We used inertial sensor technology associated with new software that provides real-time computer feedback to assess CCF movement control through two independent test protocols, the progressive consecutive stages protocol (including progressive incremental stages of ROM) and the random stages protocol (providing dynamic and less predictable movement patterns). We determined intra and inter-rater reliability and standard error of the measurement for both protocols. The participants rated their usability was analysed through the System Usability Scale (SUS) and possible secondary effects associated with the tests were registered.

Results

The progressive consecutive stages protocol and the random stages protocol were safe and easy to use (SUS scores of 82.00 ± 11.55 in the pain group and 79.56 ± 13.36 in the asymptomatic group). The progressive consecutive stages protocol demonstrated good inter-rater reliability (intraclass correlation coefficient [ICC] ≥ 0.75) and moderate to good intra-rater reliability (ICC 0.62–0.80). However, the random stages protocol exhibited lower intra-rater reliability, especially in the neck pain group, where the reliability values were poor in some cases (ICC 0.48–0.72).

Conclusion

The CFMCT (progressive consecutive stages protocol) is a promising instrument to evaluate CCF motor control in patients with chronic neck pain. It has potential for telehealth assessment and easy adherence for exercise prescription and proved to be a safe and usable tool for patients with pain and asymptomatic
individuals. Its use as a test or for exercise needs to be further investigated to facilitate its transfer to clinical practice.

**INTRODUCTION**

Neck pain represents one of the main social, economic and health problems. It is ranked as the fourth leading cause of disability worldwide with a prevalence of 15–75.1%. Although the history of recovery from neck pain appears to be favourable, the rate of chronicity and recurrence is up to 30%, and neck pain can persist for more than 6 months in 14% of patients (1).

Patients with recurrent neck pain have shown impaired sensorimotor control during analytical or functional movements, evidenced as altered kinematics or proprioception compared with asymptomatic individuals (2–5). Craniocervical flexion (CCF) is one of the most studied movements within the neck because its performance is associated with activation of deep cervical neck flexors (DCF): longus colli and longus capitis. These muscles are believed to be crucial for stability of the cervical region, which is commonly altered in patients with neck pain and associated disorders (6–9).

The Craniocervical Flexion Test (CCFT) specifically assesses the degree of activation and performance of the DCFs during CCF. The participant is placed in the supine position with a pressure biofeedback sensor between the cervical segment and a bed. Five pressure stages (from 20 to 30 mmHg) must be reached with a progressive increase of 2 mmHg at each stage by attaining the cervical lordosis and avoiding compensatory movements of the superficial muscles (10–14). In this way, the isolated upper craniocervical movement as well as the activity of the DCFs are determined.

The construct validity of the CCFT regarding activation of the DCFs has been investigated in both asymptomatic people and patients with neck pain by using nasopharyngeal electromyography (EMG) (9, 11, 12, 15). Correct performance of the test involves reaching the stages without excessive activation of the superficial flexors, a phenomenon that frequently occurs in patients with neck pain compared with asymptomatic individuals (12, 13, 16). Another construct of the CCFT is based on the progressive increase in the CCF range of motion (ROM) with each incremental stage of the test, which has been evidenced by using digital imaging methods (7, 17) or inertial sensor technology (18, 19). The clinical implementation of the CCFT does not require EMG or ROM analysis technology, however, clinicians need to be trained and skilled to be able to monitor the compensatory activation of the superficial flexors by palpation and the progressive increase ROM visually, and to detect any other compensatory strategies such as head retraction.

In the recent review, Romeo et al., (20) investigated the psychometric properties of the CCFT concluding that, although it is a useful test for detecting impairments in DCF muscles’ control, its suitability in clinical practice is restricted due to its limited reliability and standard error of the measurement (SEM) values. The authors associated this limited reliability to the difficulties in detecting compensations and the subjectivity of the criteria to detect ‘excessive activity’ of the superficial cervical flexors.
Due to the belief that poor CCF performance is linked to impaired activation of DCFs and is a characteristic of patients with neck pain or headache (21–26), other studies have investigated alternative ways to evaluate CCF performance and control based on visual inspection (27, 28) or inertial sensor analysis (29). However, these alternative methods did not show clear usefulness to discriminate between patients with pain and asymptomatic controls when evaluating CCF.

