Reliability of the RUST Score for the Assessment of Union in Simple Diaphyseal Tibial Fractures

Sabir Ali¹, Ajai Singh²*, Avinash Agarwal³, Anit Parihar⁴, A A Mahdi⁵ and R. N. Srivastava²

¹Ph. D Scholar, Department of Orthopaedics, King George’s. Medical University, Lucknow, India
²Professor, Department of Orthopaedics, King George’s. Medical University, Lucknow, India
³Associate Professor, Department of Medicine, King George’s. Medical University, Lucknow, India
⁴Associate Professor, Department of Radiology, King George’s. Medical University, Lucknow, India
⁵Professor & Head, Department of Biochemistry, King George’s. Medical University, Lucknow, India

*Correspondence Info:
Prof. Ajai Singh,
Department of Orthopaedics,
King George’s. Medical University, Lucknow -226018, India
E-mail: as29762@gmail.com

Abstract
Objective: The aim of our study is to check the reliability of recently developed Radiographic Union Score for Tibial fractures (RUST Score) for determining the healing progression using X-rays of simple diaphyseal tibial fracture patients.

Methods: Total sixty-five patients with a simple diaphyseal tibial fracture are enrolled. Each of them are regularly follow-up at 6th, 10th, 16th, 20th, 24th post fracture weeks as per as their routine treatment protocol. To measure inter-observer reliability, an orthopaedician and a radiologist independently scored the bony union using plain X-ray by RUST scoring method. Further to measure, intra-observer reliability, each of the radiographs was reassessed and rescored at average of 06th weeks, from the initial day of scoring, without revealing patients identity or fracture duration. Finally to check the reliability of RUST score, Intraclass correlation coefficient (ICC) with 95% confidence intervals (CI) was measured.

Results: Overall agreement was statistically significant. The test-retest reliability was statistically significant, having value of 0.87, 0.95, 0.94, 0.98 and 0.97 respectively at each follow-up. The overall inter as well as intra-observer reliability of RUST score was collectively significant at each follow-up.

Conclusions: The RUST score shows substantial improvements in reliability and produces reproducible results, either scored by an orthopaedician or a radiologist.

Keywords: RUST Score, Interobserver reliability, Intraobserver reliability, Simple tibial fractures, diaphyseal fracture of tibia.

1. Introduction

The subcutaneous position of the tibial bone leads to higher incidence of fractures and due to less soft-tissue coverage, higher incidence of impaired healing occurs.¹ Only the rate of incidence of tibial nonunions is constitute 2-10% of all tibial fractures.²,3 Fracture impairment leads to prolonged disability, associated with substantial pain, put extra burden on the patient.⁴,⁵ According to National Centre for Health Statistics has reported that close to 500,000 tibia and fibula fractures occur each year in the United States. Determination of healing progression in diaphyseal tibia fractures is routinely based on clinic-radiological findings. Although many studies,⁶,⁷ has been revealed different grades or scores to evaluate fracture progression radiologically, but unfortunately due to lack of objectivity or having too many limitations, till now there is no such universally accepted standard method to evaluate fracture progression radiographically. But on the same, substantial agreement was seen with quantifying the number of cortices bridged by callus with or without a definite fracture line.⁸,⁹ Later on Whelan et al. (2010) developed a radiographic scaling system to determine fracture union.¹⁰

The purpose of this investigation was to assess the tibial fracture healing using Radiographic Union Score for Tibial fracture (RUST) score and to evaluate their inter and intra-observer reliability while evaluating by an orthopaedic surgeon and by a radiologist separately.

2. Materials and Methods

This was a longitudinal cohort study in with total 65 adult patients in the age group between 18 to 45 years with simple, traumatic diaphyseal fractures (AO A1, A2 & A3 as per as AO Muller classification) of both bones of leg managed conservatively were included.

