REPRODUCIBILITY OF TILE'S CLASSIFICATION OF ACETABULAR FRACTURES

Marcos Almeida Matos[1], Adriano Moura Costa de Viveiros[2], Bruno Garcia Barreto[2], Rafael Fernandes Dias Pires[2]

SUMMARY

Properly classifying acetabular fractures is crucial for a good preoperative planning and for an efficient surgical reduction. However, in order to accomplish its objectives, any classification system must be simple and reproducible. The objective of this article is to assess inter-observer reproducibility of Tile's classification concerning acetabular fractures. Thirty X-ray images of 10 acetabular fractures at Judet planes were used and assessed by 10 observers, being five hip surgery experts and five 3rd-grade orthopaedic residents. The glo-

bal consistency achieved was 72.44% to Kappa (K) = 0.52 (0.48 among resident doctors and 0.57 among experts). It was concluded that the Tile's classification of acetabular fractures reveals a moderate inter-observer consistency, with no statistically significant difference between resident doctors and experts.

Keywords: *Reproducibility of results; classification; fracture; acetabulum.*

INTRODUCTION

Acetabular fractures represent an increasingly important topic in modern orthopaedics. Its increased incidence is related to the progressive frequency of high energy trauma [1,2]. Appropriate treatment with anatomical restoration of those fractures is paramount for assuring a good prognosis and for preventing severe and disabling sequels [1-6].

Approach determination is directly influenced by deviation, stability and kind of fracture. Tile[7], in 1980, classified those fractures into 3 types, each one presenting specific subdivisions, based on the involvement of acetabular walls and columns. According to him, the accurate knowledge of involved bone structures is crucial for choosing the most appropriate access port, providing the elements for anatomical reduction.

To precisely classify acetabular injuries is crucial for a good preoperative planning and for an efficient surgical reduction[1,5,6,8]. However, to be able to reach to its objectives, any classification system must have as a key point the exact description of the injury [9] in a simple and reproducible manner, with agreeable results between different examiners [10,11].

The objective of this article is to evaluate the inter-observer reproducibility of the Tile's classification for acetabular fractures and to determine the differences between expert observers and trainee observers.

MATERIALS AND METHODS

Pelvic X-ray images of 10 acetabular fractures at three pla-

nes (AP and Judet oblique), totaling 30 images, randomly selected among patients diagnosed with acetabular fracture, who were enrolled at the orthopaedics service of the Santa Casa de Misericórdia da Bahia (Hospital Santa Izabel), in the period comprehending March 2003 to November 2004. The study was approved by the Committee on Ethics in Research of the hospital.

X-ray films were taken with a Sony Cyber-Shot 3.2 megapixels digital camera and the images were deployed in a CD-ROM. The authors reviewed the images in order to assure quality of visualization and the correct positioning of planes.

Ten observers, composed of five hip-expert orthopaedic surgeons and five third grade resident doctors in orthopaedics, have been chosen to classify fractures according to Tile's classification.

Each observer received the CD-ROM with images and a booklet containing a descriptive scheme of kinds of fractures according to the classification under study, including figures of each type and subtype. Combined to the booklet, a questionnaire was delivered, in which the observers should classify ten fractures. All observers received guidance as to take their times to analyze images in detail and consistently to the material described on the booklet, returning the questionnaire duly filled in. In case an observer considered a fracture was not in accordance with classification, that item should be left blank.

Data gathered were distributed and presented in tables in the form of descriptive statistics. The evaluation of inter-

observer agreement was performed by obtaining a Kappa (K) index stratified by Landis and Koch[12]. For continuous variables comparison, the t test was employed, with a significance level of 0.05. Data were marked with an asterisk (*) for significant cases, and with a sharp mark (#) for non-significant cases.

RESULTS

None of the observers regarded image quality as insufficient to perform the classification task. Only one observer was not able to classify fracture number three, with all remaining fractures being classified by the 10 observers. The agreements achieved are shown on Table 1.

The disagreement between observers was further divided into nine unconformities, ranging from zero to eight. Each level consisted of a difference between a classification by observer A against observer B, and so on. Tile's classification was regarded as ordinal distinct distribution and each subtype differing from subsequent subtype in one level. For example, the subtype II-A differs from subtype II-B in one level and from subtype II-C in two levels. When two observers similarly classified a fracture within a given subtype, then we regarded it as level zero of disagreement. Disagreement frequency was distributed according to the nine possible levels, as shown on Table 2. A gradual decreasing pattern of disagreement frequency was seen from level zero to level eight.

On Table 3, the analysis is distributed according to the average, standard deviation, and standard error for disagreement average, with those data being obtained from the overall classification by each observer.

DISCUSSION

Reproducibility of classifications in orthopaedics is still challenging among experts. Gusmão et al.[11], by analyzing the Garnden's classification of femoral neck fractures, found a poor inter-observer agreement (K = 0.32). Sidor et al.[13], by assessing Neer's classification for proximal humerus, identified a poor-to-moderate agreement (K = 0.48 to 0.52). Thomsen et al.[14] showed a poor reproducibility for ankle fractures.

Table 1 – Percentage of global agreement and Kappa index for Tile´s classification.