Previous research has proposed the possibility of using new inertial sensor technologies to consider CCF ROM as the guide to assess or train DCF based on the rationale of the original CCFT, by defining potential targets of ROM equivalent to each pressure targets of the CCFT (19). This approach allows the possibility of developing new computer-based tools with real-time biofeedback systems; these tools may be more precise and less influenced by the subjectivity of the rater to detect compensations during the test. These new technologies could potentially be used with patients in positions other than supine and may allow for telerehabilitation through the use of sensors to evaluate a patient or to deliver rehabilitation at a distance, thus improving accessibility to healthcare services (30).

In the present study, we propose a novel inertial sensor-based system to test CCF movement control. It should be highlighted and acknowledged that this system is based on the rationale and previous research on the CCFT performed by the developers of this test (23, 31–34).

The main objectives of this study were:

1. to design and develop a beta version of a new test to assess CCF movement control and performance through a system of inertial sensor technology and computer-based visual feedback, namely the Craniocervical Flexion Movement Control Test (CFMCT), and

2. to determine the safety, usability and inter-rater and intra-rater reliability of this new tool in asymptomatic individuals and patients with chronic neck pain.

Methods

Design

This is a methodological reliability study of a novel system to evaluate CCF movement control using the CCF ROM to define the targets. The study followed a reproducibility protocol with three phases: training, overall agreement and study [35]. The training and overall agreement phases have already been finished and were the basis of the design used in the present study. Therefore, this study addresses the study phase of the reproducibility protocol, as well as the analysis of its safety and usability. In the training and overall agreement periods, the examiners designed and developed the new tool and discussed and agreed on how the test should be performed, the use of the inertial wearable sensors and the associated software. The procedure described in this section to assess the reliability was based on the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) methodology [36].
The study was approved by the Research Ethics Committee of CEU San Pablo University (236/17/08) and performed in accordance with the Declaration of Helsinki and World Health Organization (WHO) standards (37). The participants received a complete verbal description of the procedures and the purpose of the study and gave their written informed consent before enrolling in the study. They were able to withdraw at any time during the study.

Sample and selection

Both asymptomatic participants and patients with chronic neck pain between 18 and 65 years old were invited to participate. They were recruited through convenience sampling via e-mails and informational posters at the CEU San Pablo University and the CEU San Pablo University clinic. The inclusion criteria for patients with chronic neck pain included the following: > 3 months of pain duration, a score of > 4 on the Neck Disability Index (NDI) [38] and a score of > 3 on the visual analogue scale (39,40). Healthy participants had to have a score of 0 on the NDI and the visual analogue scale. The exclusion criteria were previous surgery in the neck or head area, previous diagnosis of headaches and temporomandibular disorders and any neurological deficits. Immediately after the participant signed the informed consent and was screened based on the selection criteria, they provided information about their demographic characteristics (age, weight, height and occupation).

CFMCT measured by inertial sensor technology

The construct of this novel test is the flexion movement control capacity through different levels of the whole CCF ROM of the participant. The tests require the subject to accurately achieve and maintain different CCF ROM targets guided by software displayed on a computer/tablet screen as biofeedback. This section describes the inertial sensor-based system used for the novel CFMCT based on the list of components of outcome measurement instruments of the COSMIN methodology.

Preparatory actions preceding data collection

Preparation of the participant by the examiner

The participant was placed in the supine position with their arms resting on their abdomen, their knees bent and their feet resting on the table. The procedure for standardising the initial neutral cervical position was to ask the subject to perform three maximum CCF repetitions, relaxing the flexors to return from each maximum CCF to neutral while avoiding contraction of the extensor muscles. The initial position of the test was the one achieved after relaxing the flexors in the previous repetition. This procedure was designed to avoid a lack of standardisation of the initial position based on anatomical landmarks, which are highly influenced by anatomical variability between subjects.

Preparation of equipment by the examiner

Once the initial position was achieved, a single 4 × 4 × 8 cm wireless wearable inertial sensor (Werium Solutions©, Madrid, Spain) was placed on the participant’s forehead. Before starting each measurement,
the inertial sensor was calibrated to zero. This inertial sensor allows real-time tracking of the progressive ROM increase while performing CCF. This instrument has previously demonstrated good to excellent inter- and intra-rater reliability in the assessment of global neck movement (41) and CCF movement (42).

