ASSESSMENT OF INTEROBSERVER RELIABILITY OF SINGH INDEX FOR GRADING OF OSTEOPOROSIS

Syed Imran Bukhari, Raja Irfan Qadir, Shandana Khan, Adnan Khan
Department of Orthopaedics & Spine Surgery, NorthWest General Hospital & Research Centre, Peshawar - Pakistan

ABSTRACT

Objective: Assessment of the reliability of Singh's Index for measurement of osteoporosis between different observers.

Material & Methods: It was a cross-sectional validation study, conducted in the Orthopaedics and Radiology Departments, NorthWest General Hospital & Research Centre, Peshawar, Pakistan, involving 120 post-menopausal women from June 2015 to December 2015. X-ray Pelvis was undertaken in all patients.

Results: Mean age was 62.91 years ± 9.6 SD. Single measurement was 43%. Overall agreement among the observers was 69.5% with p value .000, which is significant. 95% Confidence interval was 0.587 to 0.779. The kappa value ranged from .171 to .261 with a slight to fair agreement.

Conclusion: Singh Index has 69.5% inter-observer reliability.

Key Words: Singh Index, Osteoporosis.

This article may be cited as: Bukhari SI, Qadir RI, Khan S, Khan A. Assessment of interobserver reliability of singh index for grading of osteoporosis. J Med Sci 2017; 25: (1) (Supplement) 175-178.

INTRODUCTION

Osteoporosis is characterized by low bone mass, deterioration of bone tissue and disruption of bone architecture, compromised bone strength and an increase in the risk of fracture¹.

Osteoporosis is a silent disease until fracture occurs after minimal trauma. These fractures are common with a predilection for hip/wrist/vertebra and place an enormous medical and personal burden on aged population and a major economic toll on the national exchequer. In particular, hip fractures result in 10 to 20 percent excess mortality within one year²; additionally, hip fractures are associated with a 2.5 fold increased risk of future fractures. Approximately 20 percent of hip fracture patients require long-term nursing home care, and only 40 percent fully regain their pre-fracture level of independence³.

Osteoporosis can be diagnosed before the catastrophic complication, i.e, fracture occurs. Plain radiographs allow qualitative and semi-quantitative evaluation of osteoporosis, whereas other imaging techniques allow quantification of bone loss (e.g, dual-energy x-ray absorptiometry and quantitative computed tomography [CT]), assessment for the presence of fractures (morphometry), and the study of bone properties (ultrasonography)⁴.

Singh index is a diagnostic classification which has been subject of scrutiny for long regarding its reproducibility and reliability⁵-⁷. Proponents cite its cost-effectiveness for evaluation of osteoporosis, however after the introduction of new imaging modalities, its significance has waned. We have tried to establish its accuracy by assessing the interobserver reliability.

MATERIAL & METHODS

One hundred & twenty post-menopausal women with clinical risk factors (Age, Low body mass index, prior or fracture after age 50, parental history of hip fracture, current smoking habit, current or past use of systemic steroids, alcohol intake > 2 units daily, Rheumatoid arthritis) for osteoporosis were included in the study. Written informed consent was taken for inclusion in the study. X-ray Pelvis was done in all patients. Singh index was calculated from plain X-rays (Figure 1 and 2). Radiographs were observed by an orthopaedic surgeon and two radiologists (3 observers). The data was analyzed with SPSS 16.0. Interclass Correlation Coefficient was measured. Mean and Standard deviation were used for quantitative variables.
Assessment of interobserver reliability of singh index for grading of osteoporosis

RESULTS

Mean age of the patients was 62.91 ± 9.6 SD with minimum age 43 and maximum age of 88. Mean BMI was 28.54 ± 6.62 SD with minimum of 16.80 and maximum of 59.80. Table 1 shows the grading of osteoporosis on plain X-rays by 3 different observers. Interobserver agreement was reached in 28/120 (23.33%) radiographs.

Single measurement was 43%. Overall agreement among the observers was 69.5% with p value .000, which is significant. 95% Confidence interval was 0.587 to 0.779. (Table 2). Cross-tabulation was done between observer 1 and 2. Kappa measurement was .186 with a p value of .000, which is significant (Table 3). Cross-tabulation was done between observer 1 and 3. Kappa measurement was .171 with a p value of .000, which is significant (Table 4). Cross-tabulation was done between observer 2 and 3. Kappa measurement was .261 with a p value of .000, which is significant (Table 5).

