The reliability and validity of the Unified Classification System of periprosthetic femoral fractures after hip arthroplasty

Jie-Feng Huang, Jun-Jie Chen, Jian-Jian Shen, Wen-Xi Du, Fu-Cun Liu, Pei-Jian Tong

From Department of Orthopedics, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China

The Unified Classification System has expanded and updated the Vancouver Classification System and applied treatment principles to all periprosthetic fractures. This study assessed the reliability and validity of the Unified Classification System for femur fractures after hip arthroplasty. Thirty-eight radiographs were finally evaluated by 6 observers, 3 consultants and 3 trainees. Each observer read the radiographs on 2 separate occasions, independently, at least 2 weeks apart. Interobserver and intraobserver agreement and validity were analyzed, using weighted \( \kappa \) statistics. The mean \( \kappa \) value for interobserver agreement was found to be 0.849 (0.838-0.871) for consultants (almost perfect agreement) and 0.761 (0.707-0.836) for the trainees (substantial agreement). Intraobserver \( \kappa \) values ranged from 0.740 to 0.903, showing substantial to almost perfect agreement.

Validity analysis of 23 type B cases revealed 79.710% agreement within B1, B2, and B3 subgroups with a mean \( \kappa \) value of 0.694 (0.670-0.741) (substantial agreement).

Keywords: arthroplasty; femoral periprosthetic fracture; Unified Classification System; reliability; validity.

INTRODUCTION

Periprosthetic femoral fracture after total hip arthroplasty (THA) is an increasing clinical problem (1,3,6,7,9,12,17,19,21,24). They were identified as the third most frequent reason for revision following primary THA, after aseptic loosening and sepsis (9,18). Periprosthetic fractures are difficult to manage and may have a poor outcome. The treatment of periprosthetic fractures of the femur after THA remains a surgical challenge (3,6,11,16,17,19). There are many variables that need to be considered with each fracture, including fracture site and pattern, implant stability and the surrounding bone quality, complex with the more general factors such as patient’s age and functional demands (3,16,19).

Various classifications have been described, some of which depend on the site of the fracture (14,20), and others on the pattern of the fracture or relationship to the stability of the implant (2,5,13,14,20,22). These historical classification

Jie-Feng Huang and Jun-Jie Chen are co-first authors.

Jie-Feng Huang, MD.
Jun-Jie Chen, MD.
Jian-Jian Shen, MD.
Wen-Xi Du, MD.
Fu-Cun Liu, MD.
Pei-Jian Tong, MD.

Department of Orthopedics, The First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China.

Department of Orthopedics, Affiliated Cixi Hospital of Wenzhou Medical College, Cixi, China.

Correspondence: Pei-Jian Tong, Department of Orthopedics, The First Affiliated Hospital of Zhejiang Chinese Medical University, 54 Youdian Road, Shangcheng District, Hangzhou, 310006, China. E-mail: tpj637963@163.com


No benefits or funds were received in support of this study. The authors report no conflict of interests.
systems (2,5,13,14,20,22) for periprosthetic fractures of the femur have been superseded by the Vancouver Classification System (10) which has become universally accepted. The reliability and validity of the Vancouver Classification System of periprosthetic femoral fractures after hip replacement had been confirmed (3,19,21). In 2014, Duncan and Haddad (9) introduced the Unified Classification System in order to expand and update the Vancouver Classification System and apply treatment principles to all periprosthetic fractures. When applied to the femur, the Unified Classification System incorporates the previous Vancouver Classification System but is expanded to include two new fracture pattern, type D and E (9,24) (Table I).

<table>
<thead>
<tr>
<th>Type</th>
<th>Subtype</th>
<th>Fracture description</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A</td>
<td>Fracture in trochanteric region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>Fracture of the greater trochanter</td>
<td>Conservative or cable wires</td>
<td></td>
</tr>
<tr>
<td>AL</td>
<td>Fracture of the lesser trochanter</td>
<td>Conservative or cable wires</td>
<td></td>
</tr>
<tr>
<td>Type B</td>
<td>Fracture around stem or just below it</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>Well-fixed stem</td>
<td>ORIF</td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>Loose stem with good proximal bone stock</td>
<td>Revision THA</td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>Loose stem with poor-quality bone stock</td>
<td>Revision THA</td>
<td></td>
</tr>
<tr>
<td>Type C</td>
<td>Fracture occurring well below the tip of the stem</td>
<td>ORIF</td>
<td></td>
</tr>
<tr>
<td>Type D</td>
<td>Fracture of the femoral shaft between well-fixed hip and knee replacements (Type C for each)</td>
<td>ORIF</td>
<td></td>
</tr>
<tr>
<td>Type E</td>
<td>Both the femur and acetabulum fractures after THA</td>
<td>Logical individual treatment</td>
<td></td>
</tr>
<tr>
<td>Type F</td>
<td>Does not apply to the femur</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ORIF: open reduction and internal fixation; THA: total hip arthroplasty.