**Computer**

A computer screen, which displays the ROM during CCF captured by the inertial sensor, was positioned in front of the subject’s eyes at a comfortable distance. The computer screen was placed in a horizontal position parallel to the bed to standardise the initial position. This could be performed with the help of an articulated arm (Fig. 1). Real-time feedback was provided by innovative software that was created for the purpose of this study (described below).

**Characteristics and settings of the beta version of the software**

Once the participant was in the initial position, they saw two references to look at on the computer screen: first, a red line indicating the trajectory of the movement in terms of the achieved CCF ROM; second, blue stripes that they must reach and stay inside (see the supplementary video). To set the degrees of CCF ROM to be achieved in each blue stripe (levels), it was necessary to determine the maximum CCF of each participant, because the blue stripes represent the percentage of the maximum CCF ROM of each subject (see calculation below). Therefore, each participant performed three maximum CCF movements, and the software provided the examiner with the maximum CCF ROM reached by the subject.

The settings of the software to perform the CFMCT are described below (see Fig. 1).

1. **Number of stages:** this test was designed to reach a total of nine stages of a percentage of the maximum CCF ROM. The first five stages of ROM percentage targets were based on the CCFT pressure levels as reported in a previous study (19). The next four stages were added based on the mean average increment between the previous stages, so the last stage was the closest possible to the maximal CCF ROM.

2. **Stages based on ROM values:** the percentages of the maximum CCF ROM where the blue stripes are displayed on the screen were: stage 1, 20.7%; stage 2, 30.7%; stage 3, 45.2%; stage 4, 52.2%; stage 5, 61.6%; stage 6, 70%; stage 7, 79%; stage 8, 88%; stage 9: 97%.

3. **Stage level limits:** the limits for each stage were 0.3° above and below the target CCF ROM level.

4. **Duration of the stage (seconds):** the duration of each stage (blue stripes) was 3–4 seconds (depending on the protocol [see below]), while the time between the stages was 4 seconds.

5. **Appearance of the stage (seconds):** this setting dictated at what time (in seconds) on the timeline from the start each blue stripe would appear. On the screen, the subject saw how each blue stripe appeared in the previously configured stage.

**Test performance**
After performing the three CCF repetitions, the instructions for the participant were:

‘The red line you see on the screen corresponds to your movement. Your goal from now on is to reach and stay within the blue lines (stripes) that will sequentially appear on the screen performing the nodding action.’

We designed two different and independent protocols to assess CCF movement control. Both protocols were based on the stages of the CCFT explained above, but one of these protocols showed the nine stages in a progressive order (progressive consecutive stages protocol), while the other showed them in a predetermined random order (random stages protocol). The objective of both protocols was to measure the accuracy of the subjects to reach the targeted CCF ROM and be able to control the movement during the ROM.

For the progressive consecutive stages protocol (see the supplementary video), the nine stages increased progressively from 20.7% to 97% of maximum CCF ROM. After reaching each CCF stage, the subject must return to a neutral position, which is represented as a blue stripe at zero (same as the initial position where the sensor was calibrated). The stripes appeared for 3 seconds in the CCF position and for 4 seconds in the neutral position (0° of CCF; resting stages). The final score was based on the points obtained for all stages, including the neutral stages. The calculation of the final score is explained in more detail below.

For the random stages protocol, the previously mentioned stages were randomised, generating a more dynamic and less predictable movement pattern for the participant. The blue stripes appeared for 4 seconds and there were no resting stages; thus, the muscular demand was greater than for the progressive consecutive stages protocol. Once randomised, the template created with the same randomised order of the percentages was used for all participants.

Reliability design

The reliability of the CFMCT was assessed through both intra- and inter-rater reliability. During the first measurement day (first measurement), each subject performed both protocols with each examiner; therefore, each subject performed each protocol twice. This measurement protocol was repeated 1 week later (second measurement). Therefore, inter-rater reliability was calculated by comparing the measurements obtained on each measurement day between examiner 1 and examiner 2, and intra-rater reliability was calculated by comparing the measurements of each examiner obtained for the first and second measurement. The examiners were blinded regarding the protocol to minimise any potential influence or bias in the assessment. When one examiner was conducting both protocols with the subject, the other examiner was not present in the room.

Data processing and storage

Each measurement was automatically saved in the software and exported as an Excel file with the following information:
1. degrees of ROM flexion at each time point – indicates the CCF ROM every 20 milliseconds, when the inertial sensor records movement information;

2. stage reached – indicates whether the CCF ROM achieved at each time point was inside the blue stripe of the corresponding stage displayed at that time on the computer biofeedback.

