Article

Proposal of an Alternative to the AMA Guidelines for the Evaluation of the Cervical ROM

María Jesus Martínez-Beltrán [1,*] , **Alberto Fidalgo-Herrera** [2] , **José A. Moreno-Ruiz** [3] **and Julio C. de la Torre-Montero** [1]

[1] San Juan de Dios School of Nursing and Physical Therapy, Comillas Pontifical University, Avda. San Juan de Dios 1 Ciempozuelos, 28350 Madrid, Spain; juliodelatorre@comillas.edu

[2] Fisi(ON) Health Group, Biomechanics, Calle Pollensa, 2, Las Rozas, 28232 Madrid, Spain; afidalgo@fisi-on.com

[3] Departamento de Informática, Universidad of Almería, Carretera Sacramento, Calle San Urbano, s/n, La Cañada, 04120 Almería, Spain; jaruiz@ual.es

* Correspondence: mjesus.martinez@comillas.edu

check for updates

Abstract: The cervical spine is one of the most frequently injured joints in a car accident. References for the range of motion (ROM) that should be expected in a person are needed to stage the injuries. The two main objectives of this paper are to clinically validate a measuring device for the cervical spine, and to assess if the use of different ROM reference values will render different results from the American Medical Association (AMA) ROM guidelines. The present study is divided into 2 phases, a validation phase with 55 subjects and a case-control phase with 80 subjects. A BTS (Bioengineering Technology and System) system and the EBI-5 (estudio biomecánico integral) system were used for the present investigation. The intraclass correlation agreement value between both measuring devices is considered very good with a Cronbach alpha up to 0.9 in every dimension. Correlations (*r*) between variables are very high, not showing any values lower than 0.887. All comparisons between using AMA ROM guidelines or normative values presented significant differences ($p < 0.05$). The EBI-5 system has exhibited good accuracy being paired to a photogrammetric system. The use of guidelines adjusted to age constitute an alternative to the use of the AMA cervical ROM guidelines. Professionals should use age-normalized guidelines as an alternative to the AMA guidelines.

Keywords: cervical spine; AMA; normative data; range of motion; traffic injuries; accelerometry

1. Introduction

Traffic car accidents are responsible for an estimated 20 to 50 million injuries each year all over the world and while the survivability has increased, so has morbidity [1,2]. One of the most frequently diagnosed and studied injuries amongst traffic car injured is whiplash associated disorder (WAD) [3]. In western countries it is estimated that from 235 to 300 new whiplash cases arise for every 100,000 inhabitants per year [4]. This information can vary greatly between countries, with some like the USA with 328 cases every 100,000 inhabitants or others like Australia with 114 cases per 100,000 inhabitants [5–7]. In Spain, between the year 2002 and 2004, 12% of patients injured in traffic accidents were diagnosed with WAD [8], and countries like Japan 37.7% from a sample of 127,956 patients were diagnosed with WAD [9].

The differential diagnosis of WAD based on the symptoms of the patient represents a challenge for health professionals. In a study conducted in northern Sweden, 57% of patients diagnosed with WAD were wrongly diagnosed according to the ICD-10 (international classification of diseases) [6]. To ease this task, the Quebec Task Force (QTF) presented a scale of severity to grade people with

WAD. However, rating the impairment from whiplash injury is complicated when symptoms/signs do not exceed grade II of the QTF [10]. Moreover, the typical diagnosis is based solely on the complaints reported by the victims, combined with information indicating they have been involved in a collision. Complaints cannot be unambiguously verified even in a specialized physical or psychological examination [11].

Insurance fraud related to whiplash injuries is a widespread problem worldwide. Whiplash injury is easy to fake and difficult to disprove, leading to a high proportion of fraudulent claims. As there are usually no demonstrable pathoanatomical signs to be detected, WAD patients have a poor reputation [12]. This scheme poses another challenge for the clinical evaluators in order to differentiate the injured people from the uninjured ones.

Evaluators have classically relied upon visual inspection or inaccurate tools such as the goniometer, however these tend to reflect the subjective beliefs of the evaluator [13]. New trends among the professionals trained in biomechanics are surpassing this way of evaluation, promoting accurate tools such as photogrammetric systems or accelerometry [14–17]. Photogrammetric systems constitute a good source for this information, nevertheless their cost and high dependency on trained human resources make them unaffordable for widespread clinical use. There is a need for a cheaper evaluation system which can yield as much accuracy as a photogrammetric system. Inertial motion unit systems (IMUs) have demonstrated clinical validity in other joints in static and dynamic assessments of the range of motion [16]. However, the other systems rely only on the use of one recording probe and, therefore, the movement of other parts of the body artifacts can occur. The use of two reference probes could solve this problem [18,19].

Furthermore, despite the use of such tools, a good standard to compare the obtained results with is also needed.

The American Medical Association (AMA) guidelines of cervical range of motion (ROM) constitute one of the first intents to determine how much ROM should be expected in the cervical spine of an individual [20]. These guidelines have widespread use amongst clinical practitioners and legislators but are flawed in their understanding of human physiology. The AMA reference values just offer cut-offs for the minimal ROM to be expected for the cervical spine, and do not take into account variability present in the general population nor the loss of ROM due to aging [21]. With multiple studies reporting that age and the previous health status are main risk factors for the development of chronic WAD [22], it seems of great importance to adjust the references of movement to the variability present in the common population. These new references should include enough individuals in each age group to act as representative normative data of the population. Some authors have presented normative data from the cervical spine [21,23–25], however, it remains unknown if these normative data can be used across populations from different countries.

This article has three main purposes:

- To assess if the selected IMU system is valid for clinical practice.
- To assess if the use of different ROM reference values will render different results to the AMA ROM guidelines.
- To determine the cut-off where the ROM should be considered pathologically limited. Over 90% of healthy subjects should be considered as healthy.

2. Materials and Methods

2.1. Study Design

This study was divided into 2 phases. In the first phase, the precision of the measuring device was clinically validated, and in the second one, a case-control study was conducted to face the other 2 objectives.

A total of 55 subjects were initially measured for the validation phase of the study. To address how many more subjects would be needed, sample size calculations were performed with preliminary data.

The tool used to perform this calculus was the GRANMO calculator, which is free software available online [26]. The results of the calculation determined that only 32 subjects were needed to achieve reliable data with a confidence of the 95% and a power of 80%. According to this result, no more subjects were needed to be measured.

The validation group was skewed for the age of the participants, being most of them between the range of 18- to 30-year-olds. This outcome, although irrelevant for the validation, was important for the case-control phase of the study as age could potentially influence the global range of motion of a person's neck. It was decided to include in the control group only subjects between 18–30 years of age. This age group was selected to count on a sample that we would expect to show the highest amount of ROM. The AMA guidelines for neck ROM should consider most of the subjects as healthy. With the age constraint, only 40 subjects were eligible for the control group.

Data for the case group was gathered randomly from a database with more than 5000 trials of patients treated by Health Group Fisi(ON) in 90 measuring spots across Spain. All these patients were measured with the EBI-5 (estudio biomecánico integral) during normal care of their injuries. With the aim of having matching numbers between groups, 40 cases were extracted from the database.

The study had the approval of the ethics committee of Hospital Clínico San Carlos in Madrid, Spain. All principles of the Declaration of Helsinki were considered for the design and performance of this study.

2.2. Participants

Inclusion criterion for the validation phase was healthy people being between 18 and 60 years old. Exclusion criteria were: trauma surgical interventions; systemic diseases; traumatological, rheumatological or neurological diseases that had affected the cervical spine one year before; vestibular disorders; treatment with drugs that affect balance; mental alterations; fear to the procedures; and pregnancy.

The participants for the case-control phase were obtained as follows:

Data for the case group was gathered randomly from a database with more than 5000 trials of patients treated by Health Group Fisi(ON) in 90 measuring spots across Spain. The inclusion criteria for the case group were: age between 18–30 years old; injury due to a traffic car accident; no history of multiple traffic car accidents; diagnosis of cervical spine WAD or equivalent diagnosis.

