ABSTRACT

Objective: This study evaluates the manual and digital reliability and reproducibility of five methods of measuring deformity (kyphosis) in thoracolumbar burst fractures. Method: Ninety (90) tomographic images were evaluated and, in each case, kyphotic deformity was measured, both manually and digitally, through the five most relevant methods described in the literature. For the assessment of intraobserver error, 20 cases were measured again. Results: The results show that all five methods are highly reliable and digitally reproducible, with estimated error near or lower than that indicated in the intraobserver error analysis. Cobb’s method had the highest concordance (96%) while the sagittal index had the lowest concordance (75%). It is also suggested that digital assessment is more reliable than the manual method. Conclusion: All tested methods are highly reliable and digitally reproducible.

Keywords: Spinal fractures; Kyphosis; Spine.

INTRODUCTION

Thoracolumbar burst fractures are currently a very frequent diagnosis, associated with the growing incidence of high-energy traumas. Radiographic evaluation of the patient with this type of injury should be done in a detailed way, looking for signs of possible damage to the bone support and muscle-ligament structures next to the vertebral spine.1

The main characteristics of thoracolumbar burst fractures, as described by Denis in 1983, are involvement of the middle spine with fracture associated with dislocation or rotation of the posterior cortical of the vertebral body, compressing the spinal canal and altering the stability of the vertebral spine. Axial trauma concentrates a high load directly on the middle column, in the region between the vertebral pedicles, which results in them moving apart, enabling retropulsion of fragments from the posterior portion of the vertebral body towards the medullary canal and kyphosis deformity.2

Detailed study of the trauma mechanism and the morphology of burst fractures has recently gained importance. The classification of fractures has been an important point, not only for the establishment of their diagnosis, but also for their treatment.2 The use of the classifications proposed in the literature so far assists in the standardization of conduct and facilitates communication between the medical services. The classifications that exist today take into account various factors, namely: mechanism and energy of the trauma, and the complexity of the lesion in the bone structures and neighboring structures, enabling more objective conditions for establishing the therapeutic conduct and prognosis of these lesions.3

Kyphosis deformity plays a key role in the evaluation of patients with thoracolumbar burst fracture, and there are different methods for measuring it, including the Cobb method, the Gardner method, and the sagittal index.4

The evaluation of kyphosis angle in thoracolumbar burst fractures is used to assist in the indication of surgical procedures. The
angle of deformity can be measured regionally, from the segment or from just one vertebra. Although there are different methods of measuring the angle of kyphosis deformity, the reliability and reproducibility of these methods are not well-defined. The treatment of thoracolumbar burst fractures involves many factors, including the resulting kyphosis angle. Although various methods have been described, no study has directly compared these methods in terms of reliability and reproducibility. The advent of tomography, introduced by Hounsfield apud Neves and Avanzí in 1973, enabled a closer analysis of these characteristics, particularly in relation to fragments dislocated to the interior of the canal, enabling the percentage of post-traumatic stenosis to be evaluated. Various authors have analyzed and classified thoracolumbar fractures using this imaging method. Decompression of the vertebral canal, associated with stabilization with correction of the kyphosis, still offers the best results, as we know that retropropulsion of the fragments towards the vertebral canal is related to neurological damage, and the correlation between the narrowing of the vertebral canal and the severity of the neurological deficit is also recorded in the literature. Various studies have measured the bone changes evidenced in the computed tomography manually, with a millimeter ruler. The present study is necessary because, with the advent of technology, the use of digital programs provides better visualization and measurement of the kyphosis deformity in computed tomography, and the use of this exam has increased in the daily routine of hospital settings. This study therefore aims to evaluate the manual and digital reliability and reproducibility of the five methods used to measure kyphosis deformity in thoracolumbar burst fractures.

**MATERIALS AND METHOD**

After approval by the institution’s Ethics Committee, 90 tomographic images were taken in sagittal section of the thoracolumbar spine of patients with a diagnosis of thoracolumbar burst fracture, treated and followed up in the Department of Orthopedics and Traumatology of the Santa Casa de Misericórdia de São Paulo. Exams of patients with other morphological types of thoracolumbar fractures, fractures in two or more vertebrae, and those with poor quality tomographies for checking measurements were excluded. In each case, the angular value of the kyphosis deformity was measured, through the five most relevant methods described in the current literature. The Cobb method (method I); the Gardner method for measuring segmental deformity (method II), which measures the angle between the inferior plateau of the fractured vertebra and the superior plateau of the superior vertebra; the sagittal index (method III), which measures the angle between the posterior walls of the fractured vertebra and the vertebra immediately below it; isolated kyphosis of the fractured vertebra including the intervertebral discs, i.e. the angulation between the inferior plate of the intact vertebra immediately superior to the fractured vertebra and the superior plate of the intact vertebra immediately inferior to it (method IV); and isolated kyphosis of the fractured vertebra, excluding the intervertebral discs (method V). (Figure 1)

The measurements were carried out by an orthopedic doctor who was familiar with the technique. They were done manually, using a millimeter ruler, and also digitally, using the AutoCAD software, at independent times, without knowledge of the values obtained in the measurements using the other methods. For evaluation of the error of the evaluator’s measurement, 20 tomographic exams were measured again, both manually and digitally, using the same five methods. For each measurement carried out through each method, the mean and the standard deviation were obtained and the intraclass correlation coefficients were calculated, with the respective 95% confidence intervals, to evaluate the intra-examiner agreement/reproducibility. The intraclass correlation coefficient was also used to evaluate the reproducibility between the methods for each measurement, and Bland-Altman graphs were created to evaluate any trend between the methods.