Assignment of the score

The final score was calculated as the percentage of time points the participant was within the blue stripe area and ranges from 0% (lowest movement control) to 100% (highest movement control). Therefore, if a participant is within all blue stripes during the test, the total score is 100%.

Summary of the study protocol for reliability analysis

After providing information about their pain and demographic characteristics, the entire measurement procedure was explained to the participant. The study was conducted by two examiners, who each performed both protocols for each subject. The order of the examiners and the protocols was randomised by using an online randomiser (https://www.randomizer.org).

The first examiner guided the participant through a seated warm-up of three movements in each of the three planes of neck movement: flexion-extension, right and left tilts and right and left rotations. The subject also practised the CCF movement, which the examiner described as a gentle and slow head-nodding action of ROM flexion(10). The examiner gave the following instructions:

‘Please lower your chin towards your neck (nodding movement) in a controlled manner without moving your neck, such that the head rotates slightly, reaching as far as possible. The posterior side of the head will slide smoothly on the bed during the movement and the head should not separate from the bed or push into it during the movement’.

If necessary, the examiner corrected the movement verbally and manually (29)

Once the participant was lying on the bed, they performed the CCF movement as a training exercise. Once the subject had properly executed the nodding movement without any compensations after warming up, the measurement started. As mentioned above, the neutral cervical starting position was obtained after three maximum repetitions of CCF. At this point the inertial sensor was placed and calibrated. Each participant performed three maximum CCF repetitions to obtain the maximum CCF for each participant and to calculate the ROM of each stage of the test. At this point, the examiner explained to the participant that their objective was to reach the blue stripes that would appear on the screen and to remain steady for as long as possible. In addition, the examiner explained to the subject that they should reach the targets by correctly performing the previously learned CCF movement, without making any compensatory movements.
The participant performed each protocol once for each examiner, with a 5-minute resting period between the examiners. Therefore, by the end of the test each participant had performed each protocol twice.

After completing the test, each participant filled out a data sheet that included the following: (a) the intensity of pain during the test based on the visual analogue Scale (43), (b) muscular fatigue during the test based on the fatigue VAS, (c) any other perceived symptoms or adverse event (e.g. nausea, dizziness) via an open question and (d) the System Usability Scale (SUS) (44). The SUS includes 10 statements about the perceived usability of the software. Each statement received a score of 1–5, where 5 corresponds to strongly agree and 1 corresponds to strongly disagree. The odd-numbered questions constitute positive statements and are scored by subtracting one point from the user's rating. The even-numbered questions represent negative statements and are scored by subtracting 5 points from the user's rating. To obtain the final SUS score, the sum of all answers is multiplied by 2.5. The resulting final score ranges from 0 (indicating low usability) to 100 (reflecting high usability) (45). A final score between 68 and 84 points indicates good usability of the system and a score > 84 points indicates excellent usability (46).

**Data analysis plan**

All data were analysed using SPSS Statistics version 27.0 (IBM Corp., Armonk, NY USA) and R version 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria). The Kolmogorov–Smirnov test was used to determine whether the quantitative data followed a normal distribution; it showed that the data did not follow a normal distribution. Quantitative variables are presented as mean ± standard deviation (SD) and categorical variables are presented as absolute values and frequencies (%). The final score of each test was calculated as the percentage of time points the participant was within the blue stripes.

The reliability of the CFMCT was determined with intraclass correlation coefficients (ICCs) based on a two-way random effects model, an average measure and an agreement definition (47) The reliability analysis included intra-rater reliability (first measure versus second measure) and inter-rater reliability (examiner 1 versus examiner 2) of the measures. Based on the 95% confidence interval, the ICCagreement was interpreted as follows: < 0.5 indicates poor reliability, between 0.50 and 0.75 indicates moderate reliability, between 0.75 and 0.90 indicates good reliability and > 0.90 indicates excellent reliability (48).