The control group data was obtained from the participants of the first phase that were between 18–30 years old. Exclusion criteria were maintained to the previously explained factors.

The time range chosen for the sampling of the controls and the cases was between March and June 2018. As only 40 subjects were available for the control group, an equivalent number of cases was chosen.

2.3. Data Acquisition

2.3.1. Validation Phase

The EBI-5 system works with the information provided by 2 inertial motion units (IMUs). The IMUs used were the 3-SpaceTM Bluetooth sensors from Yost Labs® (Figure 1).

For cervical movement, the acquisition frequency was set to 10 Hz. The subjects had to repetitively reach the maximal ROM in each plane for 45 s without stopping. The movement performed by the subjects could never surpass 5 Hz in any case, so a 10 Hz frequency ensures that no loss of relevant information will happen. One of the sensors was placed upon the occiput with an elastic headband, and the other one between the spinous processes of T2 and T3 with a double-sided hypoallergenic sticker (Figure 2).
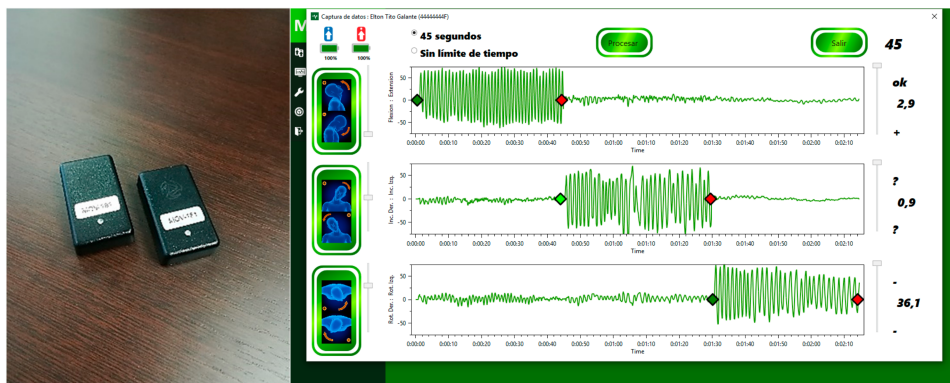
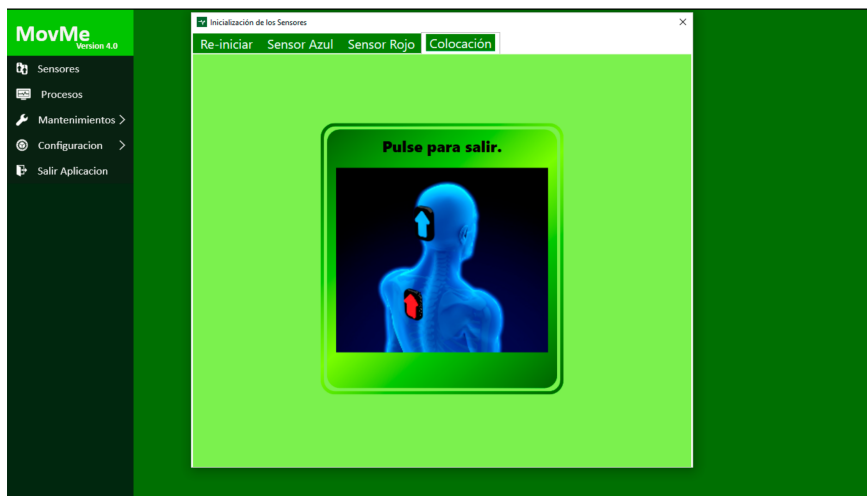**Figure 1.** Inertial motion units and EBI-5 (estudio biomecánico integral) system.

Figure 2. Diagram of the sensor placement in the subject.

All participants performed the maximum number of repetitions in a period of 45 s for each plane of motion: flexion/extension (FE), lateral bending (LB) and rotation (RT). All motions were performed in an upright sitting position at a self-determined velocity. The selection of 45 s of measuring time allowed for all patients to have time to perform enough repetitions of the movement to gain a clear picture of their actual mobility.

The photogrammetry system used for the comparison was a BTS SMART system (Bioengineering Technology and System—BTS, Milan, Italy) equipped with 6 infra-red cameras and the acquisition frequency set to 100 Hz. Reflective markers were placed upon the zygomatic processes, nasion, the suprasternal notch and in both acromion.

Researchers followed the prototypical STARD (standards for reporting diagnosis accuracy studies) as seen in Figure 3.

The variables used for the comparisons were the average range of motion of each pair of movements provided by each system in each trial.

Sanity of the data was checked before analysis and when the first trial performed by the patient was regarded as an outlier, the second trial would be chosen.
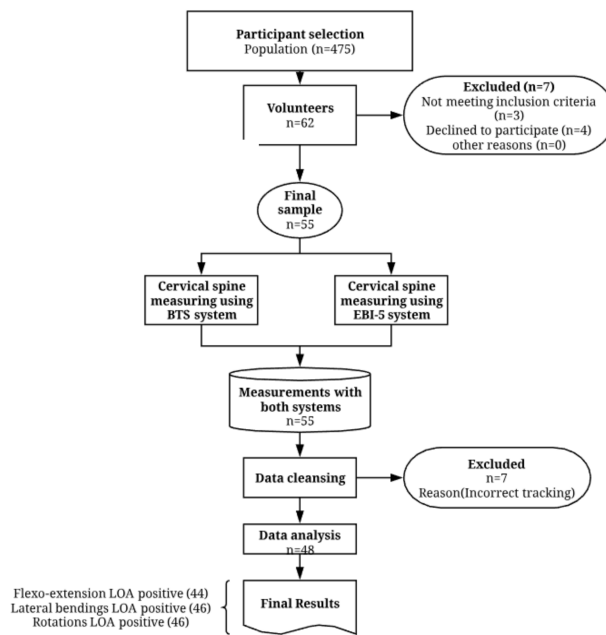
Figure 3. STARD (standards for reporting diagnosis accuracy studies) flowchart of the validation procedure.

2.3.2. Case-Control Phase

The protocol used during the validation phase was applied in this phase of the study, too. The primary outcome was the average of the maximum peaks from the ROM for each movement (flexo-extension, lateral bending movements and rotations). Secondary outcomes included: the normalized ROM in respect of three different ROM guidelines: AMA ROM guidelines, Swinkels ROM normative data [23] and our own normative data put together with our own healthy subjects. The following formula was used to calculate the normalized ROM (nROM) with the AMA guidelines for each movement:

$$nROM_{AMA} = \frac{ROM}{ROM_{AMA}} \times 100 \tag{1}$$

Since both Swinkels normative data [23] and our own data are not just cut-off values but an interval with the standard deviations, the normalized ROM was calculated with the high and low limits of the interval, and anything inside was considered a 100% of the expected ROM for that age group.

$$ROM > ROM_{Normative} + SD_{Normative} \rightarrow nROM = \frac{ROM}{ROM_{Normative} + SD_{Normative}} \times 100 \tag{2}$$

$$ROM < ROM_{Normative} - SD_{Normative} \rightarrow nROM = \frac{ROM}{ROM_{Normative} - SD_{Normative}} \times 100 \tag{3}$$

$$ROM_{Normative} - SD_{Normative} \leq ROM \leq ROM_{Normative} + SD_{Normative} \rightarrow nROM = 100\% \tag{4}$$

To assess how much limitation would be considered as pathological, two groups were created according to different thresholds in the control and the case groups. The cut-offs for this endeavor were 100%, 90%, 85%, 80% and 75% of the normalized ROM. If the subject was above the threshold, their ROM was considered as non-pathological and if below as pathologically limited.

The proportion of people considered as pathological was compared with the three different references.

2.4. Statistical Analysis

To compare the values from both systems an intra-class correlation analysis was conducted. A limits of agreement (LOA) analysis was also used to assess the agreement of both methods. Evaluating the correlation of variables is important, however, it is also as important for the evaluation of the degree of agreement between them to know the differences [27]. In addition, a Student's *t*-test for repeated measures was also performed with a significance level of 0.05.

