Heliyon 12 (2026) e45315

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# A comparison of the portable *DIERS Postura* prototype and the stationary *DIERS Formetric* system for radiation-free spinal reconstruction
## A reliability and agreement assessment

Viktoria Kiaulehn [a,*], Jörg Eschweiler [b,c], Roman Michalik [d], Miguel Pishnamaz [d], Frank Hildebrand [d], Hannah Lena Siebers [a]

[a] *Teaching and Research Area Experimental Orthopaedic and Trauma Surgery, Uniklinik RWTH Aachen, Pauwelsstraße 30, Aachen, 52074, North Rhine-Westphalia, Germany*
[b] *Department of Trauma and Reconstructive Surgery, BG Klinikum Bergmannstrost, Merseburger Str. 165, Halle (Saale), 06112, Saxony-Anhalt, Germany*
[c] *Department of Trauma and Reconstructive Surgery, University Hospital Halle, Ernst-Grube-Str. 40, Halle (Saale), 06112, Saxony-Anhalt, Germany*
[d] *Department of Orthopaedics, Trauma and Reconstructive Surgery, Uniklinik RWTH Aachen, Pauwelsstraße 30, Aachen, 52074, North Rhine-Westphalia, Germany*

ARTICLE INFO

*Keywords:*
*DIERS Formetric*
*DIERS Postura*
Azure kinect
Non-invasive spine diagnostic

ABSTRACT

*Purpose:* The *DIERS Postura* reconstructs spinal curvature from back surface topography using an algorithm derived from the established *DIERS Formetric*. By replacing rasterstereography with an Azure Kinect depth camera, it offers a cost-efficient, portable alternative. This study assesses how this change affects *Postura's* reliability and its agreement with *Formetric* to identify potential systematic errors.

*Methods:* Measurements were taken consecutively with the *Postura* and *Formetric* while the subjects maintained the same posture. Thirty-one subjects without spinal deformity were evaluated in habitual standing, with four sets of measurements per subject. Three landmarks were marked using reflective markers. Eleven spinal parameters calculated as outputs by both systems were analyzed using Intraclass Correlation Coefficient (ICC), Standard Error of Measurement (SEM), and Smallest Detectable Change (SDC). System agreement was assessed using the Bland-Altman method.

*Results:* The non-commercial *Postura prototype* (2021) showed moderate reliability for two parameters, good reliability for eight, and excellent reliability for one. SDC ranged from 0.52° to 5.11° for measured angles and from 2.06 mm to 7.46 mm for lengths. The agreement analysis showed for eight parameters a mean difference of less than 1°/1 mm, with Limits of Agreement (LoA) up to 5.6° and 8.3 mm. The kyphotic angle showed highest discrepancies, with an SDC of 5.11°, mean difference of −8.88°, and LoA of 10.70°.

*Conclusion:* Systematic errors were found in the determination of the spinous process line (SPL) and kyphotic angle. They were caused by the incompatibility of reflective markers with Postura's

* Corresponding author.
*E-mail addresses:* vkiaulehn@ukaachen.de (V. Kiaulehn), Joerg.Eschweiler@bergmannstrost.de (J. Eschweiler), rmichalik@ukaachen.de (R. Michalik), mpishnamaz@ukaachen.de (M. Pishnamaz), fhildebrand@ukaachen.de (F. Hildebrand), hannah.siebers@rwth-aachen.de (H.L. Siebers).

https://doi.org/10.1016/j.heliyon.2026.e45315
Received 10 September 2024; Received in revised form 7 July 2025; Accepted 12 August 2026
Available online 2 September 2026
2405-8440/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

depth sensor. However, for parameters not directly relying on the SPL, *Postura* achieved *Formetric's reliability* classification for four parameters, one classification lower for others, and good inter-system agreement for eight parameters. This study represents the only evaluation of a Postura system to date, providing insight into its accuracy and limitations, and a framework for the further validation of the released commercial version.

## 1. Introduction

The *DIERS Formetric* system (DIERS International GmbH, Schlangenbad, Germany) is an established but stationary device for postural analysis that uses surface measurements of the patient's back to reconstruct the 3D curvature of the spine. Building on this radiation-free technology, the *DIERS Postura* system (DIERS International GmbH, Schlangenbad, Germany) introduces a more compact design, further reducing costs, and stands out due to its portability. This makes it particularly beneficial for smaller clinics with limited space and budget. For posture analysis, both the *Formetric* and *Postura* system begin with the acquisition of the surface topography of the patient's back. Once the surface topography has been determined, both systems use this data as input to calculate a 3D reconstruction of the spinal curvature and ultimately generate a set of output parameters, such as the kyphotic and lordotic angles, which provide the user with feedback on the measured posture. This reconstruction algorithm was originally developed by Drerup and Hierholzer in the 1980s, for surfaces obtained through rasterstereography [1] - a technique in which parallel beams of light are projected onto the patient's back, and a camera system captures the resulting distortion to map the surface. Implemented in the *Formetric* system, the developed reconstruction algorithm, in combination with rasterstereography, has since been evaluated for repeatability and accuracy. In 2020, Degenhardt et al. assessed intra/inter-rater and intra/inter-day reliability [2].

The compact design of the *Postura* system was achieved by employing an alternative surface measurement technique. Rather than rasterstereography, an Azure Kinect (Microsoft Corporation, Redmond, Washington, USA) with a 1-megapixel time-of-flight depth camera is utilized for the *Postura* system. This depth camera projects a grid of infrared light onto the patient's back and measures the light's round-trip travel time to determine the surface topography. While this approach enables greater portability and reduced cost, this implementation compromises on resolution compared to rasterstereography. While *Formetric* attains $\pm$0.5 mm surface, $\pm$0.1 mm depth, and $\pm$0.8 mm fixation point precision [3], the depth camera of the Azure Kinect reports a standard deviation of $\leq$ 17 mm and a distance error of $\leq$ 11 mm (up to 3.5m distance) [4].

This study examines potential discrepancies arising from the transition in surface mapping technology, which may affect the reconstruction algorithm and measurement procedure, both previously validated with rasterstereography in the *Formetric* system. Since the *Postura* system's 3D spine reconstruction algorithm is based on the *Formetric* algorithm, both systems generate identical parameters as output. Therefore, when measuring the same patient in the same posture, the *Postura* system's results should closely align with the *Formetric* system. To assess this, the non-commercial *Postura* prototype from 2021, which uses an Azure Kinect depth sensor, was compared to the established commercial *Formetric* system that employs rasterstereography.

## 2. Material and methods

### 2.1. Test setup

To investigate inter-system agreement and to identify systemic errors resulting from the switch from rasterstereography to a depth sensor, we conducted our study in a cohort without spinal deformity. The eligibility criteria are given in Appendix C. Each subject was

**Table 1**
Description of the DIERS parameter investigated [2]. Parameters were given as output by both systems. VP: vertebra prominence, DM: midpoint between the dimples (DL, DR), ICT: inflection point cranial-thoracic, ITL: inflection point thoracic-lumbar, ILS: inflection point lumbar-sacral, CA: cervical point, KA: kyphosis apex, LA: lordosis apex.

| parameter | description |
|---|---|
| **angular [ °]** | |
| sagittal imbalance VP-DM | the angle between the connecting line VP-DM and an external line of gravity in the sagittal plane (+: subject leans forward) |
| coronal imbalance VP-DM | the angle between the connecting line VP-DM and an external line of gravity in the frontal plane (+: subject leans to the right) |
| kyphotic angle ICT-ITL | the angle between the surface tangents of ICT and ITL |
| lordotic angle ITL-ILS | the angle between the surface tangents of ITL and ILS |
| vertebra rotation (+max) | the maximum value of the horizontal components of the surface normals on the symmetry line to the right |
| vertebra rotation (-max) | the maximum value of the horizontal components of the surface normals on the symmetry line to the left |
| **length [mm]** | |
| pelvic obliquity | the vertical spatial distance between the dimples. (+: DR is higher than DL) |
| flèche lombaire | the horizontal spatial distance between the LA and the perpendicular through the KA |
| flèche cervicale VP | the horizontal spatial distance between the CP and the perpendicular through KA |
| lateral deviation VP-DM (+max) | the maximum deviation of the midline of the spine from the direct connection VP-DM in the frontal plane to the right |
| lateral deviation VP-DM (max) | the maximum deviation of the midline of the spine from the direct connection VP-DM in the frontal plane to the left. |

measured consecutively with both systems and eleven parameters provided by each system as feedback for characterizing the measured posture were assessed. These parameters, detailed in Table 1, were then analyzed to assess the repeatability of each system and compared to assess the level of agreement between the two systems.

The choice of a cohort without deformity was based on the assumption that systematic errors would be present in both healthy and pathological groups. However, because systems were evaluated by comparing the variance between measurements to the variance within the group, the use of a cohort with lower within-group variance allowed for easier identification and root cause analysis of these errors.