The SEM was calculated using the formula SEM = SD × \[\sqrt{(1 – ICCagreement)}\] (47)

The sample size was based on international guidelines for reliability studies that have reported 40 subjects are sufficient for reproducibility.(49)

**Results**

Eighty-four participants met the inclusion criteria; however, we excluded four participants who did not attend the second test session 1 week later. Eighty participants (40 with neck pain, 40 asymptomatic) successfully completed the study protocol and we included them in the study. The neck pain group
included 25 women and 15 men. Their mean ± SD age was 35.7 ± 14.1 years. In addition, in the last month, the participants with neck pain had a mean ± SD VAS score of 4.6 ± 1.7 and a mean ± SD NDI score of 10.37 ± 6.11. The asymptomatic group included 19 women and 21 men. Their mean ± SD age was 33.05 ± 14.63 years. All patients had a score of 0 for the VAS and NDI. The mean total percentage scores (% SCORE) and SDs of the CFMCT are presented in Table 1a and 1b for the neck pain and asymptomatic groups, respectively.

We noted variations in inter-rater reliability between the two groups and the progressive consecutive stages and random stages protocols (Table 2a). In the neck pain group, the progressive consecutive stages protocol showed moderate reliability for the first measurement and good reliability for the second measurement. The inter-rater reliability was moderate for the random stages protocol and, similarly to progressive consecutive stages, there was a trend for increased reliability for the second measurement. The inter-rater reliability was higher for the asymptomatic group than for the neck pain group, with ICCs of 0.83–0.89.

The intra-rater reliability results revealed some similarity between the examiners (Table 2b). The intra-rater reliability was lowest for the random stages protocol (0.48–0.51), indicating poor reliability in the neck pain group. In contrast, the reliability for the protocol in progressive consecutive stages was moderate, with ICCs of 0.62–0.68 in the neck pain group. The intra-rater reliability was better in the asymptomatic group: the random stages protocol demonstrated moderate intra-rater reliability and progressive consecutive stages protocol showed good intra-rater reliability, with ICCs ranging from 0.72 to 0.80.

Among the individuals in the neck pain group, six participants reported experiencing dizziness at the end of the test. Nevertheless, they were able to complete the test successfully. None of the participants in the healthy group reported adverse effects. For the Fatigue VAS scores, the analysis revealed distinct patterns in the groups. In the asymptomatic group, the mean ± SD fatigue score was 0.34 ± 0.77, indicating mild fatigue [50]. In contrast, the mean ± SD fatigue score in the neck pain group was 3.06 ± 2.48, showing moderate fatigue.

The SUS results indicated good usability in the neck pain and asymptomatic groups, with a mean ± SD score of 8 ± 11.55 and 79.56 ± 13.36, respectively.
Table 1

a. Mean total percentage scores (% SCORE) and SDs of the CFMCT protocols in patients with neck pain by the two assessors.

<table>
<thead>
<tr>
<th>Patients with Neck Pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol in progressive consecutive stages</td>
</tr>
<tr>
<td>1st Measure</td>
</tr>
<tr>
<td>Examiner 1</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>72.83</td>
</tr>
<tr>
<td>Random stages protocol</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>73.35</td>
</tr>
</tbody>
</table>

b. Mean total percentage scores (% SCORE) and SDs (SD) of the CFMCT protocols in asymptomatic participants.

<table>
<thead>
<tr>
<th>Asymptomatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol in progressive consecutive stages</td>
</tr>
<tr>
<td>1st Measure</td>
</tr>
<tr>
<td>Examiner 1</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>75.92</td>
</tr>
<tr>
<td>Random stages protocol</td>
</tr>
<tr>
<td>1st Measure</td>
</tr>
<tr>
<td>Examiner 1</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>75.32</td>
</tr>
</tbody>
</table>

We noted variations in inter-rater reliability between the two groups and the progressive consecutive stages and random stages protocols (Table 2a). In the neck pain group, the progressive consecutive stages protocol showed moderate reliability for the first measurement and good reliability for the second measurement. The inter-rater reliability was moderate for the random stages protocol and, similarly to progressive consecutive stages, there was a trend for increased reliability for the second measurement.
The inter-rater reliability was higher for the asymptomatic group than for the neck pain group, with ICCs of 0.83–0.89.