After obtaining ethical clearance (Ref. Code: 57 E.C.M. IIAP/4) from Institutional ethical review committee, all 65 patients were included in this study from 2011 to 2013. Informed consent was taken. The demographic data was collected of all admitted enrolled patients. All were managed conservatively (reduction setting and above knee POP were applied under regional / general anesthesia). All patients were discharged after 24 – 48 hours with a standard advice written on discharge card. Further the radiological examination of patients was done at 6th, 10th, 16th, 20th, 24th weeks. Radiological examination was done using RUST score by taking standard plain radiographs of the AP and Lateral views. The radiographic scoring was done by an orthopaedic surgeon and by a radiologist separately and blindly.

In the RUST scoring system,¹⁰,¹¹ based on the assessment of fracture healing at each of the four cortices (i.e., medial and lateral cortices on the anteroposterior X-ray, anterior and posterior cortices on the lateral X-ray), each cortex receives a score of 1 point, if presence of fracture line with no callus; 2 points, if there is callus present but a fracture line is still visible; and 3 points, if there is bridging callus with no evidence of a fracture line. The individual cortical scores are added to give a total score of 4, which being the minimum score indicating that the fracture is definitely not healed and 12 being the maximum score indicating that the fracture is definitely healed. Radiographic fracture union was defined¹² when bony callus was evident on at least 3 cortices in standard AP and Lateral views and with RUST score ≥ 7.

The examiners were blinded to patient history, age of the fracture, and other clinical information related to their sets of radiographs. Interobserver reliability was evaluated by comparing the scores given by the observers separately (Orthopaedician & Radiologist) at the initial viewing of the radiographs. To calculate intraobserver reliability, each observer was asked for the RUST scoring of the radiographs again at an average of 06 weeks (range, 04-08 weeks).
weeks) after the initial assessment. Intraclass correlation coefficients (ICC) with 95% confidence intervals (CI) were used to measure agreement in the observer’s RUST scores.

2.1 Statistical analysis

The data were entered in Microsoft Excel and statistical analysis was performed using SPSS software for Windows program (15.0 version). The continuous variables were evaluated with mean (±SD) or range value when required. For comparison of the means between the observers, Intraclass correlation coefficient (ICC) with its 95% confidence interval was used. A p value less than 0.05 or 0.001 were regarded as significant.

3. Results

Among the 65 patients, total 325 radiographs were obtained during their treatment follow-up. The table-1 shows the range of RUST score between 4 to 12, with their mean score ± SD (also see fig-1). The test-retest reliability was statistically significant, having value of 0.87, 0.91, 0.93, 0.98 and 0.97 respectively at each follow-up. The interobserver reliability of RUST score among the two different observer (an orthopaedician and a radiologist) was collectively significant at each follow-up (ICC: 0.67, 0.91, 0.93, 0.98 and 0.96 respectively), as describe in table-2. To determine the intraobserver reliability of RUST scores within the observer, the reviewers evaluated again the same radiographs at an average of 8 weeks (range, 6–12 weeks) from the original session. The overall intraobserver ICC for the group of reviewers was again statistically significant at each follow-up (ICC: 0.87, 0.89, 0.90, 0.91 and 0.94 respectively).

Table-1: Test retest reliability of the two raters at different time intervals of RUST score

<table>
<thead>
<tr>
<th>Follow-up period</th>
<th>Observer I (Orthopaedician)</th>
<th>Observer II (Radiologist)</th>
<th>Test-retest reliability (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th week</td>
<td>5.96±0.38 (4-7)</td>
<td>6.11±0.32 (4-7)</td>
<td>0.87, p=0.0001*</td>
</tr>
<tr>
<td>10th week</td>
<td>7.28±1.31 (4-8)</td>
<td>7.37±1.35 (4-9)</td>
<td>0.95, p=0.0001*</td>
</tr>
<tr>
<td>16th week</td>
<td>7.88±1.39 (4-10)</td>
<td>8.02±1.42 (4-10)</td>
<td>0.94, p=0.0001*</td>
</tr>
<tr>
<td>20th week</td>
<td>9.28±1.84 (5-11)</td>
<td>9.35±1.84 (5-11)</td>
<td>0.98, p=0.0001*</td>
</tr>
<tr>
<td>24th week</td>
<td>10.16±2.08 (5-12)</td>
<td>10.31±2.07 (5-12)</td>
<td>0.97, p=0.0001*</td>
</tr>
</tbody>
</table>