Agreement/Kappa	Experts	Resident doctors	Total
Agreement achieved Kappa (C.I.)	76,87%# 0,57(0,29 – 0,84)#	66,88%# 0,48(0,20 – 0,79)#	72,44% 0,52(0,30 – 0,80)

Note: C.I. = Confidence Interval.

Table 2 – Distribution by frequency of discrepant classifications for experts and resident doctors.

Discrepancy level	Experts	Resident doctors	Overall Total (%)
0	26	18	106(23,56%)
1	37	23	119(26,44%)
2	10	21	71(15,78%)
3	9	13	48(10,67%)
4	6	7	31(6,89%)
5	3	8	24(5,33%)
6	3	2	22(4,89%)
7	4	6	21(4,67%)
8	2	2	8(1,78%)
Total	100	100	450(100%)

Note: For each distribution (expert, resident doctor and total) the total of possible combinations was used as a denominator, i.e., five experts combined to each other for each fracture (10 possibilities), multiplying by total fractures to evaluate (10), resulting in 100 combinations. For overall total, we have 10 observers for each fracture (10 fractures) in a total of 450 combinations.

Table 3 – Distribution by frequency of errors on Tile´s classification.

	Experts	Resident Doctors	Total
Average	1,85#	2,42#	2,14
Standard deviation	2,05	2,11	2,11
Standard error	0,21	0,21	0,21

We didn't find any article in literature assessing Tile's classification reproducibility for acetabular fractures. The agreement on the classification by Letournel et al.[15] for those fractures, which is based on the same principles as the Tile's classification, presents controversial results in literature.

Visutipol et al.[16] conducted a study in which X-ray images at AP and Judet oblique planes of 20 acetabular fractures evaluated by means of Letournel's classification were assessed by nine orthopaedic surgeons, not informing their level of expertise. Inter-observers reproducibility found was poor, with K = 0.24. Sancineto et al.[17], also using the Letournel's classification, detected poor reproducibility when assessing 30 fractures by six observers, being three resident doctors in orthopaedics and three pelvic surgery experts.

The result of this study reveals a moderate global agreement (K = 0.52), achieving a higher level of inter-observer agreement than the above mentioned studies using the Letournel's classification, but a little below the results found by Beaulé et al.[10]. That author used 65 acetabular fractures assessed by nine observers, being three hip surgeons trained by Letournel, which fact may have influenced on such good reproducibility achieved.

We also noticed that the disagreement between observers was more frequent in only one or two levels of discrepancy (Standard error, Table 3). None of the studies mentioned above evaluating the reproducibility of acetabular fractures classification has addressed the frequency of discrepancy levels. We believe that this fact provides clinical applicability to the classification, because discrepancies at low levels show value on selecting surgical access and prognosis. For example, even when fractures are classified as I-A (posterior column) by an observer and as I-A (posterior wall) by another one, they will have the same surgical access chosen, enabling the correct surgical treatment for both. Fractures classified as type III, even when a subtype discrepancy exists, are fractures of the two columns, with a similar prognosis, but possibly worse than types I and II (which would represent major discrepancies).

We believe that the moderate agreement found in this study,

among other factors, is basically due to the complex acetabular architecture combined to the uncountable types and subtypes described by Tile. Previous studies show improved reproducibility when methods for simplifying a classification are used. Petrisor et al.[18] concluded, in a study published in 2003, that Letournel's classification reproducibility is significantly improved when using only parameters addressing the involvement of 6 X-ray lines on AP plane, instead of the 10 subtypes proposed by the classification. Gusmão et al.[11] also identified a higher Kappa index by reducing Garden's classification to two types, dislocated and non-dislocated. Another relevant factor is that, despite of the increased incidence, acetabular fractures are still relatively uncommon on orthopaedic practice, making difficult to accrue a vast experience with this kind of injury.

Computed tomography (CT), potentially able to provide a better evaluation in acetabular fractures in 3-d, represents an additional possibility to better understand those complex hip injuries. This test adds important supplementary data to conventional X-ray images, such as evaluation of small degrees of pre- and postoperative deviations, injuries in an isolate column with little deviation, fragments and intra-joint impactions [6,10,19,20,21] among others. In spite of this, its role as an increment to improve agreement on those fractures classification, by adding the CT in the assessment, particularly between observers with little experience with acetabular surgery. Sancineto et al.[17] reported also an increased reproducibility with the use of CT. This difference does not agree with the findings by Visutipol et al.[16] and Beaulé et al.[10]: both were not able to identify an agreement improvement when CT was added to conventional X-ray images.

We didn't find any statistically significant difference on inter-observer reproducibility for Tile's classification between the resident doctors group and the hip experts group. This result is consistent to those obtained by Sancineto et al.[17], but contrasts to the majority of published studies[10,18,22] on the matter, which show a significant increase of reproducibility between more experienced observers, trained in hip surgery. We believe that the agreement achieved by resident doctors demonstrates the relatively high frequency of those fractures in our service, which, for being a tertiary care Hospital, absorbs acetabular fractures from other institutions, enabling resident doctors in orthopaedics to become familiar to such injuries. The complex interpretation and x-ray classification of those fractures may be a difficulty factor even for more experienced doctors.