Statistical analysis was conducted using IBM SPSS, Version 20.0. Descriptive statistics of all the variables were presented.

With the purpose of comparing the ROM of the cases and controls, a Student's *t*-test of independent samples was performed.

Since normal distribution was not met by each of the normalized ROM variables, contrasts of hypothesis were made using Friedman's test with a post hoc Wilcoxon's signed-rank test with a Bonferroni adjustment of $\alpha\_altered = 0.05/3 = 0.01666 \approx 0.017$, which did not exceed $\alpha\_critical = 1 - (1 - 0.017)^3 = 0.033$.

For the third objective, tables of the relative frequency of each group with every cut-off are presented.

All the statistical tests used in the current investigation are presented in Table 1.

Table 1. Statistical tests used.

Statistical Test	Phase Used	Objective
Intra-class Correlation		Correlation of measures between systems
Limits of Agreement	Validation phase	Agreement of measures between systems
Student's *t*-test for Repeated measures		Comparison of the measure between systems
Student's *t*-test for Independent Samples		Comparison of the ROM (range of motion) of the case and the control group
Friedman's Test	Case-control phase	Comparison of the Normalized ROM which each guideline at each group
Wilcoxon's Signed-Rank Test		Post hoc tests

No missing data had to be addressed in the study. The groups were matched by age; gender was not controlled because previous studies have demonstrated no statistical differences between men and women in cervical ROM [23].

3. Results

3.1. Validation Phase

A group of 62 eligible participants was selected out of the 475 potential candidates from a student population. Six of them were excluded for different reasons, including: not meeting inclusion criteria (*n* = 2), and decision to decline participation in the study (*n* = 4). From the final measured 55 subjects, 7 of them were withdrawn due to a lack of confidence in the values obtained during the experiment. The final analysis sample consisted of 48 subjects.

Demographic results showed a balanced population sample in anthropometric measures: male *n* = 21, 43.8%, age 23.7 years. (SD 4.5) height 176.3 cm (SD 6.2) and weight 77.2 (SD 8.2); female *n* = 27, 56.3% age 22.1 years. (SD 5.3) height 165 cm (SD 6.3) and weight 62.2 (SD 11.4). Table 2 shows the mean measurements for each system.

Table 2. Statistical results of the measures from the BTS (Bioengineering Technology and System) and EBI-5 systems.

Movement	BTS	EBI-5	BTS EBI-5 Difference	Mean Diff (95% CI)		Cronbach's Alpha	IC Correlation(r) *	p-Value
				Upper	Lower			
Flexion (F)	48.3 (8.3)	49.7 (8.4)	1.3 (2.5)	0.61	2.12	0.976	0.953	0.001
Extension (E)	49.6 (9.4)	49.3 (9.8)	−0.35 (2.7)	−1.16	0.45	0.979	0.961	0.383
Flexo-Extension Total (FE Total)	98.0 (13.0)	99.0 (13.5)	1.02 (3.44)	0.01	2.03	0.983	0.967	0.046
Right Bending (RB)	40.1 (5.8)	39.8 (6.1)	−0.2 (2.5)	−1.02	0.48	0.951	0.908	0.473
Left Bending (LB)	42.1 (7.5)	41.2 (7.3)	−0.8 (3.5)	−1.91	0.15	0.940	0.887	0.092
Bending Total(B total)	82.2 (12.1)	81.0 (11.9)	−1.1 (3.7)	−2.22	−0.03	0.975	0.951	0.044
Right Rotation (RR)	71.0 (6.1)	71.4 (6.1)	0.3 (2.4)	−0.3	1.1	0.960	0.923	0.260
Left Rotation (LR)	74.0 (6.5)	72.8 (7.1)	−1.2 (1.9)	−1.77	−0.64	0.979	0.963	<0.001
Rotation Total(R total)	145.0 (11.2)	144.2 (11.7)	−0.8 (2.6)	−1.63	−0.7	0.986	0.974	0.033

* $p < 0.005$.

The intraclass correlation (ICC) agreement [28] value was considered very good with a Cronbach alpha up to 0.9 in every dimension, and a strong relationship is observed between the targeted variables. Correlations (r) between variables were very high, not presenting any value lower than 0.887. Statistically significant differences (p values ≤ 0.005) were found for all measurements between the two systems.

Figures 4–6 show the Bland and Altman plots for the flexo-extension, lateral bending, and rotation. Most of the results fall within the range of the performed limits of agreement analysis. The smallest percentage of samples entering the LOA in any of the performed analyses is 90%. The additional Bland-Altman plots are presented in Appendix A.
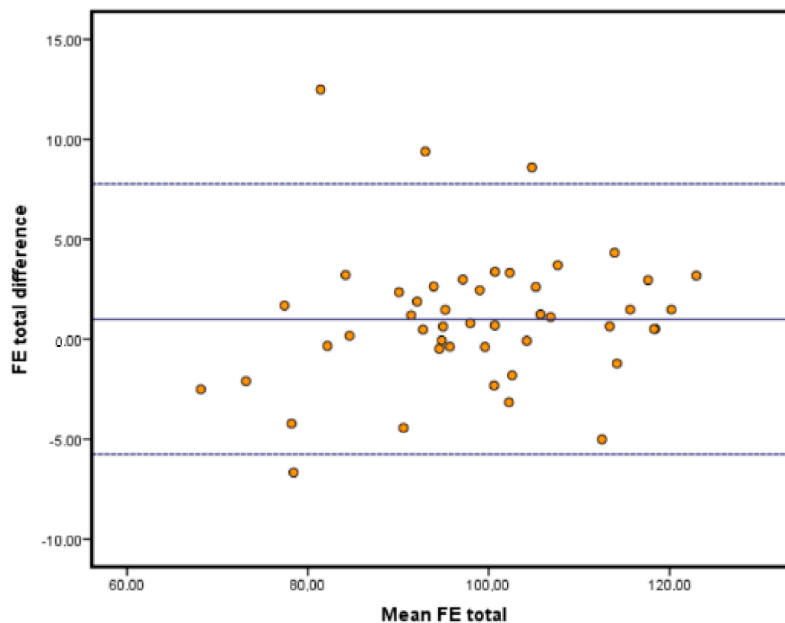
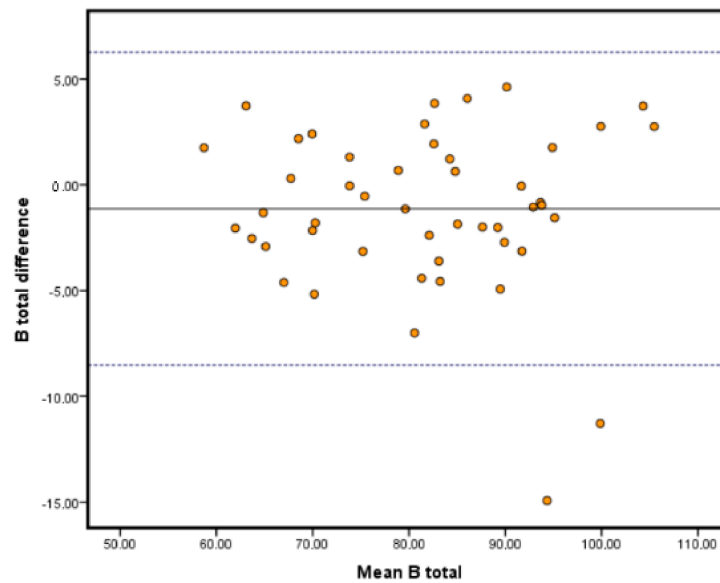**Figure 4.** Bland-Altman plot flexo-extension ROM.