For examination the systems guidelines were followed. The subject adopted a habitual standing posture: arms hanging sideways, shoulders neutral, and head facing forward. The subjects were undressed except for their trousers, which were lowered so that the rima ani was visible. Reflective adhesive dots were used to mark three anatomical landmarks: vertebral prominence (VP), left lumbar dimple (DL) and right lumbar dimple (DR). The landmarks DL and DR were used to calculate the midpoint between the dimples (DM). The systems were set up so that the subject could be measured at the same location, without the need to move between measurements. Each measurement set took about 30 s and consisted of two consecutive measurements: a *Postura* measurement followed by a *Formetric* measurement. During a measurement set, the subject was instructed to remain as still as possible to minimize postural variation. Each subject completed four of the described measurement sets, along with one backup set. All sets were conducted by the same examiner on the same day. The *Formetric* measurements were performed with DiCAM 3.17 in 4D averaging mode and the *Postura* measurements with DiCAM v3.17.0.27. Following the system's guidelines the lighting conditions were adjusted between measurements, darkened for *Formetric* and not darkened for *Postura*.

All participants provided written informed consent for participation and data use. The study was approved by the local ethics committee (EK 377/21) and registered in the German Clinical Trial Register under the number DRKS00026675. Recruitment, measurement, and data collection were conducted from December 2021 to July 2022 at Uniklinik RWTH Aachen, Aachen, Germany. The study was conducted by the STROBE guidelines (Appendix D).

### 2.2. Repeatability of each system

The repeatability of each system was estimated with the Intraclass Correlation Coefficient (ICC 2.1) using the software SPSS v28. The ICC was classified as poor, moderate, good, or excellent according to Koo et al. [5] (Table 2). The assumption of homogeneity of variance was tested with Levene's test, calculated with SPSS v28, and was accepted with a level of significance of $\alpha = 0.05$.

For each parameter (Table 1) the variation $\overline{\sigma}^2_{scan}$ (Eq. (1)) was estimated as the mean squared error (*MSE*) of the within-subject variance. This estimation was carried out using a one-way ANOVA, based on the $m = 4$ repeated measures $x_j$ and the corresponding mean value $\overline{x}$ per patient among the $n = 31$ subjects.

$$\overline{\sigma}^2_{scan} = \frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(x_j - \overline{x}\right)^2 \tag{1}$$

To evaluate each system's capability to detect a true change and distinguish it from measurement error, the standard error of measurement (SEM) (Eq. (2)) and the smallest detectable change (SDC) (Eq. (3)) were calculated for both systems. The SEM quantifies the expected variability in a measured score due to measurement imprecision, while the SDC defines the minimum change required to confidently determine that the observed difference reflects a true change rather than a measurement error. Python v3.9 was used to perform the ANOVA and to calculate the *SDC* and *SEM*.

$$SEM_{scan} = \sqrt{\frac{\overline{\sigma}^2_{scan}}{m}} \tag{2}$$

$$SDC_{scan} = 1.96 \cdot \sqrt{2} \cdot SEM_{scan} \tag{3}$$

In a clinical context, the *SDC* defines the threshold beyond which true clinical changes can be distinguished from measurement errors. This is particularly relevant for the Postura system, given its potential application in follow-up assessments. Only changes that exceed the system's *SDC* can be considered real and used for diagnosis and treatment decisions. Numerous studies have evaluated the *SDC* for biplanar X-ray measurements of the Cobb angle, showing values of $SDC < 5°$ for angle parameters [6–8] and $SDC < 10$ mm for length parameters [9–11]. Given the status of biplanar radiography, these values were used as a reference to evaluate the ability of *Formetric* and *Postura* to detect real changes.

**Table 2**
ICC classification based on Koo and Li [5].

| ICC range | classification |
|---|---|
| $0.00 \leq ICC < 0.50$ | poor |
| $0.50 \leq ICC < 0.75$ | moderate |
| $0.75 \leq ICC < 0.90$ | good |
| $0.90 \leq ICC < 1.00$ | excellent |

### 2.3. Agreement between systems

The agreement of the *Postura* (denoted with $_P$) with the established *Formetric* (denoted with $_F$) was determined with the Bland-Altman method for repeated measurements [12,13]. The difference between a pair of measurements was denoted for subject $i$ and the repeated measurements $j$ as $D_{ij}$ (Eq. (4)) with $n = 31$ and $m = 4$.

$$D_{ij} = Y_{iP_j} - Y_{iF_j} \tag{4}$$

with $i = 1, \ldots, n$ and $j = 1, \ldots, m$

As proposed by Bland and Altman, the variance $\sigma_D^2$ (Eq. (5)) of the differences was estimated via a one-way ANOVA, which was performed with Python v3.9, resulting in an estimation of the mean squared error within a subject ($MSW_D$) and the mean square error between the subjects ($MSB_D$). The overall mean difference $\mu_D$ was calculated by averaging over the indexes i and j (Eq. (6)). Accordingly, the 95% lower and upper Limits of Agreement ($LoA$) were defined (Eq. (7)).

$$\widehat{\sigma}_D^2 = MSW_D + \frac{MSB_D - MSW_D}{\frac{\left(\sum_{i=1}^{n} m_i\right)^2 - \sum_{i=1}^{n} m_i^2}{(n-1)\sum_{i=1}^{n} m_i}} \tag{5}$$

$$\mu_D = \frac{1}{n \cdot m} \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij} \tag{6}$$

$$LoA = \mu_D \pm 1.96 \cdot \sigma_D \tag{7}$$

## 3. Results

### 3.1. Collective description

A total of 33 subjects were recruited, of which 31 subjects (16 male, 15 female) were included in the study based on the criteria (Appendix C). The cohort's mean height was 175.03 cm (SD = 9.57, range = 158 to 193 cm), the mean weight was 70.90 kg (SD = 12.60, range = 52 to 95 kg), and the mean age was 25.68 years (SD = 7.48, range = 18 to 58 years). The overall mean and standard deviation of the collective determined by the *Postura* and *Formetric* system are given in Appendix A. For comparison, the results of Degenhardt et al. [2], who studied a comparable collective with the *Formetric* system, are also given.

### 3.2. Repeatability of systems

The *Postura* system demonstrated good repeatability for eight out of eleven parameters (Table 3). Excellent repeatability was observed for lordotic angle ITL-ILS, while lateral deviation VP-DM (-max) and vertebral rotation (+max) showed moderate repeatability. The repeatability of the *Formetric* system was rated good to excellent, with excellent results for kyphotic angle ICT-ITL, lordotic angle ITL-ILS, vertebral rotation (-max), as well as pelvic obliquity, flèche lombaire and flèche cervicale VP.

**Table 3**
ICC, classification and CI for the repeated measurements with the DIERS Postura and DIERS Formetric and the comparative study by Degenhardt et al. ([2], ICC-R in Table 2]). Results are classified according to Koo and Li [5] as: 0.00-0.5 poor, 0.50-0.75 moderate, 0.75-0.90 good, 0.9-1.0 excellent.

| parameter | Postura | Formetric | Degenhardt |
|---|---|---|---|
| **angular [ ° ]** | | | |
| sagittal imbalance VP-DM | 0.877/good [0.801; 0.932] | 0.891/good [0.823; 0.940] | 0.92/excellent [0.85–0.95] |
| coronal imbalance VP-DM | 0.875/good [0.798; 0.931] | 0.897/good [0.831; 0.943] | 0.67/moderate [0.50–0.79] |
| kyphotic angle ICT-ITL | 0.795/good [0.681; 0.884] | 0.929/excellent [0.883; 0.962] | 0.96/excellent [0.92–0.97] |
| lordotic angle ITL-ILS | 0.945/excellent [0.908; 0.970] | 0.957/excellent [0.927; 0.977] | 0.91/excellent [0.83–0.94] |
| vertebra rotation (+max) | 0.605/moderate [0.439; 0.757] | 0.861/good [0.777; 0.923] | 0.66/moderate [0.49–0.78] |
| vertebra rotation (-max) | 0.801/good [0.691; 0.887] | 0.927/excellent [0.878; 0.960] | 0.69/moderate [0.53–0.80] |
| **length [mm]** | | | |
| pelvic obliquity | 0.825/good [0.724; 0.901] | 0.929/excellent [0.881; 0.961] | 0.84/good [0.73–0.90] |
| flèche lombaire | 0.897/good [0.831; 0.943] | 0.900/excellent [0.836; 0.945] | 0.95/excellent [0.91–0.97] |
| flèche cervicale VP | 0.887/good [0.815; 0.938] | 0.916/excellent [0.860; 0.954] | 0.93/excellent [0.87–0.96] |
| lateral deviation VP-DM (+max) | 0.763/good [0.637; 0.863] | 0.861/good [0.777; 0.923] | 0.85/good [0.74–0.90] |
| lateral deviation VP-DM (-max) | 0.703/moderate [0.558; 0.824] | 0.865/good [0.783; 0.925] | 0.71/moderate [0.55–0.81] |

The estimated *SEM* and *SDC* of the systems are given in Table 4. For the angular parameters, the *Formetric* showed an SDC ranging from 0.52° to 3.11°, while the *Postura* system had an SDC range of 0.52° to 5.11°. Notably, the kyphotic angle ICT-ITL [°] exhibited an increased SDC above 5° for the *Postura*. The *SDC* for the length parameter ranges from 2.06 mm to 7.46 mm for the *Postura* and from 1.22 mm to 6.89 mm for the *Formetric*.