Table-2: Inter-observer and intra-observer reliability

<table>
<thead>
<tr>
<th>Follow-up period</th>
<th>Inter-observer reliability</th>
<th>Intra-observer reliability (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6th week</td>
<td>0.87 (0.73-0.92), p=0.0001*</td>
<td>0.87 (0.76-0.91), p=0.0001*</td>
</tr>
<tr>
<td>10th week</td>
<td>0.91 (0.79-0.96), p=0.0001*</td>
<td>0.89 (0.78-0.90), p=0.0001*</td>
</tr>
<tr>
<td>16th week</td>
<td>0.93 (0.89-0.95), p=0.0001*</td>
<td>0.90 (0.76-0.98), p=0.0001*</td>
</tr>
<tr>
<td>20th week</td>
<td>0.98 (0.96-0.98), p=0.0001*</td>
<td>0.91 (0.81-0.99), p=0.0001*</td>
</tr>
<tr>
<td>24th week</td>
<td>0.96 (0.94-0.98), p=0.0001*</td>
<td>0.94 (0.87-0.98), p=0.0001*</td>
</tr>
</tbody>
</table>

Fig-1: Distribution of the RUST scores among 325 radiographs given by the two different observers.

4. Discussion

Till date, there is no reliable method is available to measure the bony healing progression on plain X-rays after fracture. Commonly in practice, clinico-radiological parameters are used to measure the bony healing progression, but mostly we are not able quantify the progression of healing radiologically, especially when the clinical parameters are confusing.1,2 However, other techniques rather than X-rays, like computed tomography, radionuclide imaging3-9,10,11 are also available to overcome the above problem as above said, but unfortunately these techniques are very cost effective and are not available in every hospital.

Previously, in few studies, substantial agreement was seen with quantifying the number of cortices bridged by callus and/or with a definite fracture line.12-16 So, based on above parameters Whelan et al. (2010) developed a radiographic scaling system to determine fracture union in tibial fractures, pronounced as RUST score, and also try to validate this score in their own study.17 Later on, Bhandari et al. (2011) also find the substantial reliability of this scoring system.18 Moreover, Bhandari et al. (2013), the same principle of the RUST scoring were applied to assess the hip fracture, also find the substantial ICC agreement and termed this scoring system as Radiographic Union Score for Hip fracture (RUSH).19

In the present study, we were test the reliability of RUST scoring system. The inter-observer reliability was tested by the observed RUST score given by expert separately (an orthopaedican and a radiologist). The intra-observer reliability was evaluated by observing the same radiographs again at an average of 06 weeks (range, 04-08 weeks) from the initial day of scoring by the observers. The study results confirmed the substantial reliability of RUST score in clinical application that suggests that the cortical continuity, as determined by quantifying the number of cortices with bridging callus or with a visible fracture line, was more reliable to assess, than any of the other methods evaluated. Also in respect to previous studies, substantial reliability has been found while analysing the tibial fracture healing progression, despite of treatment methods, which may be either operative or conservative.

However, there were some limitations to this study such as, there were a limited number of observers used for scoring the healing status using RUST score, which would needed to be further verified with a larger sample size with observers of different fields. But, despite of these limitations, the present study result shows the initial steps of a full validation of the RUST scoring system. And because there is currently no “gold standard” tool for quantifying the fracture healing radiologically, the RUST score may use as a complimentary tool for the clinicians, to assess fracture healing.
5. Conclusion

In conclusion, this work has presented the initial step in the development of the RUST score to grade tibial fracture healing, treated with non-operative method. Their high degree of reliability and reproducibility of this radiographic grading method and as because there is currently no “gold standard,” method to evaluating the fracture healing, the RUST scoring system will be used as a potential tool in clinical practice.

References