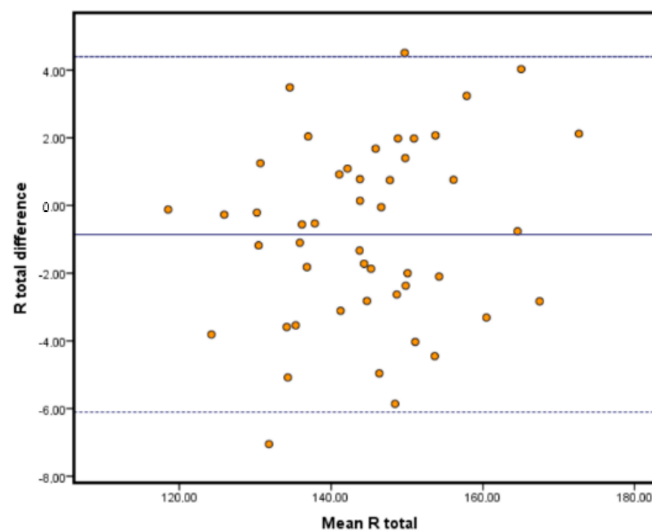**Figure 5.** Bland-Altman plot lateral bending ROM.

Figure 6. Bland-Altman plot rotations ROM.

A repeated measures Student's *t*-test was conducted for all the planes of motion. Statistically significant differences were found for the flexion, the left rotation, and the whole ROM of rotations. Although differences were found, the absolute difference between the values was not higher than 1.3°. Since in the context in which this machine was used clinically relevant differences are beyond 7°, the results of the variance analysis were discarded.

Mean differences between both systems were low, with equally low confidence intervals, which supports the good concordance of the outcome values from both systems.

3.2. Case-Control Phase

The control group (*N* = 40) presented a mean age of 21.55 years with an SD of 2.59 years. ROM values followed a normal distribution with 49.3° of extension (SD 13.38°), 55.4° of flexion (SD 11.27°), 42.10° of left side bending (SD 7.30°), 38.30° of right side bending (SD 7.63°), 71.40° of left rotation (SD 8.55°), and 73.70° of right rotation (SD 9.56°).

Normalized ROM with the AMA criteria followed a normal distribution; nevertheless, the normalized ROM with Swinkels and our own criteria were negatively skewed. Descriptive data are presented in Tables 3–5, respectively.

Table 3. Control group normalized values of ROM using the AMA guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	83.43	20.95	83.60	−0.04	90.6	65.82	99.67
Flexion	40	110.63	22.80	108.90	−0.17	87.3	92.65	131.35
Left Side Bending	40	94.14	15.98	96.45	−0.10	71.1	81.15	103.67
Right Side Bending	40	85.2	17.12	87.60	−0.18	69.7	70.02	96.95
Right Rotation	40	89.80	10.21	88.05	0.55	43.50	83.70	97.07
Left Rotation	40	92.72	11.56	92.70	0.321	54.90	84.57	101.45

Table 4. Control group normalized values of ROM using Swinkels guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	76.45	17.83	77.55	−0.39	62.20	61.10	92.45
Flexion	40	95.97	9.59	100.00	−2.32	42.10	94.45	100.00
Left Side Bending	40	96.44	8.03	100.00	−1.59	39.50	94.92	100.00
Right Side Bending	40	92.64	11.68	100.00	−1.59	43.10	83.97	100.00
Right Rotation	40	97.19	5.44	100.00	−1.40	25.30	95.60	100.00
Left Rotation	40	97.16	6.98	100.00	−0.79	39.20	94.80	100.00

Table 5. Control group normalized values of ROM using our own guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	99.38	8.58	100.00	−1.08	57.70	100.00	100.00
Flexion	40	98.94	8.01	100.00	−2.10	41.60	100.00	100.00
Left Side Bending	40	98.76	6.86	100.00	−0.43	43.00	100.00	100.00
Right Side Bending	40	99.07	7.67	100.00	−1.59	44.50	100.00	100.00
Right Rotation	40	100.50	4.00	100.00	1.529	24.00	100.00	100.00
Left Rotation	40	99.89	4.64	100.00	2.52	32.00	100.00	100.00

Friedman's test for the comparison of each of the normalized ROM in the control group were all significantly different ($p < 0.05$) except for the left side bending values ($p > 0.05$).

Wilcoxon tests were undertaken afterwards. All tests presented significant differences ($p < 0.017$).

The values normalized using Swinkles guidelines (SGN) were significantly higher than the values normalized with the AMA guidelines (AMAN) for all movements except for extension. The values normalized with our own guidelines (OGN) were significantly higher than AMAN in all movements except for the flexion, which was significantly lower than the flexion in AMAN. OGN were significantly higher than SGN for all movements.

Case group ($N = 40$) presented a mean age of 24.2 years with an SD of 0.44 years. The values for the ROM follow a normal distribution with 41.53° of extension (SD 4.35°), 40.43° of flexion (SD 4.92°), 34.99° of left side bend (SD 2.67°), 34.99° of right-side bend (SD 3.54°), 61.54° of left rotation (SD 3.07°), and 61.12° of right rotation (SD 3.91°).

Normalized ROM with the AMA criteria followed a normal distribution, nevertheless, SGN and OGN were negatively skewed. Descriptive data are presented in Tables 6–8 respectively.

Table 6. Case group normalized values of ROM using AMA guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	69.22	27.53	68.95	0.65	138.00	54.77	80.15
Flexion	40	80.86	31.12	81.90	−0.09	126.80	57.72	102.60
Left Side Bending	40	76.86	16.87	80.65	−0.86	73.50	65.70	90.30
Right Side Bending	40	77.77	22.37	78.65	−0.25	102.40	65.42	92.92
Right Rotation	40	76.93	19.38	79.10	−0.70	110.40	64.85	89.32
Left Rotation	40	76.40	24.75	81.80	−0.72	103.00	62.30	93.70

Table 7. Case group normalized values of ROM using Swinkels guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	56.09	21.91	56.75	−0.11	99.20	42.50	67.47
Flexion	40	69.15	26.43	74.30	−0.69	90.60	48.57	90.85
Left Side Bending	40	79.24	18.58	80.25	−1.11	83.40	68.10	97.20
Right Side Bending	40	79.98	20.97	87.10	−1.00	78.40	68.27	100.00
Right Rotation	40	77.31	20.37	81.40	−1.00	97.80	63.10	92.92
Left Rotation	40	75.89	24.72	85.40	−1.00	92.90	62.17	97.62

Table 8. Case group normalized values of ROM using our own guidelines.

	N	Mean	Standard Deviation	Median	Skewness	Range	Percentile 25	Percentile 75
Extension	40	87.01	24.64	100.00	−1.71	125.90	77.00	100.00
Flexion	40	74.87	27.21	82.60	−0.91	92.20	54.00	100.00
Left Side Bending	40	84.92	17.84	88.85	−1.68	81.60	75.35	100.00
Right Side Bending	40	89.34	19.25	100.00	−1.582	87.30	83.67	100.00
Right Rotation	40	83.77	20.38	90.70	−1.366	104.80	70.32	100.00
Left Rotation	40	81.17	24.62	95.00	−1.304	93.30	69.10	100.00

Friedman's test results for the comparison of each of the normalized ROM conditions in the case group were all significantly different ($p < 0.05$).

Wilcoxon tests were conducted afterwards. All the tests presented significant differences ($p < 0.017$) except OGN/AMAN for flexion, SGN/AMAN and OGN/AMAN for left side bend, SGN/AMAN for right side bend and SGN/AMAN for left rotation and right rotation. SGN was significantly higher than AMAN for extension and significantly lower than AMAN for flexion, respectively. OGN was significantly higher than AMAN for extension and significantly lower than AMAN for right side bend, left rotation, and right rotation. OGN were significantly higher than SGN for all movements.

Student's *t*-tests were statistically significant between the controls and the cases for all the ROM ($p < 0.05$).