### 3.3. Agreement between systems

The largest mean differences were found for the kyphotic angle ICT-ITL at −8.88°, the lordotic angle ITL-ILS at −1.64° and the fléche lombaire with −5.26 mm. The remaining angular parameters showed a mean difference between −0.29° and −0.79° and the remaining length parameter between −0.38 mm and +0.38 mm. The *LoA* between the systems ranged between ±0.9° and ±10.70° for the angular and between ±3.12 mm and ±12.83 mm for the length parameter. Agreements above ±5.0° or ±10 mm were found for the kyphotic angle ICT-ITL, the lordotic angle ITL-ILS, the vertebra rotation (-max), and the fléche lombaire with a *LoA* of ±10.70°, ±7.47°, ±5.60° and ±12.83 mm respectively. The estimated mean differences $\mu_D$ along with its standard deviation $\widehat{\sigma}_D$ and the *LoA* are presented in Table 5. In addition, the underlying within-subjects mean square errors (*MSW*) and between-subjects mean square errors (*MSB*) estimated from the one-way ANOVA are given. A Bland Altman plot for each parameter can be found in Appendix B.

## 4. Discussion

The *non-commercial Postura* prototype (2021) adapts the reconstruction algorithm and measurement protocol of the stationary *Formetric* system, while achieving portability by replacing the rasterstereography with an integrated Azure Kinect depth sensor for surface detection. To assess the uncertainties arising from these design changes, reliability analyses of both systems using ICC, SEM and SDC, together with Bland-Altman analysis for inter-system agreement, were performed on a cohort of 31 subjects without spinal deformity. The findings demonstrated good reliability and agreement for certain parameters, but revealed systematic errors in others, particularly in the kyphotic angle ICT-ITL and the lordotic angle ITL-ILS. To provide a basis for optimizing the *Postura* and other devices using the Azure Kinect depth sensor for surface-based posture analysis, the feasibility of the parameters and the causes of the identified errors are further discussed and recommendations to enhance accuracy are given.

For the inter-system comparison, the Formetric system was selected as the ideal reference. Given that both systems inherently contribute to measurement uncertainties, the Formetric results were further evaluated against those reported in a previous study by Degenhardt et al. [2]. These measurements were deemed a valid reference, as they demonstrated intraclass correlation coefficients (ICCs) that were comparable to those observed in the study by Degenhardt (Table 3). The comparison of ICC scores between the *Postura* and *Formetric* showed that, while the *Postura* system generally yielded lower scores, it still demonstrated good repeatability for eight out of eleven evaluated parameters. In contrast, the *Formetric* system exhibited excellent repeatability for six parameters and good repeatability for five. The *Postura* system matched the *Formetric* system's classifications for four parameters: sagittal imbalance, coronal imbalance, lordotic angle ITL-ILS, and lateral deviation (+max). For the remaining parameters, the *Postura* system was rated one classification level lower. The most notable differences in ICC scores were observed in lateral deviation (-max), vertebral rotations (±max) and kyphotic angle ICT-ITL.

The impact on reliability is also reflected in the finding that *Postura* tends to have higher SEM and SDC than *Formetric*. The *Postura* system achieved almost identical SEM and SDC values only for sagittal and coronal imbalance VP-DM. For the lordotic angle ITL-ILS and vertebral rotation (±max), the SEM increased slightly by 0.18° to 0.34°, resulting in a moderate increase in SDC of no more than +1°. The impact on the reliability of the length parameter was also considered moderate, with a SDC increasing by up to 1 mm for all measurements, except for lateral deviation, where the SDC increased by +1.16 mm. The most notable increase was observed again for the kyphotic angle ICT-ITL, with the SDC rising by +2° and reaching 5.11°, indicating the presence of a systematic error.

**Table 4**
SEM (Standard Error of Measurement) and SDC (Smallest Detectable Change) of DIERS Postura (P), DIERS Formetric (F) and Degenhardt et al. (D) [2]. Bold: parameters that exceed radiology reference SDC < 5° for angular parameter and SDC < 10 mm for length parameter.

| parameter | SEM F | SEM P | SEM D | SDC F | SDC P | SDC D |
|---|---|---|---|---|---|---|
| **angular [ ° ]** | | | | | | |
| sagittal imbalance VP-DM [°] | 0.40 | 0.38 | 0.35 | 1.11 | 1.06 | 0.97 |
| coronal imbalance VP-DM [°] | 0.18 | 0.19 | 0.3 | 0.50 | 0.52 | 0.83 |
| kyphotic angle ICT-ITL [°] | 1.12 | 1.84 | 1.15 | 3.11 | **5.11** | 3.19 |
| lordotic angle ITL-ILS [°] | 0.92 | 1.18 | 1.53 | 2.54 | 3.27 | 4.24 |
| vertebra rotation (+max) [°] | 0.45 | 0.75 | 1.45 | 1.24 | 2.08 | 4.02 |
| vertebra rotation (-max) [°] | 0.55 | 0.89 | 1.07 | 1.54 | 2.47 | 2.97 |
| **length [mm]** | | | | | | |
| pelvic obliquity [mm] | 0.44 | 0.74 | 1.52 | 1.22 | 2.06 | 4.21 |
| fleche lombaire [mm] | 2.49 | 2.69 | 1.53 | 6.89 | 7.46 | 4.24 |
| fleche cervicale VP [mm] | 1.87 | 1.80 | 1.96 | 5.20 | 4.99 | 5.43 |
| lateral deviation VP-DM (+max) [mm] | 0.92 | 1.33 | 1.44 | 2.54 | 3.70 | 3.99 |
| lateral deviation VP-DM (-max) [mm] | 0.65 | 0.86 | 1.17 | 1.81 | 2.39 | 3.24 |

**Table 5**
Mean diff (estimated mean difference), SD (standard deviation of the difference), LoA (Limits of Agreement), MSW (within-subject mean squared error), and MSB (between-subject mean squared error), Bold: LoA exceed the radiology reference of 5° for angular parameter and of 10 mm for length parameter.

| parameter | mean diff | SD | LoA | MSW | MSB |
|---|---|---|---|---|---|
| **angular [ ° ]** | | | | | |
| sagittal imbalance VP-DM | −0.29 | 1.01 | ±1.97 | 0.86 | 1.46 |
| coronal imbalance VP-DM | −0.63 | 0.21 | ±0.9 | 0.18 | 0.30 |
| kyphotic angle ICT-ITL | −8.88 | 5.46 | **±10.70** | 13.30 | 79.34 |
| lordotic angle ITL-ILS | −1.64 | 3.81 | **±7.47** | 6.68 | 38.12 |
| vertebra rotation (+max) | −0.79 | 1.96 | ±3.84 | 2.54 | 7.75 |
| vertebra rotation (-max) | −0.70 | 2.86 | **±5.60** | 3.38 | 22.57 |
| **length [mm]** | | | | | |
| pelvic obliquity | 0.38 | 1.59 | ±3.12 | 2.57 | 2.41 |
| fleche lombaire | −5.26 | 6.54 | **±12.83** | 29.91 | 81.55 |
| fleche cervicale VP | −0.38 | 4.23 | ±8.30 | 15.80 | 24.32 |
| lateral deviation VP-DM (+max) | 0.26 | 4.00 | ±7.84 | 8.10 | 39.70 |
| lateral deviation VP-DM (-max) | −0.06 | 2.52 | ±4.93 | 4.89 | 10.66 |

The highest agreement with the *Formetric* system was found for the sagittal imbalance VP-DM with a mean difference of −0.29° and a LoA of ±1.97°. The coronal imbalance VP-DM showed a slightly higher mean difference of −0.63° but exhibited the lowest LoA (±0.9°), indicating as well strong agreement. Combined with the observed reliability, these results demonstrate that the *Postura* achieves consistent outcomes for sagittal and coronal imbalance VP-DM, despite the transition to a depth camera. The results of the Bland-Altman analysis overall indicate that the *Postura* system underestimates all parameters, except for lateral deviation (+max) and pelvic obliquity, when compared to the *Formetric* system. However, for eight out of the eleven parameters, the mean difference was less

