

ORIGINAL ARTICLE

FRACTURES OF THE NECK OF THE TALUS: EVALUATION OF REPRODUCIBILITY OF HAWKINS' CLASSIFICATION

MADSON LOBATO DRUMMOND FILHO, MARCOS AURÉLIO VERZANI, ANDRÉ FRAZÃO ROSA, CIRO JABUR PIMENTA, JEAN GRYNWALD, ALBERTO CLIQUET JUNIOR

ABSTRACT

Objective: To evaluate the intraobserver and interobserver reproducibility of Hawkins' classification for fractures of the neck of the talus. **Methods:** 20 random cases of fracture of the talus were selected, to be defined according to the classification of types by eight orthopedic surgeons, 13 orthopedic residents and 15 radiology residents. **Results:** Using the statistical test of Landis and Koch, measurements of 0.627 and 0.668 were obtained in the

first and second evaluations, respectively. These values define a satisfactory agreement for Hawkins' classification. **Conclusion:** We conclude that this classification is reproducible between observers, with better values for the more experienced observers. **Level of Evidence I, Study Diagnostic - Investigating a diagnostic test.**

Keywords: Fractures, bone/classification. Fractures, bone/complications. Talus.

Citation: Drummond Filho ML, Verzani MA, Rosa Af, Pimenta CJ, Grynwald J, Cliquet Junior A. Fractures of the neck of the talus: evaluation of reproducibility of Hawkins' classification. *Acta Ortop Bras.* [online]. 2012;20(3): 170-3. Available from URL: <http://www.scielo.br/aob>.

INTRODUCTION

Fractures of the neck of the talus constitute a challenge to the orthopedic surgeon. They are recognized by the considerable frequency of unsatisfactory results, with a high incidence of severe complications, such as osteonecrosis.¹

Hawkins' classification is based on the talar neck fracture according to deviation and congruity with the subtalar (talus and calcaneus) and tibiotalar joints.

This classification initially described as three types and subsequently modified to include a fourth type according to observations made by Canale and Kelly and Pantazopoulos et al.² is the following:

Type I: Vertical fracture of the neck without deviation;

Type II: Fracture of the neck with subluxation or dislocation of the subtalar joint (the ankle joint continues aligned);

Type III: Fracture of the neck with tibiotalar and subtalar dislocation;

Type IV: Fracture of the neck with talonavicular dislocation.

One of the most feared complications of the talar fracture, which is osteonecrosis, is closely correlated to Hawkins' classification.

Type I fractures range from 0-13%, type II from 20-50% and type III from 83 to 100%, with the average between 21 and 58%, which means that this complication is also the most common in this type of fracture.²

Nowadays, the classification used most often for talar neck fractures is that proposed by Hawkins.³ Its importance resides in the fact that this classification allows the standardization of conducts for the types described, estimates the prognosis and allows a comparison of results obtained with other publications. It is essential for its concordance to be high, both in the case of an evaluation between different observers (interobserver) and the same observer at different times (intraobserver).

This study aims to evaluate the reproducibility of Hawkins' classification in talar neck fractures in the intra- and interobserver aspects.

MATERIALS AND METHODS

Twenty cases of talar neck fracture with pathological fractures were selected, according to the following exclusion criteria: pathologic fractures, associated malleolus fractures or deformities in the ankle secondary to other pathological processes, to be classified by 36 observers. Eight of these observers were orthopedic surgeons and 28 orthopedic and radiology residents (1st to 3rd year).

The radiographs selected were the same used to define the patient's treatment, thus reproducing the conditions of daily practice at the institution. The professionals only used anteroposterior (AP) and lateral (L) radiographs of the ankle taken

All the authors declare that there is no potential conflict of interest referring to this article.

Universidade Estadual de Campinas DOT-UNICAMP – Campinas, SP, Brazil

Study conducted at the Department of Orthopedics and Traumatology of Universidade Estadual de Campinas, SP.

Mailing address: Rua Tessália Vieira de Camargo, n°126, Cidade Universitária Zeferino Vaz, Barão Geraldo, Campinas, SP, Brazil, ZIP-CODE 13083-970, Caixa Postal 6111.