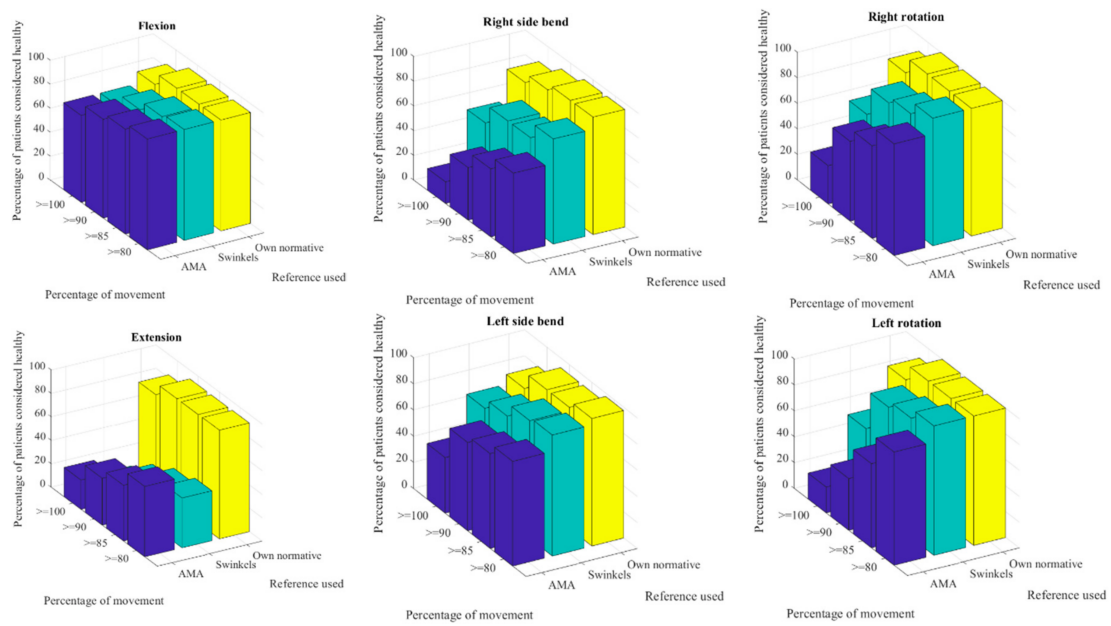The percentage of subjects included in the study depending on the fixed threshold are represented in Figures 7 and 8.

Figure 7. Percentage of controls included as non-pathological with each cut-off.
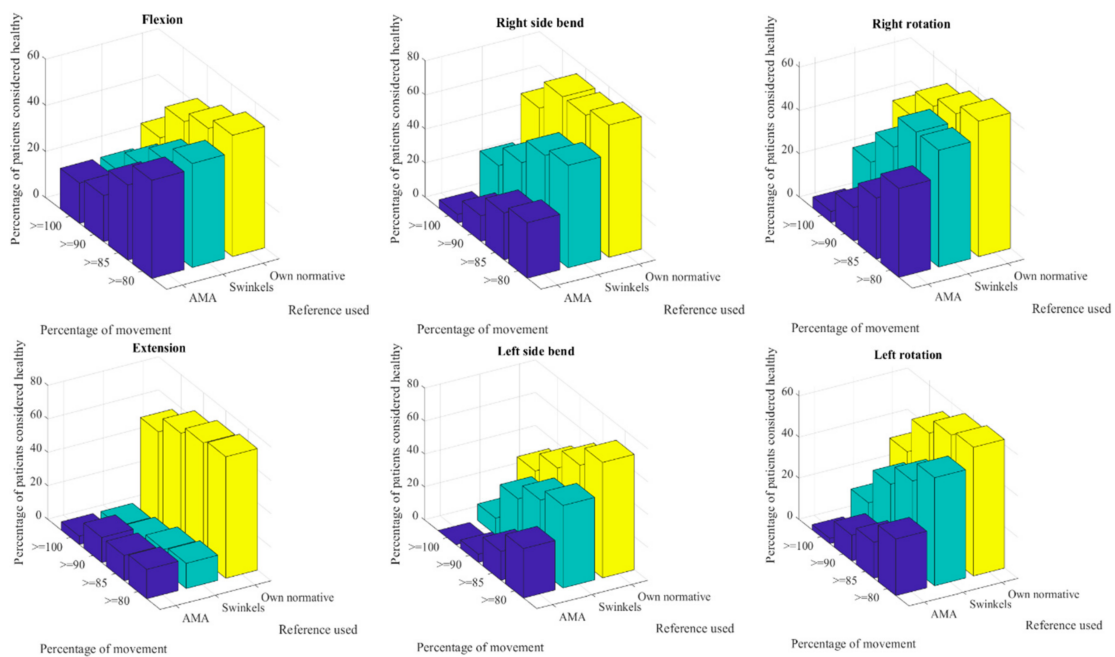

Figure 8. Percentage of cases included as non-pathological with each cut-off.

4. Discussion

4.1. Validation

The present study evaluates the performance of a new measuring device based on inertial motion units (IMUs). IMU technology has experienced a progressive improvement over the past few years. However, at the present time few IMU systems have evaluated the concordance of their measurements with previously established gold standards.

The correlation of the new system with an already established gold standard was necessary to ensure that the new device to be implemented would return similar values. ICC test scores showed that there was an excellent correlation between both systems, with values ranging from 0.887 to

0.974. These results were in accordance with some other systems in the field and even show better performance than others [29,30]. It is generally accepted that results falling above 0.7 of correlation are to be considered as good correlation, and with the lowest value almost close to 0.9, both systems can be regarded as excellently correlated. The correlation of IMUs with the cervical range of motion (CROM) device has been calculated also by other researchers. In this investigation, the results indicated that the pairing of the IMUs with the CROM was excellent [27]. According to their results and ours, the IMUs showed good accuracy in comparison with already established reliable methods of measurement.

However, correlation does not imply agreement between the measurements. To assess the degree of agreement, a *t*-test for related samples and a limits of agreement analysis were conducted. Statistically significant differences were found for all the comparisons between the EBI-5 system and the BTS system. Although the variance analysis returned significant *p* values, the maximum mean difference between the measures was 1.3° which in other studies has been considered as a good agreement [30].

The limits of agreement analyses are depicted in the Bland-Altman plots in Figures 2–4. This type of analysis shows how much agreement is to be expected from the comparison of both systems. Our results showed that the vast majority of our samples fell within the range of the limits of agreement, and those limits were similar to those reported in the literature, if not improving them [28]. It is important to remark that the highest limits were around 7º of discrepancy. The differences found can be explained in different ways. Some authors determine that the inner working of IMUs might return slightly different values than the calculation of the movement of reflective markers in the space [29]. Another hypothesis is that different configurations of markers and IMUs might result in differences in the results obtained. Furthermore, the difference in the sampling rates can affect the values obtained [31]. Nevertheless, mean differences between both systems show low values and, therefore, comparisons between systems can be made. Some authors argue that even with slight discrepancies between measuring devices, validated tools with low disagreement are acceptable for clinical and investigatory purposes [32].

4.2. Case-Control

The main problem when the quantification of the degree of injury is needed is to address the reference against which to compare the ROM of the patient. The AMA guidelines have long been used by health-related professions to answer this need. However, at the time they were conceived, precise measuring devices were not widespread and, thus, an approximation to the normal values a person should achieve was a fair comparison.

As our study has shown, this fair comparison becomes insufficient when precise ROM measurements are to be classified into pathologically limited or non-limited cases. It has been a surprising result that when the AMA guidelines were used over a healthy-subjects population (the control group) more than half of these would be under the expected ROM and, therefore, they should be categorized as pathologically limited. Only when the references were changed to normality intervals adjusted for age did the number of people considered to be healthy start to match reality. We would not expect any of the healthy subjects to present significant ROM limitations since the exclusion criteria would have prevented such an event. Other studies have investigated the effect of age on the ROM and all concluded that ROM consistently decreases with each decade of life [21,23]. Our results show that an unadjusted age reference (control group compared to AMA guideline for cervical ROM) leads to poor results and induces one to think that healthy people might have a significant limitation of ROM.

Previous research found statistical differences between using the AMA guidelines for cervical ROM or age-adjusted references for every decade of life. However, this study was not conducted on a healthy population hence it was not clear which guidelines were best to determine the degree of limitation [33]. The present results show that AMA guidelines for cervical ROM were unable to determine that a healthy population had no ROM impairment. These results also question the alleged sensitivity of the AMA cervical ROM guidelines to detect different degrees of injury across the decades of life.