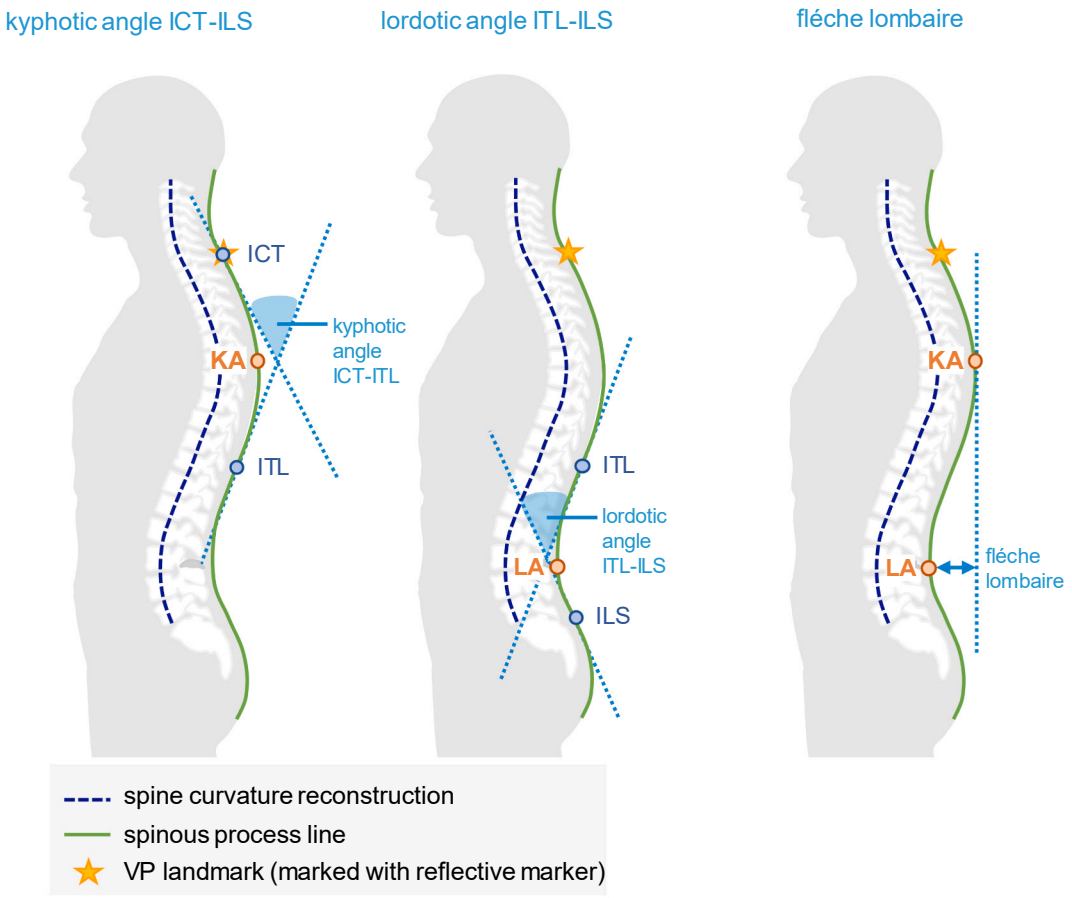



**Fig. 1.** Visualization of the calculation of the kyphotic angle ICT-ITL (left), the lordotic angle ITL-ILS (middle), and the fléche lombaire (right). The spinous process line (green) is determined from the surface measurement. The VP landmark (yellow star) is one of three anatomic landmarks, marked with reflective markers and used as key references by the reconstruction algorithm. Image adapted from Haimerl et al. [14].

than ±1° for angular measurements and less than ±1 mm for length parameters. This suggests that the lower surface accuracy of the depth camera only has a mild effect on the reliability of these parameters. Considering that the shift to a depth camera with lower accuracy compared to rasterstereography results in disagreements of up to 1°/1 mm, we identified three parameters exhibiting significantly higher mean differences and limits of agreement (LoA), indicating systematic errors: the kyphotic angle ICT-ITL (−8.88°, ±10.70°), the lordotic angle ITL-ILS (−1.64°, ±7.47°), and the fléche lombaire (−5.26 mm, ±12.83 mm).

A collaborative investigation with DIERS International GmbH identified inaccuracies in surface detection around the VP region as the root cause of these inconsistencies. These discrepancies were traced to the use of reflective markers, which are standard for anatomical landmarks in the *Formetric* system. However, when used with the Azure Kinect depth sensor, these markers introduced noise, leading to errors in surface detection around the marker's location in the *Postura* system. Since these landmarks establish two key reference lines, the midline (VP-DM) and the gravity line at VP, any inaccuracies in their detection can propagate through the reconstruction process, potentially impacting all calculated parameters. A significant impact was observed in the determination of the spinous process line (Fig. 1, green line) with the most notable errors occurring in the cervical region, where the curvature was almost flattened for some patients, as shown in Fig. 2. The parameters found to have the greatest discrepancies are precisely those that directly rely on the spinous process line for value determination.

As illustrated in Fig. 1, the kyphotic angle ICT-ITL is calculated as the angle between the surface tangents at the cranial-thoracic inclination point (ICT) and the lumbar-thoracic inclination point (ITL). Since the ICT point lies within the flattened region at the VP,

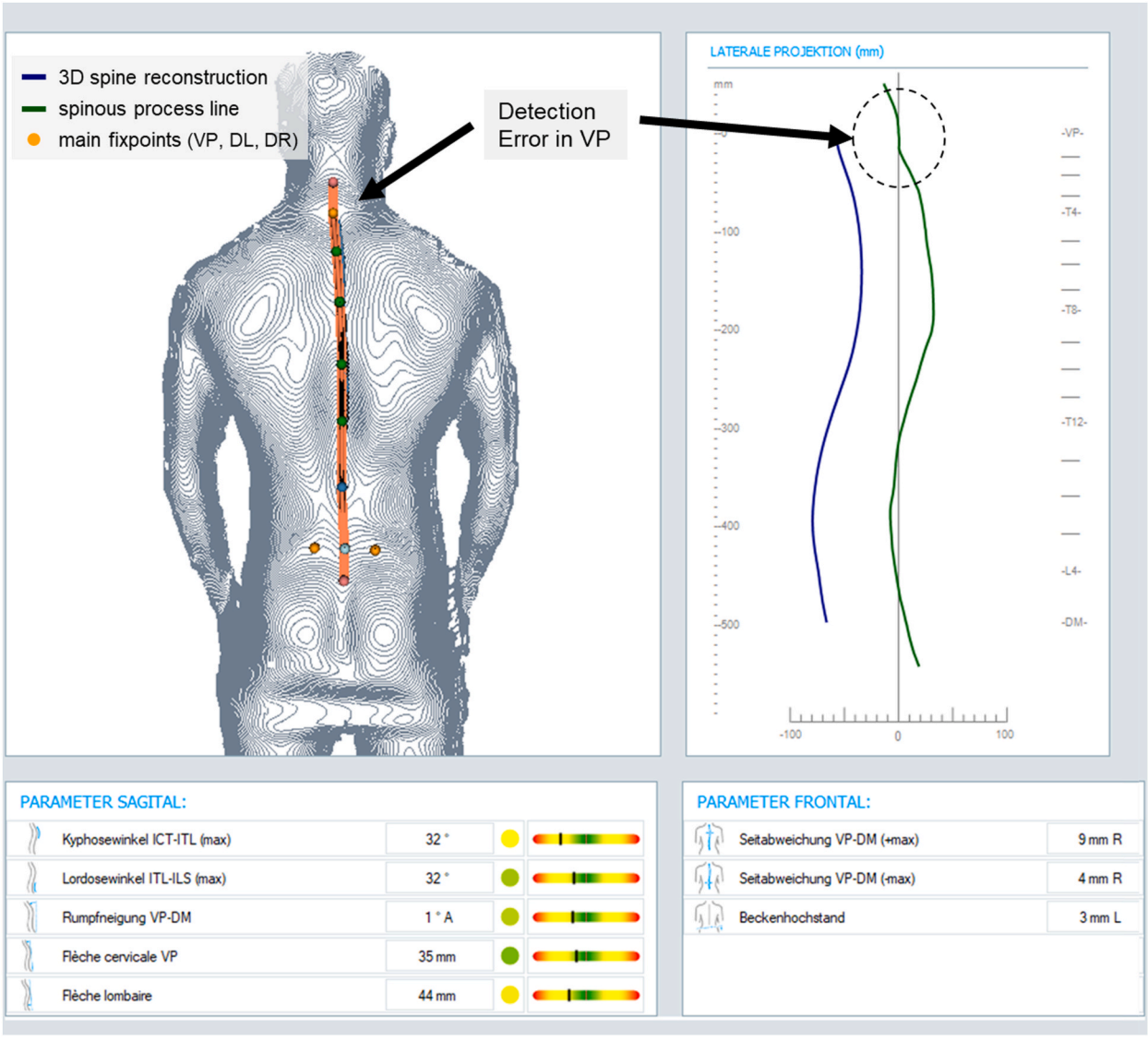



**Fig. 2.** Example of the *Postura* result overview: Erroneous spinous process line (green line) due to the erroneous detection of the surface in the area of the VP fixed point (yellow point).

the tangent at this point changes, resulting in a significant underestimation of the kyphotic angle ICT-ITL by the Postura system and contributing to high SEM and SDC values for this parameter. Additionally, the considerable discrepancies in both the lordotic angle ITL-ILS, calculated as the angle between the tangents at ITL and ILS, and the flèche lombaire, calculated as the distance between the lordotic and kyphotic apexes, suggest that the error is propagated throughout the determination of the spinous process line. As the spinous process line could not be correctly determined, particularly influencing the kyphotic angle ICT-ITL and lordotic angle ITL-ILS, it prevents the 2021 non-commercial Postura system from being a viable alternative to the Formetric system.

Based on this study, the commercial version of the Postura now uses non-reflective markers to address the source of the identified errors. Considering that the non-commercial prototype of the Postura has already shown good reliability and good system agreement for the less error-prone parameters, the commercial version could be a promising portable alternative to the Formetric. The measurement accuracy could be further improved by incorporating a 40-60 min warm-up period, as recommended by Tolgyessy et al. and Kurillo et al. Without such a warm-up phase, the temperature sensitivity of the Azure Kinect could lead to depth measurement errors of up to 2 mm [15,16]. However, the practical feasibility of implementing a warm-up period in clinical routine remains questionable.

To date, no published studies have evaluated either the commercial or non-commercial versions of the DIERS Postura. Thus, this study provides valuable insight into the technical and methodological challenges of replacing raster stereography with a depth sensor and discusses systematic errors introduced when using reflective markers with the Azure Kinect. These findings are particularly relevant for researchers aiming to refine or validate the Postura system and for the broader community using Kinect-based technologies in posture and gait analysis. Since Postura's overall measurement procedure is similar to Formetric's, users familiar with Formetric should be made aware of the key differences to prevent potential inaccuracies. With the main sources of error now addressed in the commercial version, further validation studies are needed to confirm its accuracy. This study offers both a proposed methodology and initial reference values for selected parameters, forming a basis for future validation efforts.