Figure 1. — Anteroposterior radiograph of the left hip and proximal femur showing a B1 type periprosthetic fracture. The stem is well-fixed.
For any classification system to be useful, it should not only guide treatment and help in the decision-making process, but also be reliable and valid. The objective of this study is to independently assess the interobserver and intraobserver reliability and validity of the Unified Classification System, in our hospital.

MATERIALS AND METHODS

The study was approved by IRB (Institutional Review Board) of the hospital. We retrospectively analyzed a cohort of 41 consecutive patients who had sustained a periprosthetic fracture of the hip between January 2000 and October 2014. Each subtype applied to the femur of the Unified Classification System was represented at least once. The full extent of the fracture and prosthesis needed to be clearly visible on the available radiograph in at least 2 planes for inclusion in the study. Suitability for inclusion was ultimately judged by the first author (J.F. Huang), who was independent of all observers and not involved in the judgment process. Only patients with a pre-operative radiological diagnosis and complete operative documentation of the type of fracture were included. Cases were excluded if any of the following were present: migration or fracture of an old nonunited trochanteric osteotomy; revision Exeter component (impaction cancellous allograft), and fracture of the femoral prosthesis. Three cases were excluded. Thirty-eight patients were tested in our study. There were 15 men and 23 women with a mean age of 75.7 years (56.4 to 89.2).

The study was done in 2 parts. Part 1 of the study involved the analysis of interobserver and intraobserver reliability of the Unified Classification System. Six observers were involved in this study, including 3 consultants and 3 trainee surgeons (registrars). All participants reviewed the Unified Classification System before the study. Radiographs of all 38 patients were classified by all 6 participants independently, on 2 separate occasions, at least 2 weeks apart. Part 2 of the study assessed of validity of the classification by comparing the subgroup classification of type B fractures with the intraoperative findings, as retrieved from the operative notes. The senior consultant’s classification results were used for validation purpose and were compared with the intraoperative findings.
cessed for the first reading among all the possible pairs of consultants and trainees (Table II). The mean $\kappa$ value was 0.849 (0.838-0.871) for consultants (almost perfect agreement) and 0.761 (0.707-0.836) for the trainees (substantial agreement), more so among the consultants than the trainees. Intraobserver $\kappa$ values ranged from 0.740 to 0.903, showing substantial to almost perfect agreement. The mean $\kappa$ value was found to be 0.891 for consultants and 0.783 for trainee surgeons.

Validity analysis involved 23 type B cases and, when compared with the operative findings, showed 79.710% agreement within the B1, B2, and B3 subgroups, with a mean $\kappa$ value of 0.694 (0.670-0.741) (substantial agreement) and a standard error of 0.121 (Table III).

**DISCUSSION**

Periprosthetic fractures are difficult to manage and may have a poor outcome. Historically, the treatment of periprosthetic femoral fractures has been associated with a high frequency of complications and reoperations (12,17). The treatment of periprosthetic fractures of the femoral after THA remains a surgical challenge (3,6,11,16,17,19). Consequently, a robust, reliable classification which can establish the principles of management for surgical treatment of periprosthetic fractures is needed. Several authors (2,5,13,14,20,22) have written about the need for a standardized classification and an adequate treatment algorithm for late femoral periprosthetic fractures. Parrish et al (1964) (20) classified periprosthetic fractures in relation to their location, as: trochanteric; proximal; middle; or distal third fractures. Johansson et al (1981) (14) classified fractures in relation to the implant as proximal to the tip of the stem; extending beyond the tip of the stem; or entirely beyond the tip. Bethea et al (1982) (2) expanded on these classifications to incorporate position and pattern: type A occurring distal to the tip of the stem; type B as spiral fractures around the stem; and type C as comminuted fractures around the stem. Cooke et al (1988) (5) described four types of periprosthetic fracture and emphasized the need for early revision in the comminuted type 1 fracture, while the type 2 transverse fracture around the stem

**RESULTS**

Thirty-eight cases were finally included in this study. Interobserver agreement was separately ac-

---

*Fig. 4.* Anteroposterior radiograph of the right hip and proximal femur showing a type E periprosthetic fracture of the proximal femur with an associated fracture of the acetabulum.
could be managed either conservatively or by revision. They proposed that fractures at or distal to the tip of the stem (type 3 and 4 respectively) required internal fixation. Jensen et al (1988) (13) emphasized the importance of implant stability at the time of injury as a predictor for revision of the prosthesis. Fractures were classified as type 1 when located around the proximal two-thirds of the femoral component, type 2 when the fracture extended proximal and distally from the stem tip, and type 3 when the fracture extended distal to the tip of the stem. Roffman et al (1989) (22) described on the basis of the predicted stability of the implant of ipsilateral femur fractures after THA.