CONCLUSION

We conclude that the Tile's classification for acetabular fractures presents a moderate inter-observer reproducibility (K = 0.52). We found no statistically significant difference on reproducibility between resident doctors and hip surgery experts. This classification shows a low level of discrepancy (0.15) and, thus, despite of the moderate reproducibility, it can be used relatively safely in daily clinical practice.

REFERENCES

1. Aristide RSA, Honda E, Polesello G, Fernandez MS. Fratura em "T" do acetábulo: análise de 45 casos. Rev Bras Ortop. 1996; 31:919-24.
2. Alonso JE, Volgas DA, Giordano V, Stannard JM. A review of the treatment of hip dislocations associated with acetabular fractures. Clin Orthop Relat Res. 2000; 377: 32-43.
3. Köberle G, Miranda JB, Yamanaka E, Grava ALS, Belangero WD. Fraturas do acetábulo: estudo dos resultados clínicos do tratamento cirúrgico precoce e tardio. Rev Bras Ortop. 1993; 28:361-6.
4. Moed BR, Carr SEW, Watson JT. Results of operative treatment of fractures of the posterior wall of the acetabulum. J Bone Joint Surg Am. 2002; 84:752-8.
5. Knop T, Silva LHP, Laghi R. Fraturas acetabulares: resultados de tratamento cirúrgico. Rev Bras Ortop. 1996; 31:825-30.
6. Mears DC, Velyvis JH, Chang C. Displaced acetabular fractures managed operatively: indicators of outcome. Clin Orthop Relat Res. 2003; 407:173-86.
7. Tile T. Fractures of The Acetabulum. Orthop Clin North Am. 1980; 11:481-506.
8. Honda H, Aristide RSA, Polesello G, Lima LEA, Favaro T, Laino VEA. Fratura do acetábulo que acomete a coluna anterior: Alternativa de osteossíntese. Rev Bras Ortop. 1998; 33:759-64.
9. Murphy WM. "Classificação de Fraturas: significado biológico". In: Princípios AO do Tratamento de Fraturas. Porto Alegre-RS : AO Publishing; 2000. p. 45-6.
10. Beaulé PE, Dorey FJ, Matta J. Letournel Classification for Acetabular Fractures: Assessment of Interobserver and Intraobserver Reliability. J Bone Joint Surg Am. 2003; 85: 1704-9.
11. Gusmão PDF, Mothes FC, Rubin LA, Gonçalves RZ, Telöken MA, Schwartsmann CR. Avaliação da reprodutibilidade da classificação de Garden para fraturas do colo femural. Rev Bras Ortop. 2002; 37:381-6.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33:159-74.
13. Sidor ML, Zuckerman JD, Lyon T, Koayal K, Cuomo F, Schoenberg N. The Neer Classification System for Proximal Humeral Fractures: an assessment of interobserver reliability and intraobserver reproducibility. J Bone Joint Surg Am. 1993; 75:1745-50.
14. Thomsen NOB, Overgaard S, Olsen LH, Hansen H, Nielsen ST. Observer Variation in the Radiographic Classification of Ankle Fractures. J Bone Joint Surg Br. 1991; 73: 676-8.
15. Judet R, Judet J, Letournel E. Fractures of the Acetabulum: classification and surgical approaches for open reduction. J Bone Joint Surg Am. 1964; 46:1615-46.
16. Visutipol B, Chobtangsin P, Ketmalasiri B, Pattarabanjird N, Varodompun N. Evaluation of letournel and judet classification of acetabular fracture with plain radiographs and three-dimensional computerized tomographic scan. J Orthop Surg. 2000; 8:33-37.
17. Sancineto C, Barla J, Slullitel G, Della Valle, AG; Comba F; Bongiovanni S et al. Reproductibilidad de la clasificacion de Letournel para las fracturas acetabulares: estudio sobre radiografias y tomografia computadorizada tridimensional. Rev Asoc Argent Ortop Traumatol. 2004; 69:59-63.
18. Petrisor BA, Bhandari M, Orr RD, Mandel S, Kwok DC, Schemitsch EH. Improving reliability in the classification of fractures of the acetabulum. Arch Orthop Trauma Surg. 2003; 123:228-33.
19. Berton RM, Carr SE, Gruson KI, Watson JT, Craig JG. Computed tomographic assessment of fractures of the posterior wall of the acetabulum after operative treatment. J Bone Joint Surg Am. 2003; 85:512-22.
20. Borrelli JJ, Goldfarb C, Catalano L, Evanoff BA. Assessment of articular fragment displacement in acetabular fractures: a comparision of computerized tomography and plain radiographs. J Orthop Trauma. 2002 16:449-56.
21. MacLeod MD, Mears DC. The resolution of acetabular fracture displacement by plain radiography. J Bone Joint Surg Br. 1998; 80:3.
22. Hüfner T, Pohlemann T, Gänsslen A, Assassi P, Prokop M, Tscherne H. Classification of acetabular fractures. a systematic analysis of the relevance of computed tomography. J Orthop Trauma. 1999; 13:391-2.