## 5. Limitations

The study identified measurement errors by analyzing the reliability of each system and the inter-system agreement for a cohort of thirty-one middle-aged participants without spinal deformities. Since the errors are systematic, they are likely to occur in all cohorts, though they are more easily identified in cohorts with lower within-group variance. However, the reliability results depend on the within-group variance and, therefore, can at best be applied to populations with similar characteristics measured by the same user in consecutive assessments. Due to interference between the measurement methods, the measurements could not be conducted simultaneously. Instead, they were performed sequentially within a 30-s window. Despite all efforts to minimize posture changes, these could not be considered fully excluded. For the inter-system comparison, the *Formetric* system was considered the ideal reference system. However, given that the Formetric system also exhibits a non-zero SEM, it contributes to the overall measurement disagreement observed between the two systems. To evaluate the systems capability to detect clinical changes, the SDC of X-ray was used as a threshold for SDC evaluation. However, this can only serve as a rough reference, emphasizing the need for established standards. Ideally, each parameter should have a clinically defined threshold, that the SDC of a measurement system must meet if used for diagnoses.

## 6. Conclusion

This study highlights the potential and limitations of the non-commercial (2021) prototype of the *Postura* system, which incorporates a depth sensor to achieve a compact, cost-effective alternative to rasterstereography. Despite the reduced accuracy of the depth camera in surface recognition, the Postura system demonstrated excellent reliability for one, good reliability for eight, and moderate reliability for two out of the eleven parameters evaluated. The system showed good agreement with the *Formetric* for eight parameters. However, a systematic error was identified, particularly affecting the surface detection around the vertebral prominence. This error, caused by the reflective markers, affected the determination of the spinous process line, leading to deviations of directly dependent parameters, in particular the kyphotic angle ICT-ITL, lordotic angle ITL-ILS and flèche lombaire. Therefore, the non-commercial prototype could not be considered a reliable alternative to the Formetric system. Following these findings, the commercial version of Postura addressed the error source by using non-reflective markers, but further validation is needed. As the first study of any Postura system, the findings reveal methodological differences and error sources when transitioning from raster-stereography to depth sensor for surface measurement, provide an initial benchmark for measurement accuracy, and establish a framework for the future validation of the now-available commercial system.

## AI policy statement

During the preparation of this work the author used DeepL and ChatGPT-4 to improve readability and grammar of the draft. After using this service, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

## CRediT authorship contribution statement

**Viktoria Kiaulehn:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis. **Jörg Eschweiler:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Roman Michalik:** Writing – review & editing, Methodology. **Miguel Pishnamaz:** Writing – review & editing, Methodology. **Frank Hildebrand:** Writing – review & editing,

Resources, Conceptualization. **Hannah Lena Siebers:** Writing – review & editing, Project administration, Methodology, Investigation, Conceptualization.

## Ethical approval statement

The study received approval from the local ethics committee, officially named the *"Ethik-Kommission an der Medizinischen Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen"* on December 25th, 2021. It was registered under the title *"Validierung des transportablen DIERS Postura über das DIERS Formetric" (English: "Validation of the Portable DIERS Postura Compared to the DIERS Formetric")*. The internal reference number assigned is EK377/21 and the CTC-A number is 21.348. All participants provided written informed consent for participation and data use.

## Data availability statement

The raw measurement data cannot be shared, as it includes images of individuals with uncovered backs. Results derived from these individual measurements can be made available in a de-identified format upon reasonable request. Access to the data requires signing a data access agreement. The study protocol is accessible through the clinical trial registry.

## Clinical trial

The clinical trial described in this paper was registered with the Center for Translational & Clinical Research (CTC-A) at the Medical Faculty of RWTH Aachen University under registration number 21-348 as well as in the German Clinical Trial Register under the registration number DRKS00026675.

## Funding

The study was funded by the German Federal Ministry of Education and Research (BMBF) through the *KMU innovativ: Medizintechnik* for the project: *R2WS: Patientenindividuelle Analyse der Wirbelsaeule (13GW0289B)*. The industrial partner involved was *DIERS International GmbH, Schlangenbad, Germany*, which provided the prototype. The authors declare that there are no conflicts of interest.

## Declaration of competing interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Joerg Eschweiler, Frank Hildebrand reports financial support was provided by German Federal Ministry of Education and Research (BMBF) - KMU innovativ. DIERS International GmbH, Schlangenbad, Germany was the industrial partner in the BMBF-KMU-innovativ project and provided the prototype. The authors declare that there are no conflicts of interest. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Descriptive Analysis of the Collective

**Table A**
Descriptive analysis (mean ± SD, range) bold: collective differs by more than one SD from the collective of Degenhardt et al., italic:. DIERS Postura measurements differ by more than one SD from DIERS Formetric measurements.

| parameter | Postura | Formetric | Degenhardt |
|---|---|---|---|
| **angular [ ° ]** | | | |
| sagittal imbalance VP-DM | 2.02 ± 2.15 [-3.6; 7.00] | 2.31 ± 2.25 [-2.80; 6.70] | 3.17 ± 2.18 |
| coronal imbalance VP-DM | **−1.48 ± 1.05 [-3.80; 1.20]** | **−0.85 ± 1.10 [-3.40; 2.20]** | 0.15 ± 0.66 |
| kyphotic angle ICT-ITL | *39.08 ± 8.09 [20.90; 56.40]* | 47.96 ± 8.36 [29.10; 68.30] | 48.47 ± 8.32 |
| lordotic angle ITL-ILS | 37.44 ± 9.96 [14.20; 61.60] | 39.09 ± 8.71 [17.70; 60.20] | 35.42 ± 7.55 |
| vertebra rotation (+max) | **1.84 ± 2.37 [-1.35; 10.96]** | **2.63 ± 2.37 [-0.19; 9.33]** | 5.68 ± 2.10 |
| vertebra rotation (-max) | −5.61 ± 3.96 [-16.52; 1.13] | −4.91 ± 4.05 [-16.22; 0.22] | −4.51 ± 2.4 |
| **length [mm]** | | | |
| pelvic obliquity | −0.60 ± 3.51 [-9.50; 8.10] | −0.97 ± 3.27 [-9.70; 7.80] | −0.17 ± 2.92 |
| flèche lombaire | 41.66 ± 11.08 [14.50; 73.70] | 42.04 ± 11.72 [12.70; 72.60] | 37.43 ± 12.18 |
| flèche cervicale VP | 36.39 ± 15.83 [5.90; 69.20] | 41.65 ± 16.93 [11.00; 82.60] | 48.17 ± 13.19 |
| lateral deviation VP-DM (+max) | 8.19 ± 5.41 [0.00; 27.00] | 7.93 ± 4.86 [0.00; 20.00] | 8.04 ± 5.13 |
| lateral deviation VP-DM (-max) | −2.57 ± 3.13 [-16.20; 0.00] | −2.51 ± 3.51 [-17.20; 0.00] | −4.62 ± 2.92 |