Table 1: Observers grading of X-rays

<table>
<thead>
<tr>
<th>Singh Index Grading</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Observer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>42</td>
<td>47</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>23</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2: Intraclass Correlation Coefficient

<table>
<thead>
<tr>
<th></th>
<th>Intraclass Correlation</th>
<th>95% Confidence Interval</th>
<th>F Test with True Value 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>Single Measures</td>
<td>.432</td>
<td>.321</td>
<td>.540</td>
</tr>
<tr>
<td>Average Measures</td>
<td>.695</td>
<td>.587</td>
<td>.779</td>
</tr>
</tbody>
</table>

Table 3: Observer1 * observer2 Crosstabulation

<table>
<thead>
<tr>
<th></th>
<th>Asymp.</th>
<th>Approx. T</th>
<th>Approx. Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure of Agreement</td>
<td>.186</td>
<td>3.837</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Observer1 * observer3 Crosstabulation

<table>
<thead>
<tr>
<th></th>
<th>Asymp.</th>
<th>Approx. T</th>
<th>Approx. Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure of Agreement</td>
<td>.171</td>
<td>3.782</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Observer2 * observer3 Crosstabulation

<table>
<thead>
<tr>
<th></th>
<th>Asymp.</th>
<th>Approx. T</th>
<th>Approx. Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure of Agreement</td>
<td>.261</td>
<td>5.518</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DISCUSSION

Osteoporosis means “porous bone.” Viewed under a microscope, healthy bone looks like a honeycomb. When osteoporosis occurs, the holes and spaces in the honeycomb are much larger than in healthy bone. Osteoporotic bones have lost density or mass and contain abnormal tissue structure. Studies
Assessment of interobserver reliability of singh index for grading of osteoporosis

suggest that approximately one in two women and up to one in four men age 50 and older will break a bone due to osteoporosis. By 2025, experts predict that osteoporosis will be responsible for approximately three million fractures and $25.3 billion in costs annually2.

Measurement of bone mass can predict the fracture risk. Skeletal mass can be measured by different techniques; Singh index, radiogrammetry, radiographic absorptiometry, quantitative computed tomography, ultrasound, energy absorptiometry (dual/single energy X-ray absorptiometry) and single or dual photon absorptiometry11.

The use of Singh’s Index, described in 1970, in evaluating bone mineral density is fraught with controversies8. It’s an observational tool and thus has an inherent drawback. Singh et al accepted the limitation in grading with occasional interobserver variability which, however, do not amount to more than one grade. There are several limitations of plain radiographs, e.g, soft tissue shadow in obese patient and poor radiographic technique11,12.

Our study showed that there is 69.5% interobserver agreement. The kappa value ranged from 0.171 to 0.261 with a slight to fair agreement according to Landis and Koch criteria14. Koot et al13 have argued that the results were disappointing for interobserver agreement (kappa values 0.01 to 0.54). Yoo MJ et al15 showed that kappa values for interobserver agreement ranged from 0.325 to 0.423, which did not reach good reliability overall.

According to other studies16-19 the inter-observer (and weighted) reliability k was 0.10 (95% CI: 0.00-0.20) and 0.26. Their conclusion was that inter-person reliability of the Singh index is poor; the index needs to be simplified to establish its reliability as a screening tool for osteoporosis. According to Supradeeptha Challa16, there is poor to fair inter-observer reliability; interobserver agreement was 0.21±0.2 (kappa value).

ACKNOWLEDGMENT

Sincere thanks to Dr. Iqbal Wahid for help in statistics.

CONCLUSION

Interobserver reliability of singh index for grading of osteoporosis is very useful method for grading of osteoporosis.

REFERENCES

15. Yoo MJ, Park HG, Yang SW. Reliability of Singh Index in Female Patients Older than 65 years with Proximal
Assessment of interobserver reliability of Singh index for grading of osteoporosis


Korean.


CONFLICT OF INTEREST: Authors declare no conflict of interest

GRANT SUPPORT AND FINANCIAL DISCLOSURE NIL

AUTHOR'S CONTRIBUTION

Following authors have made substantial contributions to the manuscript as under:

Bukhari SI: Formulation of Hypothesis, grading of radiographs and writing research article
Qadir RI: Collection of data
Khan S: Grading of radiographs and review of the article
Khan A: Grading of radiographs

Authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

The Journal of Medical Sciences, Peshawar is indexed with WHO IMEMR (World Health Organisation Index Medicus for Eastern Mediterranean Region) and can be accessed at the following URL.

http://www.who.int/EMRJorList/details.aspx?docn=4468