The differences between Swinkels values and our own are around 3% except for the flexion-extension, where Swinkels values do not match our own and the difference was around 20%. While the device used in the present study compensates for the trunk motions, thus, eliminating extra range of motion from the final measure, the CROM device does not eliminate these combined movements. A recent systematic review analyzed the published normative values of cervical ROM obtained using different technologies. Their conclusions state that only the normative values obtained with the CROM device show consistency across studies. Consequently, they present the pooled normative values for the CROM device as the best normative values to use. However, their statement regarding the CROM guidelines as the only useful guidelines should be questioned. There is a high discrepancy in the extension movement between the studies using tools such as goniometers or the CROM device and the devices that use digital measuring devices. The CROM device cannot compensate for movement in associate joints, therefore, we might be evaluating the combined movement of multiple joints and not only the cervical range [34]. Digital devices usually account for these combined movements and remove them from the equation. The biomechanical model used for measuring can influence the results. Similar instrumentations or reference frames may correlate better than others that differ more than the former [27,35]. Another reason for discrepancies between normative data could be differences in the anthropometrics of the population. When the effect of anthropometrics in human motion is studied, performance is bound to subject characteristics (anthropometrics) and, therefore, should be accounted for [36]. The population from which we obtained the normative values were of university students and with a normal distribution of their ROM scores; there is no reason to think they will not follow a normal distribution in their anthropometrics for any population of Spanish students. These normative values will not be suitable for the clinical setting as they might not reflect the whole Spanish population and may not be applicable to other age decades. Nevertheless, future work will focus on obtaining these normative values for the whole population.

Once the data were normalized to the age intervals, the normality of the data was lost. One could argue that the loss of a normal distribution in the normalized values would be indicative of some bias not considered. However, the skewness resultant from this transformation suggests that that most of our results gather around the 100% mark and, thus, our healthy subjects are better classified. A similar trend is also observed in the case group with an increased skewness when normalized with the reference values that correspond to their age. This behavior matches the one that is observed in the studies of health care costs. Costs can never be lower than 0 in health and the most expensive treatments are also the least frequent, therefore, all the results are positively skewed [37]. It seems reasonable that the most frequent values are of less severity, while the worst outcomes are less frequent in the sample. In this case, the skewness reflects the nature of the clinical presentation of signs in pathology, with worse outcomes being less frequent that milder ones.

Using the normalized references with our own normative values did not result in all the healthy subjects receiving 100% of movement in all planes. If we were to consider the absence of pathology having 100% of the ROM, many healthy subjects would be incorrectly classified as injured. To include more than 90% of these subjects in a healthy consideration, the threshold of normalized movement had to be lowered to 90%. Only the right bending would have required a lower bound, however, we consider that lowering it too much will result in a greater probability of false negatives.

Lowering the bound on the case group also has an effect, as more people would be considered as non-limited in each movement. Lowering the bound does not result in the transformation of injured patients into healthy ones, it only reflects better which movements are truly limited. The AMA guidelines will classify almost all the patients as injured when in clinical practice some patients will not show any ROM limitation at all. Applying normative data adjusted by age indicates that some patients would not show ROM impairment at the assessment. Changing the way things are interpreted can allow for a better injury-oriented treatment in each patient, not wasting time treating unproblematic aspects and, therefore, improving the outcome. It should also be considered that this information ought to be crossed with the presence of other pathological symptoms such as pain. As our data show

in the control group, a person can exhibit a mobility lower than 90% of the normalized ROM and be completely asymptomatic. Having a limitation in ROM does not directly imply having an active injury. This consideration should be especially taken into consideration by clinicians, since finding a limitation does not always imply the necessity to treat it.

4.3. Limitations

Some drawbacks can be stated in the present study. The first limitation is that some degrees of difference must be accounted for when contrast is to be made with photogrammetric systems. The present devices are intended for clinical use in the rehabilitation environment. If measurements are to be taken in fields where precision must be higher, like surgery, this would not be the measuring device of choice. The second drawback is that the case-control phase focused only on the age interval of 18–30 years, it would be unwise to assume that this age group answers for all age intervals. However, we would expect to see similar results with other age intervals since what we are seeing here is the sheer amount of healthy people (control group) that would be considered injured if the AMA guidelines are used to obtain percentages of movement. However, it is reasonable that, as older individuals are evaluated, this effect would be less striking than the results obtained. Another limitation is that anthropometrics or gender were not considered, and although all distributions for ROM follow normality and gender does not affect ROM in the cervical spine, the assessment of these variables as confounding factors would favor even more the results presented herein.

5. Conclusions

In conclusion, the EBI-5 system has shown good accuracy being paired to a photogrammetric system. As a result, it should be considered validated for its accuracy. This outcome offers the possibility to use a cheaper and more easily portable device to measure ROM with objectively verifiable indicators. More accurate assessments will help direct the treatment and determine when to stop treating.

It can also be concluded that the use of guidelines adjusted to age constitutes a feasible alternative to the use of the AMA cervical spine ROM guidelines. It is patent that using ROM references adjusted to age to normalize the movement is more coherent both in our control group and case group.

Appendix A

Appendix A includes the additional Bland-Altman plots, evaluating each range of motion separately.
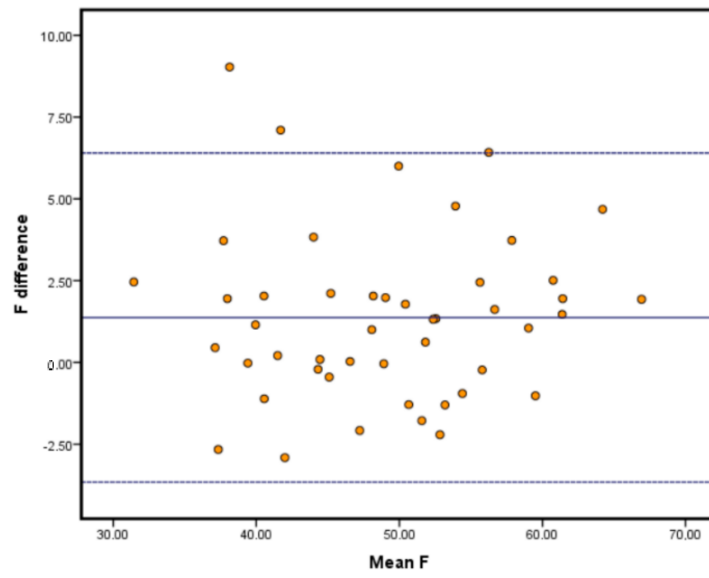
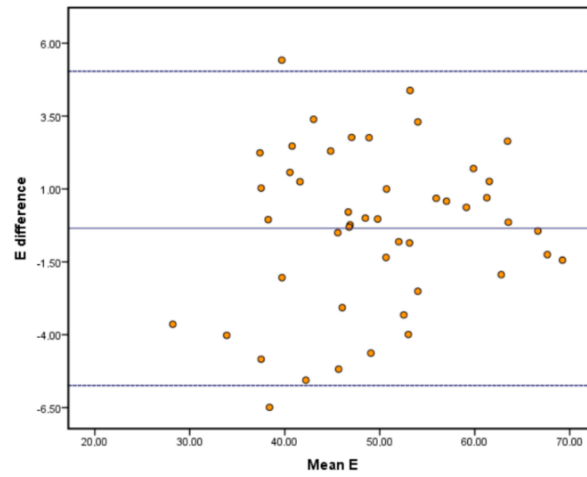**Figure A1.** Bland-Altman plot flexion ROM.
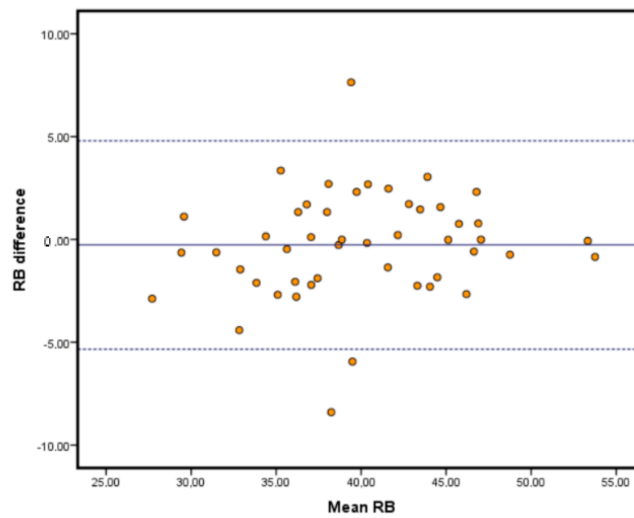
Figure A2. Bland-Altman plot extension.