The intra-rater reliability results revealed some similarity between the examiners (Table 2b). The intra-rater reliability was lowest for the random stages protocol (0.48–0.51), indicating poor reliability in the neck pain group. In contrast, the reliability for the protocol in progressive consecutive stages was moderate, with ICCs of 0.62–0.68 in the neck pain group. The intra-rater reliability was better in the asymptomatic group: the random stages protocol demonstrated moderate intra-rater reliability and progressive consecutive stages protocol showed good intra-rater reliability, with ICCs ranging from 0.72 to 0.80.
### Table 2

#### a. Inter-rater reliability and its corresponding Intraclass Correlation Coefficient (ICC) for subjects with neck pain and those asymptomatic during the CFMCT protocols and respective measurement timepoints

<table>
<thead>
<tr>
<th>Protocol in progressive consecutive stages</th>
<th>Random stages protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1st Measure</strong></td>
<td><strong>2nd Measure</strong></td>
</tr>
<tr>
<td>Neck Pain group</td>
<td></td>
</tr>
<tr>
<td>ICC (CI)</td>
<td>SEM</td>
</tr>
<tr>
<td>0.75</td>
<td>4.74</td>
</tr>
<tr>
<td>(0.52–0.87)</td>
<td>(0.65–0.9)</td>
</tr>
<tr>
<td>Asymptomatic</td>
<td></td>
</tr>
<tr>
<td>ICC (CI)</td>
<td>SEM</td>
</tr>
<tr>
<td>0.83</td>
<td>3.83</td>
</tr>
<tr>
<td>(0.67–0.91)</td>
<td>(0.71–0.92)</td>
</tr>
</tbody>
</table>

CI: Confidence Interval  
SEM: Standard Error of Measurement

#### Table 2b. Intra-rater reliability and its corresponding Intraclass Correlation Coefficient (ICC) for subjects with neck pain and those asymptomatic during the CFMCT protocols and respective measurement timepoints

**Intra-rater reliability**

<table>
<thead>
<tr>
<th>Protocol in progressive consecutive stages</th>
<th>Random stages protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Examiner 1</strong></td>
<td><strong>Examiner 2</strong></td>
</tr>
<tr>
<td>Neck Pain Group</td>
<td></td>
</tr>
<tr>
<td>ICC (CI)</td>
<td>SEM</td>
</tr>
<tr>
<td>0.68</td>
<td>5.68</td>
</tr>
<tr>
<td>(0.41–0.83)</td>
<td>(0.24–0.81)</td>
</tr>
<tr>
<td>Asymptomatic</td>
<td></td>
</tr>
<tr>
<td>ICC (CI)</td>
<td>SEM</td>
</tr>
</tbody>
</table>


Among the individuals in the neck pain group, six participants reported experiencing dizziness at the end of the test. Nevertheless, they were able to complete the test successfully. None of the participants in the healthy group reported adverse effects. For the Fatigue VAS scores, the analysis revealed distinct patterns in the groups. In the asymptomatic group, the mean ± SD fatigue score was 0.34 ± 0.77, indicating mild fatigue (50). In contrast, the mean ± SD fatigue score in the neck pain group was 3.06 ± 2.48, showing moderate fatigue.

The SUS results indicated good usability in the neck pain and asymptomatic groups, with a mean ± SD score of 82.00 ± 11.55 and 79.56 ± 13.36, respectively.

**Discussion**

Previous research has evaluated the use of inertial sensor technology to assess the CCF ROM during the performance of the original CCFT (19, 51) or during unconstrained full-range CCF (29). These computer-based tools equipped with real-time biofeedback systems have the potential to allow for testing or training the DCFs while providing improved accuracy and reduced subjectivity in identifying compensatory movements; moreover, the tools could be used in multiple positions (not just supine) (19). In the current study, we introduced a novel inertial sensor-based system with computer-based visual feedback designed to assess CCF movement control and to measure its reliability, safety and usability in asymptomatic subjects and patients with chronic neck pain. To our knowledge, there is no previous research that has developed a sensor-based system to test CCF movement control.

Another interesting aspect of the CFMCT is its potential use in terms of telerehabilitation, because inertial sensors and the software developed in this study could allow for evaluation or rehabilitation at a distance for patients with craniocervical pain. This could improve the accessibility to assessment and treatment for these patients, increase adherence levels and potentially reduce time and costs (30, 52, 53). As telehealth should be implemented to meet the specific needs of healthcare services (53), there is a lack of tools that enable clinicians to provide these services to manage craniocervical pain.

We showed that the progressive consecutive stages and random stages protocols are safe and easy to use. However, the random stages protocol exhibited lower intra-rater reliability, especially in subjects with neck pain, where the reliability values were poor. The lower reliability of the random stages protocol may be attributed to the variability in the performance of subjects between the first and second
measurements, as well as the inherently higher difficulty of the test in performing it consistently at both measurement times. Therefore, we suggest that the consecutive progressive stages protocol could become a useful test in clinical practice but perhaps the random protocol could be used as a treatment strategy to challenge the motor control system at the craniocervical level. Future research is necessary to investigate the potential use and effects of this tool as a therapeutic exercise.