## Appendix B. Bland Altman Plots

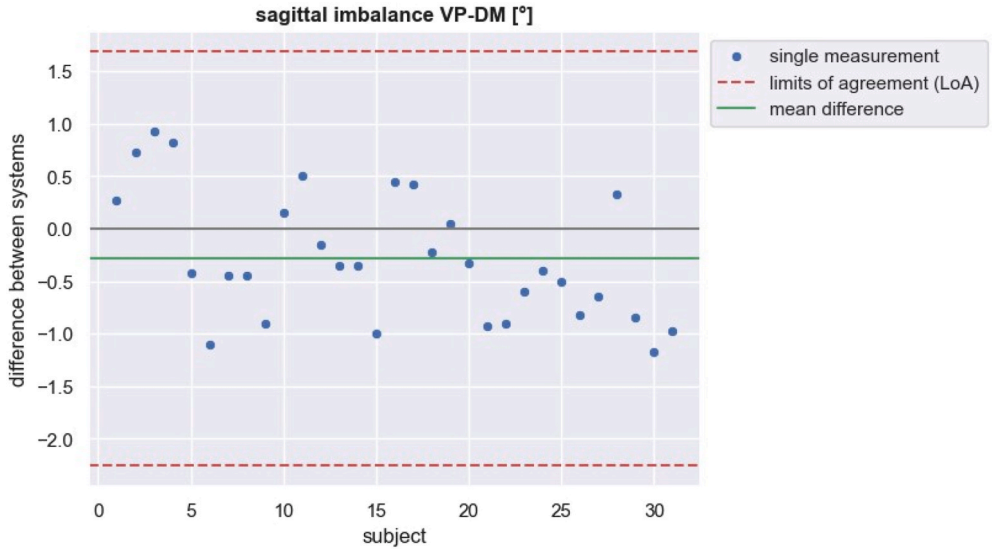



**Fig. B1.** Bland Altman plot for the sagittal imbalance VP-DM [°]

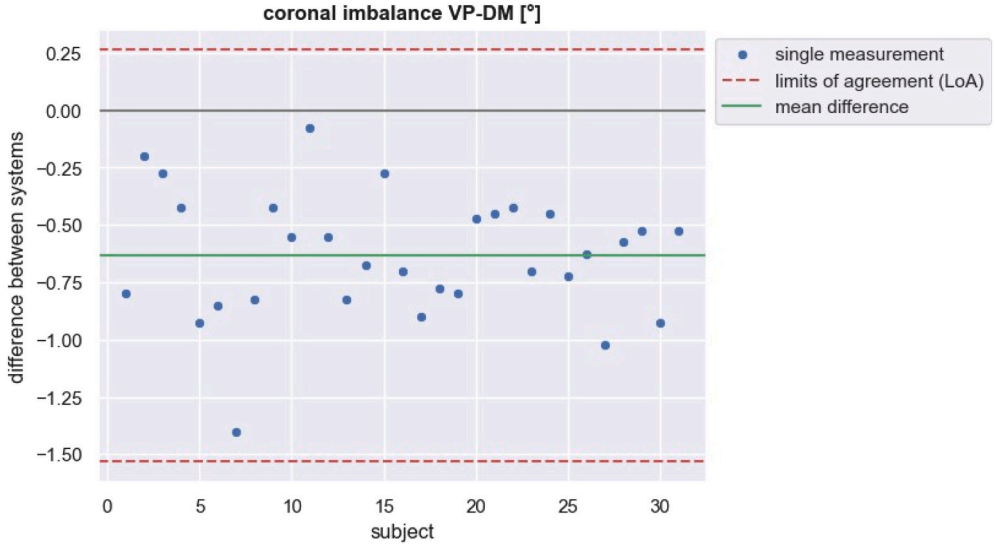



**Fig. B2.** Bland Altman plot for the coronal imbalance VP-DM [°]

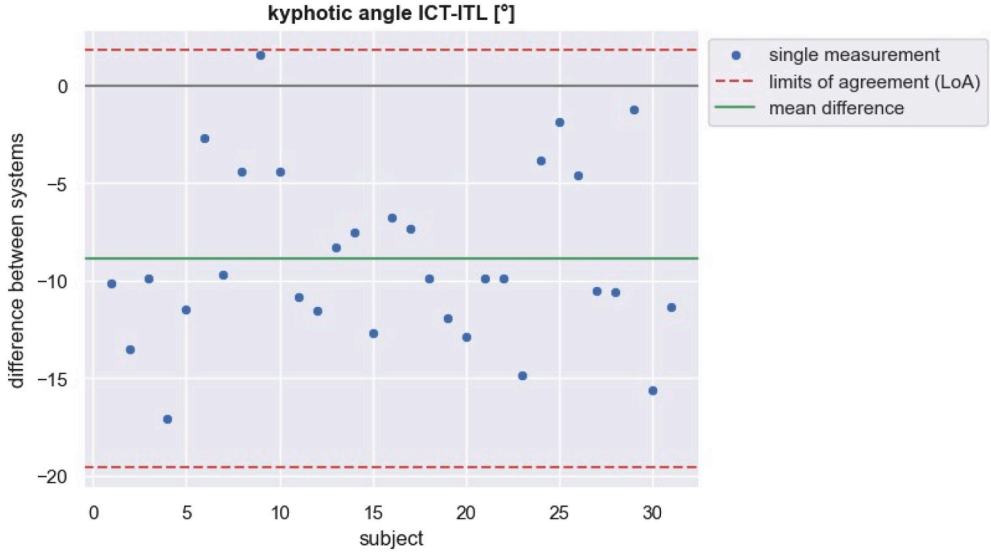



**Fig. B3.** Bland Altman plot for the kyphotic angle ICT-ITL [°]

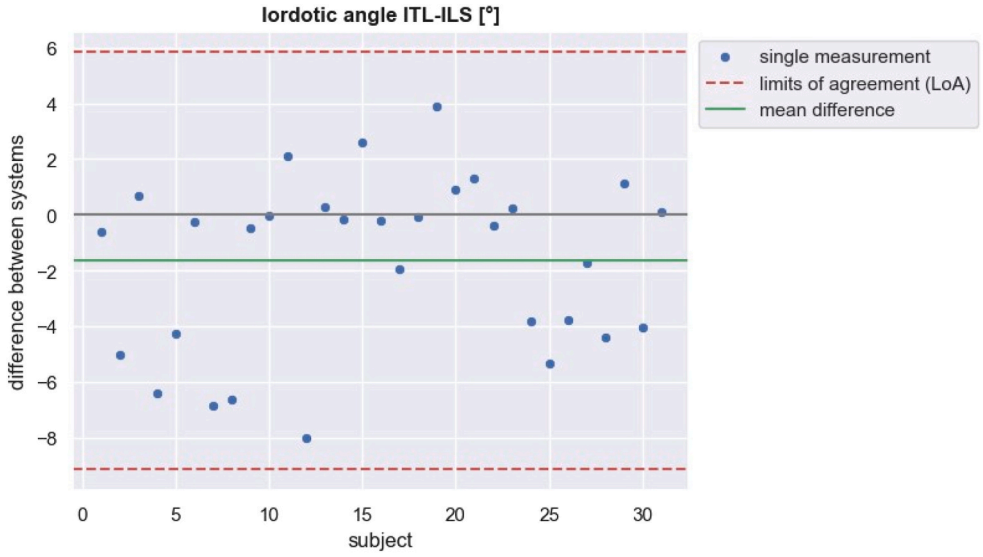



**Fig. B4.** Bland Altman plot for the lordotic angle ITL-ILS [°]

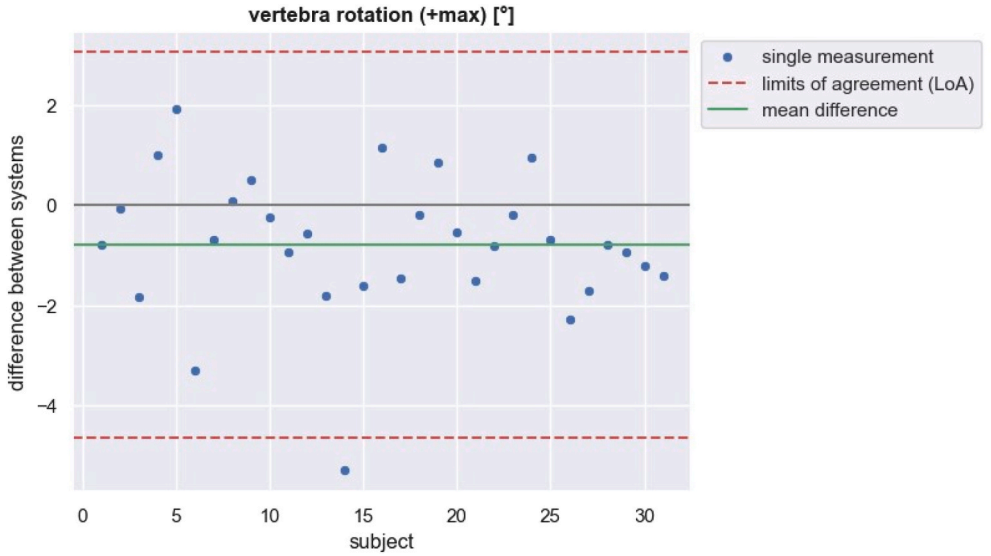



**Fig. B5.** Bland Altman plot for the vertebra rotation (+max) [°]

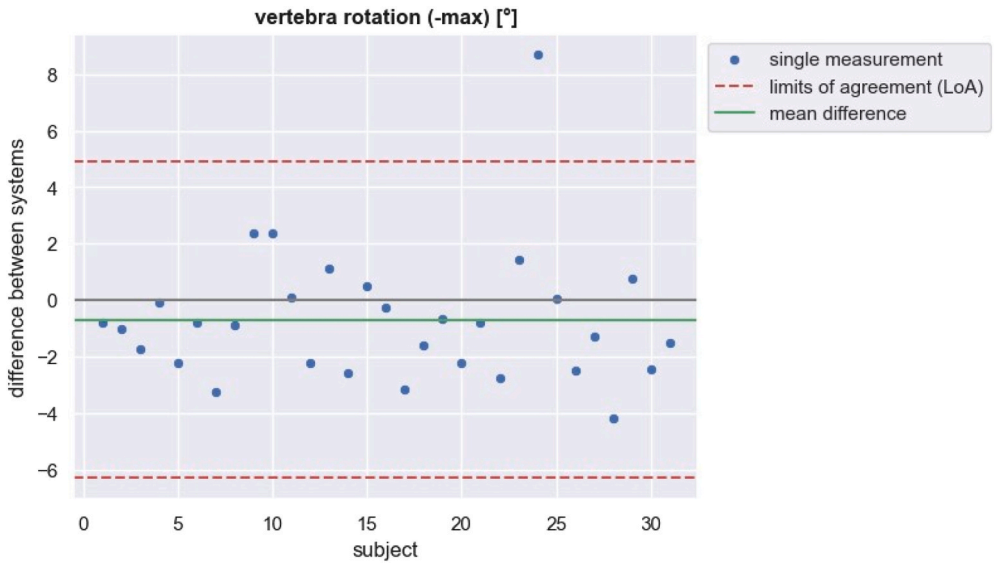



**Fig. B6.** Bland Altman plot for the vertebra rotation (-max) [°]

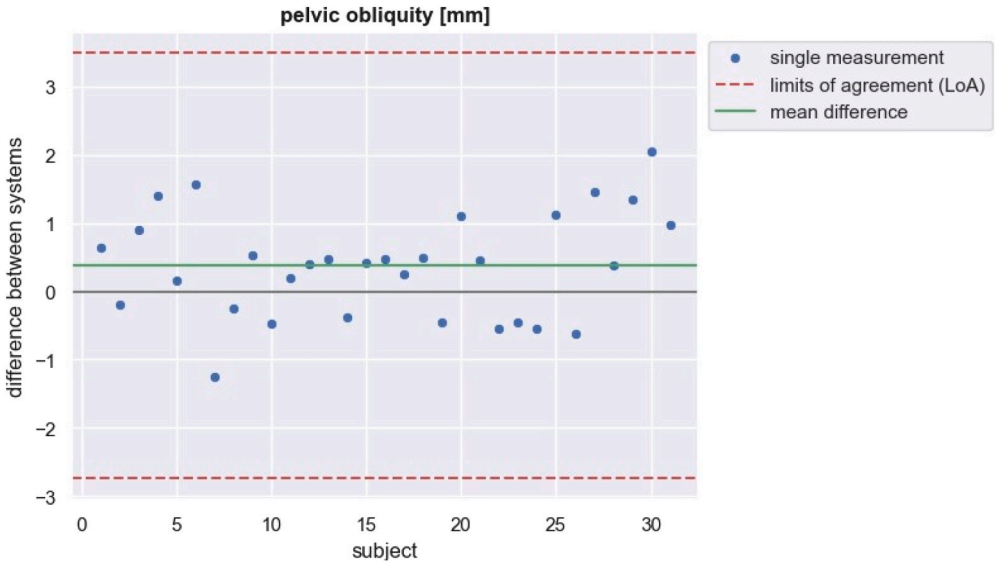