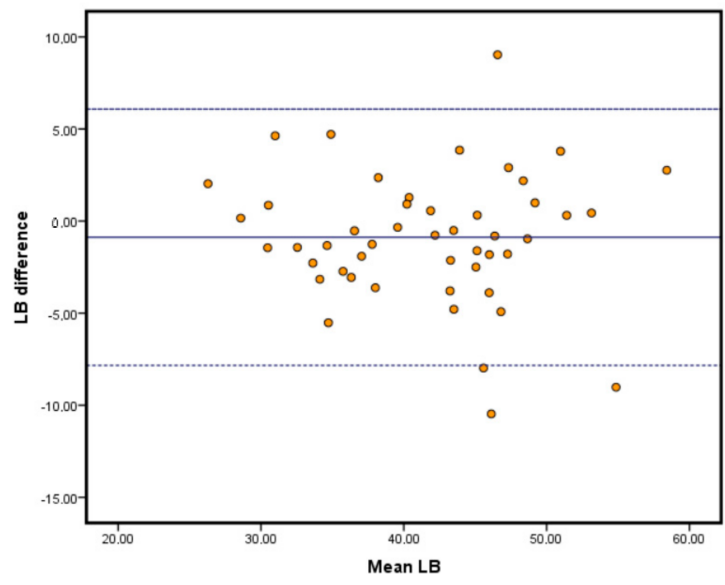**Figure A3.** Bland-Altman plot right bending.

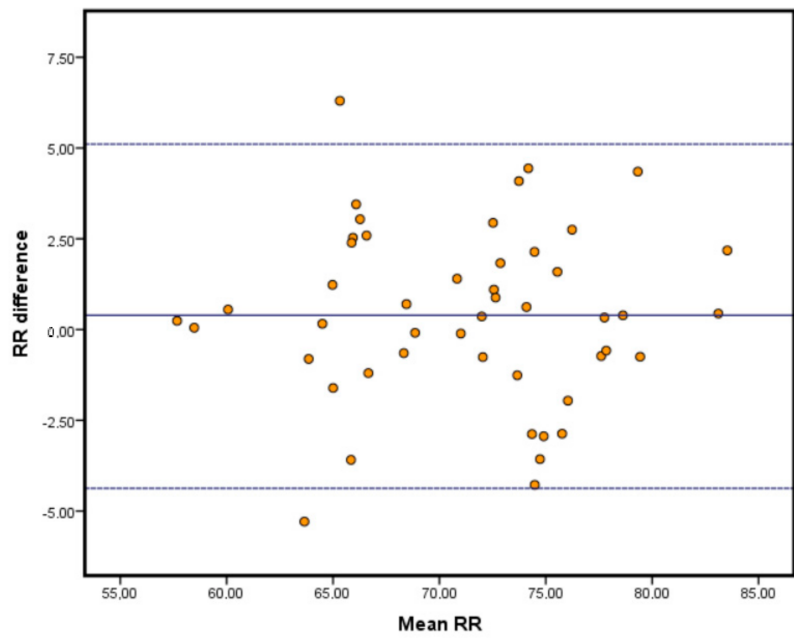**Figure A4.** Bland-Altman plot left bending.

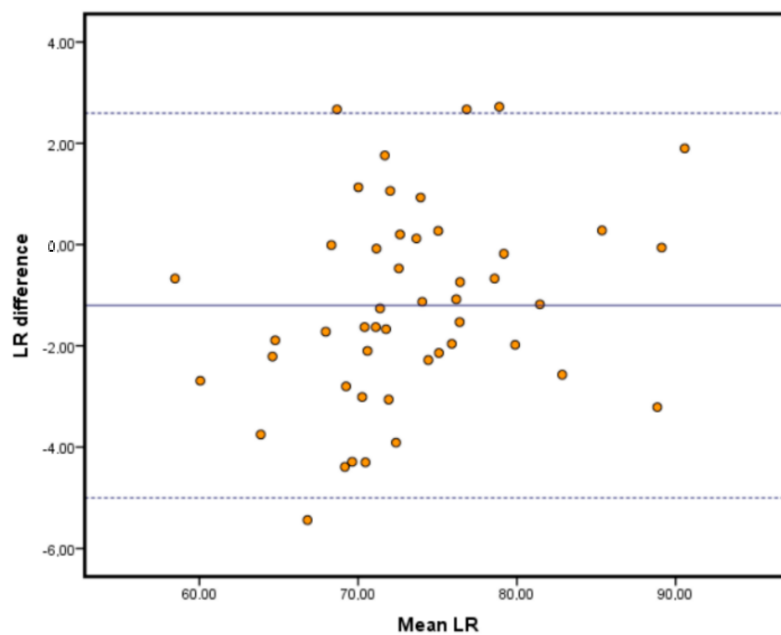**Figure A5.** Bland-Altman plot right rotation.

Figure A6. Bland-Altman plot left rotation.

References

1. Alemany, R.; Ayuso, M.; Guillén, M. Impact of road traffic injuries on disability rates and long-term care costs in Spain. *Accid. Anal. Prev.* **2013**, *60*, 95–102. [CrossRef]
2. Bortsov, A.V.; Platts-Mills, T.F.; Peak, D.A.; Jones, J.S.; Swor, R.A.; Domeier, R.M.; Lee, D.C.; Rathlev, N.K.; Hendry, P.L.; Fillingim, R.B.; et al. Effect of pain location and duration on life function in the year after motor vehicle collision. *Pain* **2014**, *155*, 1836–1845. [CrossRef]
3. Forman, J.L.; Lopez-Valdes, F.J.; Duprey, S.; Bose, D.; Dios, E.D.P.D.; Subit, D.; Gillispie, T.; Crandall, J.R.; Seguí-Gómez, M. The tolerance of the human body to automobile collision impact—A systematic review of injury biomechanics research, 1990–2009. *Accid. Anal. Prev.* **2015**, *80*, 7–17. [CrossRef]
4. Bussières, A.; Stewart, G.; Al-Zoubi, F.; Decina, P.; Descarreaux, M.; Hayden, J.; Hendrickson, B.; Hincapié, C.; Pagé, I.; Passmore, S.; et al. The Treatment of Neck Pain-Associated Disorders and Whiplash-Associated Disorders: A Clinical Practice Guideline. *J. Manip. Physiol. Ther.* **2016**, *39*, 523–564.e27. [CrossRef] [PubMed]
5. Garamendi, P.M.; Landa, M.I. *Epidemiología y Problemática médico Forense del síndrome de Latigazo Cervical en España. Cuadernos de Medicina Forense*; Scielo: Madrid, Spain, 2003; pp. 5–18.
6. Styrke, J.; Stålnacke, B.-M.; Bylund, P.-O.; Sojka, P.; Björnstig, U. A 10-Year Incidence of Acute Whiplash Injuries After Road Traffic Crashes in a Defined Population in Northern Sweden. *PM R* **2012**, *4*, 739–747. [CrossRef] [PubMed]
7. Berecki-Gisolf, J.; Collie, A.; McClure, R. Reduction in health service use for whiplash injury after motor vehicle accidents in 2000–2009: Results from a defined population. *J. Rehabil. Med.* **2013**, *45*, 1034–1041. [CrossRef] [PubMed]
8. Tournier, C.; Hours, M.; Charnay, P.; Chossegros, L.; Tardy, H. Five years after the accident, whiplash casualties still have poorer quality of life in the physical domain than other mildly injured casualties: Analysis of the ESPARR cohort. *BMC Public Health* **2016**, *16*, 13. [CrossRef] [PubMed]
9. Oka, H.; Matsudaira, K.; Fujii, T.; Tanaka, S.; Kitagawa, T. Epidemiology and psychological factors of whiplash associated disorders in Japanese population. *J. Phys. Ther. Sci.* **2017**, *29*, 1510–1513. [CrossRef]
10. Dvir, Z.; Gal-Eshel, N.; Shamir, B.; Pevzner, E.; Peretz, C.; Knoller, N. Simulated pain and cervical motion in patients with chronic disorders of the cervical spine. *Pain Res. Manag.* **2004**, *9*, 131–136. [CrossRef]
11. Teresiński, G. Verification of whiplash-associated disorders in forensic medical practice. part I–assessment of the injury circumstances and biomechanics. *Arch. Med. Sadowej Kryminol.* **2013**, *63*, 69–78.
12. Oddsdóttir, G.L.; Kristjansson, E.; Gislason, M.K. Sincerity of effort versus feigned movement control of the cervical spine in patients with whiplash-associated disorders and asymptomatic persons: A case-control study. *Physiother. Theory Pr.* **2015**, *31*, 403–409.