In contrast to the random stages protocol, the progressive consecutive stages protocol demonstrated moderate to good reliability. Based on this finding, we have focused the following discussion mostly on the progressive consecutive stages protocol characteristics, limitations and recommendations for further research. The results for the progressive consecutive stages protocol showed greater agreement based on inter-rater ICCs (ranging from moderate to good) than on intra-rater ICCs (moderate). This may be caused by the fact that the performance-based CFMCT is very likely to show a learning effect (36). Therefore, most of the patients obtained a higher score on the second measurement, which may have reduced the intra-rater ICCs but did not necessarily influence the inter-rater results. Notably, reliability ICCs were higher in asymptomatic subjects (0.83–0.89) compared with patients with neck pain (0.67–0.82), suggesting that healthy individuals performed the test more consistently.

We hypothesise that modifying certain parameters such as the duration of progressive stages may influence the reliability of the test. The fact that the duration of each stage was as short as 3 seconds implies that participants might not reach a stage if they do not have good movement control or simply lose focus. However, because the test provides a total score, rather than a score for each stage, potential errors in one stage could be compensated for by better performance in other stages. Additional research could investigate whether the same test format as performed in the present study including stages with a longer duration would improve reliability scores.

The CCFT was developed to evaluate DCF performance. This construct was evidenced by EMG using nasopharyngeal electrodes to confirm that the five progressive stages of the CCFT reflected progressive activation of the DCF muscles (23, 33). Although we developed the CFMCT based on the fact that we have observed an agreement between pressure levels of the air-filled pressure sensor (indicative of DCF activation (23, 33) ) and CCF ROM stages measured by inertial sensors (19, 51), the CFMCT is focused on CCF movement control and not on DCF performance. Moreover, the first five stages of the CFMCT may be equivalent to the CCFT pressure stages, but the final four stages of the CFMCT are associated with additional CCF ROM targets that might not necessarily be associated with DCF contraction, but rather with additional activation of superficial flexors. It should be noted that we designed the CFMCT to minimise variation among the examiners regarding the subjective visual or tactile detection of superficial flexor activation, head retraction or progressive increments of the CCF ROM. Therefore, the CFMCT does not consider superficial flexor activation as something to avoid; rather, it assumes it may occur, especially in the latter stages of the test. The final score of the CFMCT is focused on measuring CCF movement control in a standardised manner guided by computer software and potentially minimising the influence of variation from the examiners. We hypothesise that CCF movement control is a relevant measure to assess impaired motion in patients with different craniocervical pain conditions, as well as a potential
friendly training exercise tool to improve the tolerance and adherence to exercise prescription. Future research should investigate its psychometric properties in its actual format or of any modified version, including its potential to discriminate between pain and asymptomatic populations. Moreover, its use as a low-load exercise for the treatment of craniocervical pain conditions may be of interest for clinicians and researchers in the field.

As a performance-based test, we designed the CFMCT so that the main challenge for patients is precisely to reach and maintain their CCF ROM within the coloured bars. As the range of each stage within the test is based on the participant’s own initially measured maximal CCF, it is assumed that the participant is capable of reaching all stages, but they might not have enough movement control to obtain high scores (measured as the number of time points within each stage-coloured bar). This idea is also different from the CCFT, because some people may not be capable of reaching some of the higher pressure levels of the test (21). In addition, the CCFT performance may be influenced by the elastic resistance of the pressure cuff that is compressed during the flattening of the cervical lordosis, while the CFMCT is performed without any cuff or resistance to the movement. Furthermore, while the CCFT generates a score based on five distinct stages, the CFMCT provides a continuous recording of the number of time points within each ROM stage. This provides a final continuous score that can be translated to a percentage of performance on a 100-point scale. We believe this novel evaluation system could enhance the precision and accuracy of sensorimotor control evaluation of the craniocervical spine and could make a significant contribution to the assessment and management modalities in the domain of healthcare technologies.