**Fig. B7.** Bland Altman plot for the pelvic obliquity [mm]

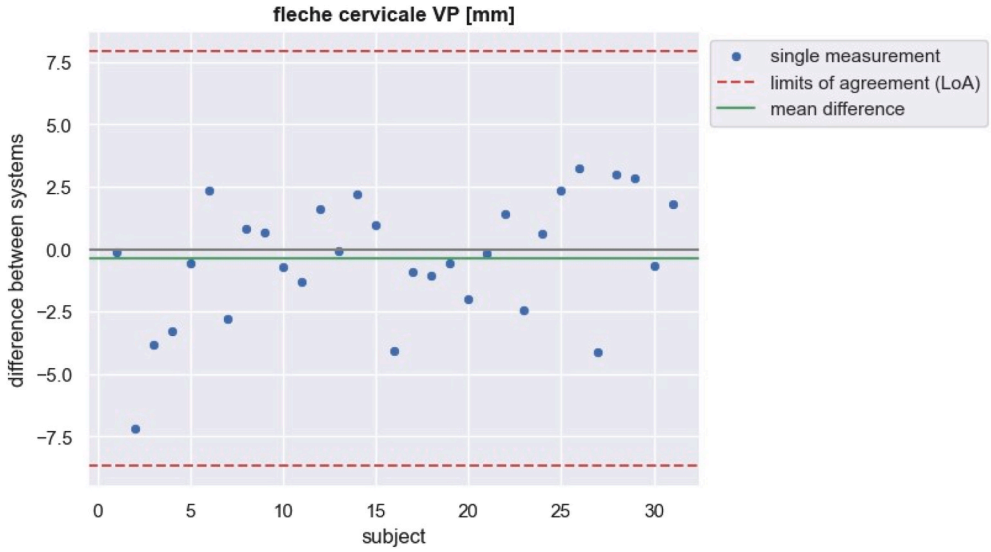



**Fig. B8.** Bland Altman plot for the fleche cervicale VP [mm]

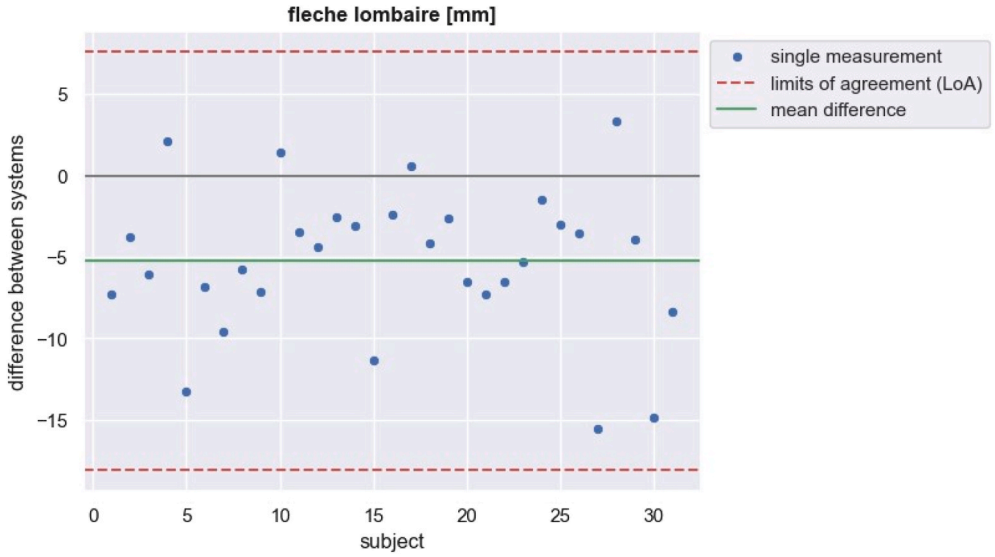



**Fig. B9.** Bland Altman plot for the fleche lombaire [mm]

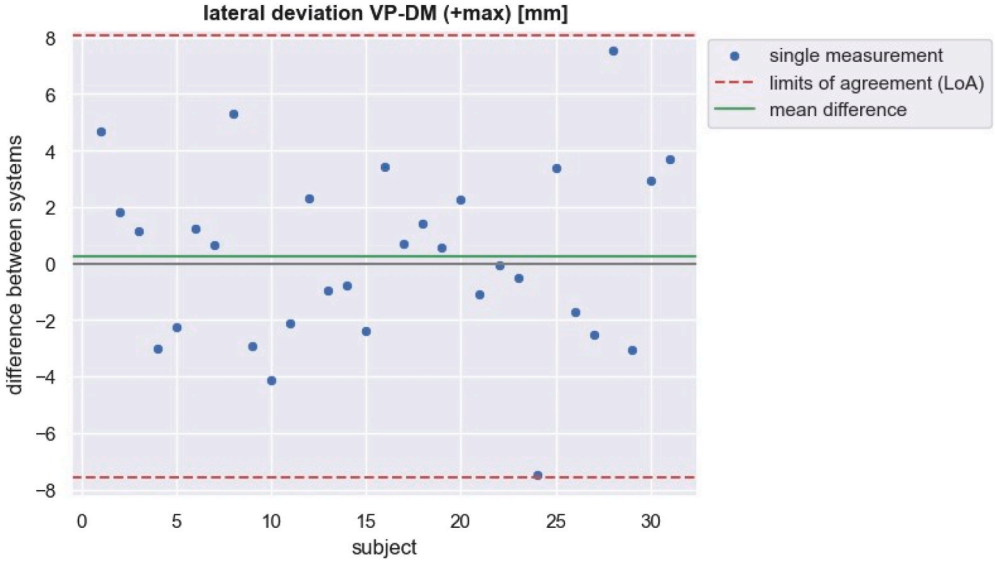



**Fig. B10.** Bland Altman plot for the lateral deviation VP-DM (+max) [mm]

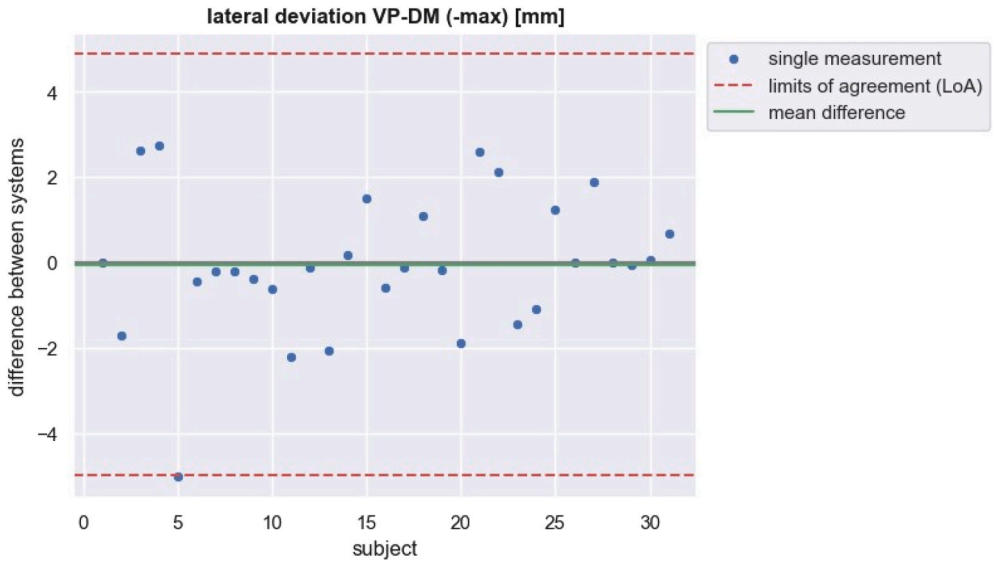



**Fig. B11.** Bland Altman plot for the lateral deviation VP-DM (-max) [mm]

## Appendix C. Eligibility Criteria

*Inclusion Criteria:*

- Healthy participants
- Participants capable of understanding the research project
- Adult participants (age > 18 years)