**RESULTS**

Table 1 shows that the intra-examiner agreement of all the measurements performed by the manual method is high, with values of ICC > 0.7. The lowest agreements/reproducibilities were obtained for measurements III and V, which resulted in the highest estimated errors. In the digital method, all the measurements presented high intraclass correlation coefficients (ICC > 0.95), with estimated errors lower than 2.5 degrees.

Table 2 shows that the measurements present high agreements/reproducibilities between the methods (ICC > 0.75), with the poorest agreement/reproducibility being that of measurement III (ICC = 0.75). The estimated errors are lower than 4 degrees in all the groups, ranging from 2.26 for measurement I and 3.76 for measurement III.

**DISCUSSION**

In the evaluation of intra-examiner agreement of the manual measurements, a higher level of agreement was found between the first and second sequences of measurements for the Cobb (method I) and Gardner (method II) methods, and isolated kyphosis including the adjacent intervertebral discs (method IV), with agreement close to 95% and estimated error close to 2 degrees. The method of isolated kyphosis excluding the adjacent intervertebral discs (method V) and especially the sagittal index (method III) obtained agreement lower than 80%, with estimated error of 3.7 and 4.7 degrees, respectively, which are representative errors, but acceptable.

It is important to observe that methods III and V, with poorer performance, are based on lines drawn on surfaces of fractured vertebrae, which apparently diminishes their precision and agreement. It is worth emphasizing that method II, specifically, because it measures the sagittal index indicated by lines on the posterior wall of the fractured vertebra and the vertebra immediately below it, is more comfortably and more accurately measured in vertebrae with a clearly-defined posterior wall, a characteristic which, due to the nature of burst fractures (involvement of the Denis posterior spine and retropropulsion of the fragment towards the vertebral canal), in the majority of cases was not present in the tomographic exams, making this method the least reliable of the five methods evaluated.

On the other hand, evaluation of the intra-examiner agreement of the digital measurement of the angles of kyphosis deformity presented unexpected agreement, higher than 95% in all five measurement methods, with estimated error of between 1.46 and 2.25 degrees. This strongly suggests that the digital measurement has greater reliability than the manual method.

In the comparison between the manual and digital measurement groups, Table 2 clearly shows again methods I, II and IV as the most reproducible, with agreement higher than 90% and estimated error of between 2.26 and 2.51 degrees, method I (Cobb) being the one with highest agreement, with 96% of reliability (CI 89 to 98%). Once again, method V and, in particular, method III (sagittal index) obtained lower agreement values, reaching 75% (CI 58 to 84%) in this
Table 1. Description in degrees of the measurements performed through each method, by the examiner, and results of intraclass correlations for evaluation of intra-examiner error.

<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement</th>
<th>Evaluation</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>ICC</th>
<th>CI (95%)</th>
<th>Estimated error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inferior</td>
<td>Superior</td>
</tr>
<tr>
<td>Manual</td>
<td>First</td>
<td>9.55</td>
<td>9.5</td>
<td>20</td>
<td></td>
<td>0.95</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>8.8</td>
<td>9.16</td>
<td>20</td>
<td></td>
<td>0.95</td>
<td>0.87</td>
<td>0.98</td>
</tr>
<tr>
<td>Digital</td>
<td>First</td>
<td>-0.05</td>
<td>10.21</td>
<td>20</td>
<td></td>
<td>0.73</td>
<td>0.44</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>-2.05</td>
<td>7.97</td>
<td>20</td>
<td></td>
<td>0.93</td>
<td>0.82</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>6.25</td>
<td>7.45</td>
<td>20</td>
<td></td>
<td>0.79</td>
<td>0.55</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>7.2</td>
<td>8.86</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>11.2</td>
<td>9.95</td>
<td>20</td>
<td></td>
<td>0.95</td>
<td>0.88</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>11.85</td>
<td>10.07</td>
<td>20</td>
<td></td>
<td>0.95</td>
<td>0.88</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>15.4</td>
<td>8.16</td>
<td>20</td>
<td></td>
<td>0.96</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>15.85</td>
<td>8.44</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>-2.1</td>
<td>10.32</td>
<td>20</td>
<td></td>
<td>0.95</td>
<td>0.88</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>-3.15</td>
<td>10.52</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>5</td>
<td>9.13</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>5.2</td>
<td>10.58</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>First</td>
<td>17.45</td>
<td>9.01</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Second</td>
<td>17.1</td>
<td>9.21</td>
<td>20</td>
<td></td>
<td>0.97</td>
<td>0.94</td>
<td>0.99</td>
</tr>
</tbody>
</table>

1. Cobb; II. Gardner segmental deformity; III. Sagittal index; IV. Isolated kyphosis including the adjacent intervertebral discs; V. Isolated kyphosis excluding the adjacent intervertebral discs.

latter, with estimated error of 3.04 and 3.76 degrees, respectively, which is probably due to the particularities of burst fracture mentioned above.

The errors estimated between the manual and digital measurement groups, independent of the variable values between the methods, are close or lower than the estimated intra-examiner errors, which indicates that the reproducibility between the methods is highly reliable.

**CONCLUSION**

All five manual methods of measuring kyphosis in thoracolumbar burst fracture are highly reliable and digitally reproducible. It is suggested that the digital measurement is more reliable than the manual measurement. Further studies are needed, with a higher number of cases and examiners, so that digital measurement can be recommended as standard.

All authors declare no potential conflict of interest concerning this article.

**REFERENCES**