Email: madseira@yahoo.com.br

Article received on 09/26/2011, and approved on 01/18/2012.

prior to the reduction when indicated, without any form of traction and with the limb unrestricted. (Figure 1)

The cases were collected retrospectively, excluding pathological fractures, associated malleolus fractures or patients with deformities in the ankle secondary to other pathological processes. The evaluations were carried out in an auditorium, with the classification presented to the survey participants by an orthopedic surgeon. Afterwards, a copy of Hawkins' original article was distributed for reading and reference during the evaluation. During the evaluation process, all the participants also received a schematic drawing of the classification. (Figure 2)



Figure 1. Lateral radiograph of ankle, exemplifying the images used during the evaluation (DOT-HC/Unicamp).
Image contributed by the Department of Orthopedics and Traumatology - HC/Unicamp.

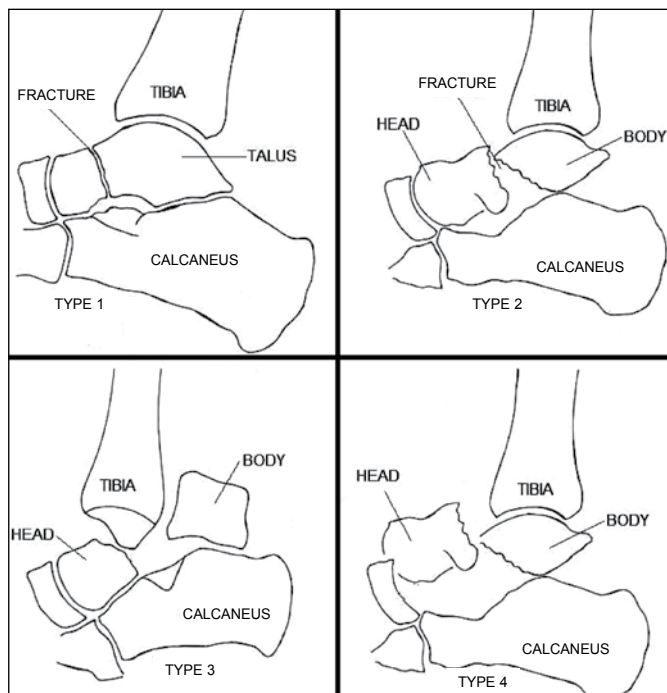


Figure 2. Graphic exemplification of Hawkins' classification (the same used during application of the study).

When evaluating the reliability of interobserver agreement it is necessary to incorporate the agreement occurring by chance in the evaluation.^{4,5} The intraclass correlation coefficient was used to verify the agreement^{6,9} and the criteria of Landis and Koch⁸ were considered for interpretation of the following strengths of agreement:

- a) almost perfect: 0.80 to 1.00;
- b) substantial: 0.60 to 0.80;
- c) moderate: 0.40 to 0.60;
- d) fair: 0.20 to 0.40;
- e) mild: 0 to 0.20;
- f) poor: -1.00 to 0.

A professional from the area was called in for the statistical calculation and to interpret the meaning of the results.^{7,8}

RESULTS

We present below Tables 1 and 2 with the intra/interobserver agreement results, based on the statistical/computer-aided calculations*.

*SAS System for Windows (Statistical Analysis System) [computer program], version 9.2. SAS Institute INC, 2002-2008, Cary, NC, USA

*SPSS for Windows [computer program], version 10.0. SPSS Inc, 1989-1999, Chicago, Illinois, USA.

The results presented in the graphs show that the correlation of Hawkins' interobserver classification presents a general mean considered "substantial" according to our coefficient, both in the first and in the second evaluation [0.627 (0.487;0.784) and 0.668 (0.532;0.813), respectively].

In analyzing the interobserver classification in groups formed by 1st, 2nd and 3rd year orthopedic residents, 1st, 2nd and 3rd year radiology residents and orthopedists, we verified an increase in agreement both in the first and in the second evaluation according to experience. The lower the level of experience, the worse the correlation of the fracture presented with the classification

Table 1. Intraclass correlation coefficients for interobserver evaluation in each professional category and in general.