With regard to the safety of the CFMCT, a total of six participants in the neck pain group reported experiencing dizziness during the test. This finding may be associated with the maintained and highly demanding use of a computer screen that some patients may have perceived as conflicting among the sensory signals of self-movement in a digital environment (54). None of these participants had to stop the test or withdraw from the study, so it seems that these sensations were mild and not incompatible with performing the test. Future research could investigate this secondary effect and how much it could influence the performance of the test. None of the participants in the asymptomatic group reported secondary effects.

The mean fatigue VAS score was nearly three points higher for the neck pain group compared with the asymptomatic group. Although the CFMCT cannot be directly associated with DCF performance, we hypothesise that patients with chronic neck pain perceive the test as more tiring compared with healthy participants due to altered activation of DCFs and impaired movement control. Previous EMG studies have shown delayed and reduced DCF activation during a cervical flexion task (55). Future research could investigate the validity of the CFMCT to discriminate between neck pain and asymptomatic populations.

The SUS scores of the CFMCT reflected high levels of acceptance and satisfaction to use the system frequently in both groups of participants. Furthermore, the scores showed that the system did not appear unnecessarily complex: the participants found it easy to use and reliable. Additionally, the effective
integration of various functions within the system, along with the perception of minimal inconsistency, reinforces the notion of strong usability. The expectation that most people could learn to use the system quickly suggests that the tool has gentle learning curve, which is crucial for successful adaptation to the system. Therefore, this tool provides a positive user experience based on accessibility, ease of use, confidence, and contributing to a high level of usability.

**Limitations of the study**

The present study has several limitations. First, the findings are limited to the characteristics of the study population. Even though patients with neck pain had a minimum VAS score of 3 and an NDI score of 4, most of the patients exhibited moderate levels of pain and disability. Moreover, the sample may have not fully represented the diverse demographics and pain conditions observed in clinical practice. Non-European populations, specific pain conditions and higher levels of pain and disability should be included in future research. This will ensure the samples are more heterogeneous and help to generalise the results. Second, the variability between demographic characteristics among the participants based on age or educational level may have been associated with difficulties in understanding and correctly performing the CFMCT. Additionally, there were no familiarisation or training sessions. The lack of a familiarisation session might have affected the reliability of the measurements. Third, the assessment of movement was supervised by the examiners though observation without the objective monitoring via EMG. Therefore, there are no data regarding the motor strategies developed by participants during the CFMCT, particularly in patients with neck pain who may present different neuromuscular alterations. Future research may explore the EMG activity of the cervical muscles during in participants while they perform the CFMCT.

**Conclusions**

The CFMCT is a promising instrument to evaluate CCF movement control in patients with craniocervical pain: it reduces variation from the examiners and could serve as a telehealth instrument for assessment and exercise prescription. It is a safe and usable instrument both in patients with chronic neck pain and asymptomatic participants. Future research should investigate its psychometric properties for evaluation and its possible use for therapeutic exercise. In addition, it would be valuable to determine which muscular strategies are performed by subjects with pain during the test and how these strategies compare with asymptomatic subjects.

**Declarations**

**Ethics approval and consent to participate**

The study was approved by the Research Ethics Committee of CEU San Pablo University (236/17/08) and performed in accordance with the Declaration of Helsinki and World Health Organization (WHO) standards. All participants provided signed informed consent before participation.
Consent for publication

**Availability of data and materials:** The datasets generated during the current study are available from the corresponding author upon request.

**Competing interests:** None declare

**Funding:** The study was financially supported by the Agencia Estatal de Investigación, Spanish Ministry of Science, Innovation and Universities (Reference number: PID2019-108616RA-I00/AEI/10.13039/501100011033). The funding body had no role in the design of the study; collection, analysis, and interpretation of data; and in writing and submitting the manuscript.

**Authors' contributions:** AM, EBC, FM, TP, SAO and SL: conception and design of the study

EB and AM: writing/editing of final manuscript

EB and AM: prepared figure and supplementary video

EB, FM and AM: recruitment and testing;

SAO, TP, FM and SL: critical review and revision of manuscript;

All authors reviewed, read and approved the final manuscript.

**References**


2. Franov E, Straub M, Bauer CM, Ernst MJ. Head kinematics in patients with neck pain compared to asymptomatic controls: a systematic review. BMC Musculoskelet Disord. 2022;23(1).


7. Cagnie B, D’Hooge R, Achten E, Cambier D, Danneels L. A magnetic resonance imaging investigation into the function of the deep cervical flexors during the performance of craniocervical flexion. J


Figures
Figure 1

Figure legend not available with this version.