These historical classification systems (2,5,13,14, 20,22) for periprosthetic fractures of the femur have been superseded by the Vancouver Classification System (Duncan et al. 1995) (10) which has become universally accepted. The reliability and validity of the Vancouver Classification System of periprosthetic femoral fractures after hip replacement had been confirmed (3,19,21). In 2014, Duncan and Haddad (9) introduced the Unified Classification System. Its purpose is threefold: to expand and update the Vancouver Classification System by the addition of three subtypes which are increasing in prevalence; to apply these groupings and treatment principles to all periprosthetic fractures regardless of the bone that is broken and the joint involved; and lastly, to propose a common language so that we may communicate with clarity and simplicity, regardless of our native tongue, preferred dialect, and subspecialty limb or joint of interest (8). When applied to the femur, the Unified Classification System incorporates the previous Vancouver Classification System but is expanded to include two new fracture pattern, type D and E (9,24).

The reliability and validity of the Unified Classification System for periprosthetic fracture of the pelvis and femur around a total hip replacement have been assessed by Vioreanu et al in 2014 (24). They tested the inter- and intra-observer agreement for the Unified Classification System as applied to the pelvis and femur using 20 examples of periprosthetic fracture in 17 patients. When applied to the femur, interobserver reliability showed $\kappa$ values of 0.805 for consultants, 0.732 for trainees (24). Our study reconfirmed the finding of the prior study with comparable result and showed intraobserver and interobserver agreement.

The study has shown that the Unified Classification System for periprosthetic femoral fracture is reliable. Intraobserver and interobserver agreement was judged to be substantial to almost perfect by $\kappa$ analysis.

In this study, the validity assessment for type B fracture subgroups showed an observed agreement of 79.710% with a $\kappa$ value of 0.694, indicating substantial agreement. However, like 3 prior stud-

<table>
<thead>
<tr>
<th>Groups</th>
<th>$\kappa$ Value</th>
<th>% Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultant A-B</td>
<td>0.838</td>
<td>86.842%</td>
</tr>
<tr>
<td>Consultant A-C</td>
<td>0.871</td>
<td>89.474%</td>
</tr>
<tr>
<td>Consultant B-C</td>
<td>0.838</td>
<td>86.842%</td>
</tr>
<tr>
<td>Mean $\kappa$ value</td>
<td>0.849</td>
<td>(95% confidence interval, 0.802-0.896)</td>
</tr>
<tr>
<td>Trainee A-B</td>
<td>0.836</td>
<td>86.842%</td>
</tr>
<tr>
<td>Trainee A-C</td>
<td>0.707</td>
<td>76.316%</td>
</tr>
<tr>
<td>Trainee B-C</td>
<td>0.740</td>
<td>78.947%</td>
</tr>
<tr>
<td>Mean $\kappa$ value</td>
<td>0.761</td>
<td>(95% confidence interval, 0.595-0.928)</td>
</tr>
</tbody>
</table>
ies (3,19,21), this one also failed to achieve perfect agreement. This indicates that implant stability can be difficult to determine from radiographs alone. This finding has an important bearing on the consequences, as choosing the wrong treatment can lead to early failure. For example, Van Houwelingen et al (2011) (23) and Capello et al (2014) (4) reported pseudo A(LT) periprosthetic fracture: it’s really a B2 (Fig. 5). The new fracture pattern was not described in the original Vancouver Classification System and the Unified Classification System.

The goal of any classification system is ultimately to improve patient outcome. To be able to accomplish this goal, the classification must guide treatment, be reliable, and valid. The Unified Classification System of periprosthetic fractures for femur has been shown to be reliable and valid in this study, but remains need to be refined in the future.

REFERENCES


Fig. 5. — Anteroposterior radiograph of the left hip and proximal femur showing a pseudo A type periprosthetic fracture, but it’s really a B2 type periprosthetic fracture. Fractures of the lesser trochanter as well as the greater trochanter, the stem is loose but there is adequate bone stock.