*Exclusion Criteria:*

- Presence of any orthopedic, neurological, or muscular disorders that affect movement
- Implanted pacemaker
- Participants are unable to provide consent or those with a legal guardian
- Age under 18 years
- Pregnant participants
- Obese participants with a BMI > 35
- Unreliable participants (e.g., alcoholism, drug addiction, mental illness)
- Participants with a potential SARS-CoV-2 infection will be excluded for the duration of the infection
- Tattoo on the spine

## *Appendix D. STROBE Statement*

| | Item No. | Recommendation | **Page** No. | Relevant text from manuscript |
|---|---|---|---|---|
| **Title and Abstract** | | | | |
| Title and abstract | 1 | (*a*) Indicate the study's design with a commonly used term in the title or the abstract | 2 | Page 2, Lines 2-6 Purpose |
| | | (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | 2 | Page 2, Lines 2-14<br>"what was done": Purpose and Methods<br>Page 2 Lines 16-30<br>"what was found": Results and Conclusion: |
| **Introduction** | | | | |
| Background/ rationale | 2 | Explain the scientific background and rationale for the investigation being reported | 3 | Page 3, Lines 2-7 Motivation<br>Page 3; Lines 8-26: State of the Art (Formetric) and new Design (Postura) |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | 3 | Page 3, Lines 27-34 |
| **Methods** | | | | |

(*continued on next page*)

(*continued*)

| | Item No. | Recommendation | **Page** No. | Relevant text from manuscript |
|---|---|---|---|---|
| Study design | 4 | Present key elements of study design early in the paper | 4,5 | Page 4 & 5 Section "Test-setup"<br>The overall test setup is explained first in Methods before explaining the statistical methods in detail. |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | 4 | Page 5, Lines 16-20 |
| Participants | 6 | Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up<br>For matched studies, give matching criteria and number of exposed and unexposed | 4 | Page 4, Lines 4-5 (Appendix C) and Lines 9-12<br>List of eligibility critria and describtion of choice<br>Not a matched study (all patients were accessed with both systems) |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | 4,5,6 | Page 4, Table 1 "Description of the DIERS parameter investigated in this study" gives a description of all studied parameter<br>Page 5: Subsection 2.2<br>Definition of the ICC classification (Table 2) is used to rate the outcome. Mathematical Expression of the Standard Error of Measurement (SEM – Eq. (2)) and the Smallest Detectable Change (SDC – Eq. (3)) as well as the selected thresholds.<br>Page 6: Subsection 2.3:<br>Mathematical Expression of the Bland Altman method for repeated measurements and the Limits of Agreement |
| Data sources/ measurement | 8* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | 3,5,6 | Page 3: Line 7-26: Explanation of both systems<br>Page 5: Line 7-15: Explanation of the setup for comparable measurements<br>Page 6: Mathematical Definition for Agreement between Systems. |
| Bias | 9 | Describe any efforts to address potential sources of bias | 5,6 | Page 5: Line 7-15: Explanation of the setup for comparable measurements<br>Page 6: Line 18-19 (using a Bland Altman method for repeated measurements) |
| Study size | 10 | Explain how the study size was arrived at | 4 | Page 4: Line 3-4 |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | 5,6 | Page 5: Subsection 2.2<br>All Equations of the calculated Variables are given.<br>Page 6: Subsection 2.3:<br>All Equations of the calculated Variables are given. |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | 5,6 | Page 5: Subsection 2.2: Explain following methods: ICC, SEM (Standard Error of Measurement) and SDC (Smallest Detectable Change)<br>Page 6: Subsection 2.3: Explains following methods: bland Altman for repeated measurements, Limits of Agreement (LoA)<br>All Equations of the calculated Variables are given. |
| | | (*b*) Describe any methods used to examine subgroups and interactions | 5,6 | See (12a) |
| | | (*c*) Explain how missing data were addressed | | No missing data |
| | | (*d*) If applicable, explain how loss to follow-up was addressed | | Not applicable (all patient did both measurements – no follow up) |
| | | (*e*) Describe any sensitivity analyses | | No sensitivity analysis |
| **Results** | | | | |
| Participants | 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analyzed | 7 | Page 7: Line 3 |
| | | (b) Give reasons for non-participation at each stage | 7 | Page 7: Line 3 |
| | | (c) Consider use of a flow diagram | | Considered not to use (only 2 stages 33 recruited, 31 measured) |
| Descriptive data | 14* | (a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders | 7 | Appendix A |
| | | (b) Indicate number of participants with missing data for each variable of interest | | No missing data |
| | | (c) Summarise follow-up time (e.g., average and total amount) | | Not applicable as the study is not a follow up study. |
| Outcome data | 15* | Report numbers of outcome events or summary measures over time | 5 | Page 5: Line 7-8: Time between the assessments with the different measurements (30 s) |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, | 7,9 | Page 7: Section 3.2<br>The ICC for each system is provided<br>Page 9: Table 5: |

(*continued on next page*)

(*continued*)

| | Item No. | Recommendation | **Page** No. | Relevant text from manuscript |
|---|---|---|---|---|
| | | 95% confidence interval). Make clear which confounders were adjusted for and why they were included | | The LoA for the systems agreement are provided Bland Altman Plots are provided for all variables (Appendix C) |
| | | (*b*) Report category boundaries when continuous variables were categorized | 5,6 | For ICC (Page 5: Table 2) For SDC (Page 6: Line 9-15) |
| | | (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | | Not relevant |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | | None other analysis |
| **Discussion** | | | | |
| Key results | 18 | Summarise key results with reference to study objectives | 14 | Page 14: Subsection 6: Conclusion |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | 13 | Page 13: Subsection 5 Limitation |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | 10-12 | Page 10-12: Subsection 4: Discussion |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | 13 | Page 13: Subsection 5: Limitations, Line 2-7 |
| **Other information** | | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | 15 | Page 14: Line 15-19 |

## References


[1] B. Drerup, E. Hierholzer, Automatic localization of anatomical landmarks on the back surface and construction of a body-fixed coordinate system, J. Biomech. 20 (10) (1987) 961–970.
[2] B.F. Degenhardt, Z. Starks, S. Bhatia, Reliability of the diers formetric 4d spine shape parameters in adults without postural deformities, BioMed Res. Int. 2020 (2020).
[3] DIERS International GmbH, M.K. Katrin Gerlach, 3D/4D Spine and PostureAnalysis - Therapeutic Consequenceswith Case Studies from Actual Practice, 2014.
[4] Microsoft: azure kinect DK hardware specification - depth camera supported operating modes. https://learn.microsoft.com/en-us/azure/kinect-dk/hardware-specification#depth-camera-supported-operating-modes (Accessed: 2021-October-15).
[5] T.K. Koo, M.Y. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, J. Chiropract. Med. 15 (2) (2016) 155–163.
[6] R.E. Will, I.A. Stokes, X. Qiu, M.R. Walker, J.O. Sanders, Cobb angle progression in adolescent scoliosis begins at the intervertebral disc, Spine 34 (25) (2009) 2782–2786.
[7] B.V. Marchetti, E. Raupp, J.A. Sedrez, R.P. Ribeiro, C.T. Candotti, The importance of clinical experience for the measurement of scoliosis curve in children by cobb technique, Fisioterapia e Pesquisa 26 (2019) 241–246.
[8] D.S. Jun, W.J. Shin, B.K. An, J.W. Paik, M.H. Park, The relationship between the progression of kyphosis in stable thoracolumbar fractures and magnetic resonance imaging findings, Asian Spine J. 9 (2) (2015) 170.
[9] C. Hasler, Beinlängendifferenzen, Orthopä 29 (9) (2000) 766–774.
[10] C. Konrads, M. Rudert, Klinische Tests Und Untersuchung in Orthopaedie Und Unfallchirurgie, Springer, 2018.
[11] X. Gao, L. Wang, S. Li, P. Wang, J. Zhang, Y. Shen, Predictors for postoperative loss of lumbar lordosis after long fusions arthrodesis in patients with adult scoliosis, Med. Sci. Monit.: Int. Med. J. Exp. Clin. Res. 24 (2018) 531.
[12] J.M. Bland, D.G. Altman, Agreement between methods of measurement with multiple observations per individual, J. Biopharm. Stat. 17 (4) (2007) 571–582.
[13] R.A. Parker, C. Scott, V. Inácio, N.T. Stevens, Using multiple agreement methods for continuous repeated measures data: a tutorial for practitioners, BMC Med. Res. Methodol. 20 (1) (2020) 1–14.
[14] M. Haimerl, A. Linkerhägner, J. Konradi, C. Wolf, P. Drees, U. Betz, Dynamic surface topography data for assessing intra-and interindividual variation of vertebral motion, Data Brief 48 (2023) 109178.
[15] M. Tölgyessy, M. Dekan, L. Chovanec, P. Hubinský, Evaluation of the azure kinect and its comparison to kinect v1 and kinect v2, Sensors 21 (2) (2021) 413.
[16] G. Kurillo, E. Hemingway, M.-L. Cheng, L. Cheng, Evaluating the accuracy of the azure kinect and kinect v2, Sensors 22 (7) (2022) 2469.