13. Ng, T.S.; Pedler, A.; Vicenzino, B.; Sterling, M. Physiotherapists' Beliefs About Whiplash-associated Disorder: A Comparison Between Singapore and Queensland, Australia. *Physiother. Res. Int.* **2015**, *20*, 77–86. [CrossRef] [PubMed]

14. Duc, C.; Salvia, P.; Lubansu, A.; Feipel, V.; Aminian, K. A wearable inertial system to assess the cervical spine mobility: Comparison with an optoelectronic-based motion capture evaluation. *Med. Eng. Phys.* **2014**, *36*, 49–56. [CrossRef] [PubMed]

15. Ernst, M.J.; Crawford, R.J.; Schelldorfer, S.; Rausch-Osthoff, A.-K.; Barbero, M.; Kool, J.; Bauer, C.M. Extension and flexion in the upper cervical spine in neck pain patients. *Man. Ther.* **2015**, *20*, 547–552. [CrossRef] [PubMed]

16. Støve, M.P.; Palsson, T.S.; Hirata, R. Smartphone-based accelerometry is a valid tool for measuring dynamic changes in knee extension range of motion. *Knee* **2018**, *25*, 66–72. [CrossRef]

17. Mourcou, Q.; Fleury, A.; Diot, B.; Franco, C.; Vuillerme, N. Mobile Phone-Based Joint Angle Measurement for Functional Assessment and Rehabilitation of Proprioception. *BioMed Res. Int.* **2015**, *2015*, 328142. [CrossRef]

18. Xu, X.; Chen, K.B.; Lin, J.-H.; Radwin, R.G. The accuracy of the Oculus Rift virtual reality head-mounted display during cervical spine mobility measurement. *J. Biomech.* **2015**, *48*, 721–724. [CrossRef]

19. Quek, J.; Brauer, S.G.; Treleaven, J.; Pua, Y.-H.; Mentiplay, B.; Clark, R. Validity and intra-rater reliability of an Android phone application to measure cervical range-of-motion. *J. Neuroeng. Rehabil.* **2014**, *11*, 65. [CrossRef]

20. Andersson-Gunnar, B.J.; Cocchiarella, L. *AMA Guides to the Evaluation of Permanent Impairment*, 5th ed.; American Medical Association: Chicago, IL, USA, 2000.

21. Niederer, D.; Vogt, L.; Wilke, J.; Rickert, M.; Banzer, W. Age-related cutoffs for cervical movement behaviour to distinguish chronic idiopathic neck pain patients from unimpaired subjects. *Eur. Spine J.* **2015**, *24*, 493–502. [CrossRef]

22. Biering-Sørensen, S.; Møller, A.; Stoltenberg, C.D.; Holm, J.W.; Skov, P.G. The return-to-work process of individuals sick-listed because of whiplash-associated disorder: A three-year follow-up study in a Danish cohort of long-term sickness absentees. *BMC Public Health* **2014**, *14*, 113. [CrossRef]

23. Swinkels, R.A.H.M.; Swinkels-Meewisse, I.E.J.C.M. Normal Values for Cervical Range of Motion. *Spine* **2014**, *39*, 362–367. [CrossRef] [PubMed]

24. Cagnie, B.; Cools, A.; De Loose, V.; Cambier, D.; Danneels, L. 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. Manip. Physiol. Ther.* **2007**, *30*, 450–455. [CrossRef] [PubMed]

25. Guo, L.-Y.; Lee, S.-Y.; Lin, C.-F.; Yang, C.-H.; Hou, Y.-Y.; Wu, W.-L.; Lin, H.-T. Three-dimensional characteristics of neck movements in subjects with mechanical neck disorder. *J. Back Musculoskelet. Rehabil.* **2012**, *25*, 47–53. [CrossRef] [PubMed]

26. Sample Size and Power Calculator GRANMO, Version 7.12, April 2012. Available online: https://www.imim.cat/ofertadeserveis/software-public/granmo/ (accessed on 22 February 2018).

27. Theobald, P.S.; Jones, M.D.; Williams, J.M. Do inertial sensors represent a viable method to reliably measure cervical spine range of motion? *Man. Ther.* **2012**, *17*, 92–96. [CrossRef]

28. Ockendon, M.; Gilbert, R. Validation of a Novel Smartphone Accelerometer-Based Knee Goniometer. *J. Knee Surg.* **2012**, *25*, 341–346. [CrossRef]

29. Chung, P.Y.M.; Ng, G.Y.F. Comparison between an accelerometer and a three-dimensional motion analysis system for the detection of movement. *Physiother. UK* **2012**, *98*, 256–259. [CrossRef]

30. Saggio, G.; Quitadamo, L.R.; Albero, L. Development and evaluation of a novel low-cost sensor-based knee flexion angle measurement system. *Knee* **2014**, *21*, 896–901. [CrossRef]

31. Pfister, A.; West, A.M.; Bronner, S.; Noah, J.A. Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *J. Med. Eng. Technol.* **2014**, *38*, 274–280. [CrossRef]

32. Milani, P.; Coccetta, C.A.; Rabini, A.; Sciarra, T.; Massazza, G.; Ferriero, G. Mobile smartphone applications for body position measurement in rehabilitation: A review of goniometric tools. *PM&R* **2014**, *6*, 1038–1043.

33. Fidalgo-Herrera, A.J.; Beltrán, M.J.M.; García-González, A.; De La Torre-Montero, J.C. P 124-Normative data to assess the degree of normality in the cervical spine movement replacing the outdated AMA guidelines. Preliminary results. *Gait Posture* **2018**, *65*, 443–444. [CrossRef]

34. Graaf, M.T.-D.; Thoomes, E.; Fernández-De-Las-Peñas, C.; Plaza-Manzano, G.; Cleland, J.A. Normative values of cervical range of motion for both children and adults: A systematic review. *Musculoskelet. Sci. Pr.* **2020**, 102182. [CrossRef] [PubMed]

35. Raya, R.; Garcia-Carmona, R.; Sanchez, C.; Urendes, E.; Ramirez, O.; Martin, A.; Otero, A. An inexpensive and easy to use cervical range of motion measurement solution using inertial sensors. *Sensors* **2018**, *18*, 2582. [CrossRef] [PubMed]

36. Demers, E.; Pendenza, J.; Radevich, V.; Preuss, R. The Effect of Stance Width and Anthropometrics on Joint Range of Motion in the Lower Extremities during a Back Squat. *Int. J. Exerc. Sci.* **2018**, *11*, 764–775. [PubMed]

37. Malehi, A.S.; Pourmotahari, F.; Angali, K.A. Statistical models for the analysis of skewed healthcare cost data: A simulation study. *Health Econ. Rev.* **2015**, *5*, 11. [CrossRef]