Category	No. of evaluators	ICC (IC95%)
1st evaluation		
R1 – radiology	6	0.485 (0.300;0.696)
R2 – radiology	6	0.730 (0.578;0.861)
R3 - radiology	3	0.738 (0.541;0.875)
orthopedists	8	0.750 (0.607;0.872)
R1 - orthopedics	5	0.494 (0.297;0.708)
R2 – orthopedics	4	0.672 (0.480;0.832)
R3 – orthopedics	4	0.770 (0.614;0.888)
General	36	0.627 (0.487;0.784)
2nd evaluation		
R1 – radiology	6	0.671 (0.503;0.825)
R2 – radiology	6	0.717 (0.559;0.853)
R3 - radiology	3	0.598 (0.318;0.802)
orthopedists	8	0.704 (0.555;0.843)
R1 - orthopedics	5	0.654 (0.472;0.817)
R2 – orthopedics	4	0.643 (0.444;0.815)
R3 - orthopedics	4	0.836 (0.706;0.923)
General	36	0.668 (0.532;0.813)

Table 2. Intraclass correlation coefficients for intraobserver evaluation.

Evaluator	ICC (IC95%)	Evaluator	ICC (IC95%)
R1A radiology	0.681 (0.349;0.861)	MOD	0.909 (0.779;0.964)
R1B radiology	0.662 (0.328;0.850)	MOE	0.701 (0.383;0.870)
R1C radiology	0.544 (0.164;0.788)	MOF	0.876 (0.663;0.953)
R1D radiology	0.805 (0.577;0.918)	MOG	0.704 (0.336;0.877)
R1E radiology	0.893 (0.748;0.956)	MOH	0.969 (0.926;0.988)
R1F radiology	0.868 (0.701;0.945)	R1A orthopedics	0.529 (0.112;0.784)
R2A radiology	1.0000	R1B orthopedics	0.695 (0.380;0.867)
R2B radiology	0.819 (0.602;0.924)	R1C orthopedics	0.789 (0.540;0.911)
R2C radiology	0.702 (0.395;0.869)	R1D orthopedics	0.752 (0.482;0.893)
R2D radiology	0.683 (0.329;0.864)	R1E orthopedics	0.374(-0.084;0.697)
R2E radiology	0.801 (0.563;0.916)	R2A orthopedics	0.622 (0.251;0.832)
R2F radiology	0.839 (0.638;0.933)	R2B orthopedics	0.840 (0.644;0.933)
R3A radiology	0.695 (0.280;0.877)	R2C orthopedics	0.490 (0.059;0.764)
R3B radiology	0.785 (0.534;0.909)	R2D orthopedics	0.894 (0.756;0.956)
R3C radiology	0.890 (0.736;0.955)	R3A orthopedics	0.641 (0.300;0.839)
MOA	0.901 (0.768;0.959)	R3B orthopedics	0.723 (0.432;0.879)
MOB	0.969 (0.926;0.988)	R3C orthopedics	0.840 (0.638;0.934)
MOC	0.471 (0.049;0.751)	R3D orthopedics	0.943 (0.864;0.977)

MO= orthopedic physician (A,B,C,D,E,F,G,H); R1= first-year resident of the specialty; R2= second-year resident of the specialty; R3= third-year resident of the specialty.

(0.485 and 0.494 for the radiology and orthopedic R1 group respectively in the 1st evaluation) while the correlation in the radiology and orthopedic R2 and R3 groups and orthopedists ranged from 0.672 to 0.770.

In the 2nd evaluation the radiology and orthopedic R1s presented results that were superior (0.671 and 0.654 respectively) to the first, yet slightly inferior to the others evaluated, whose mean values were between 0.598 and 0.836. The third-year orthopedic residents presented higher values in the two evaluations. Firstly only R1 of the two specialties did not obtain substantial values, being classified in the moderate group, and in the second evaluation only the radiology R3 group obtained a moderate result, with the orthopedic R3 group members presenting an almost perfect value.

There was an improvement in agreement from the first to the second evaluation between the radiology and orthopedic R1s and the orthopedic R3s, while there was a deterioration of correlation between the radiology and orthopedic R2 group and orthopedists.

We observed that in spite of the variations that occurred between the two evaluations, values responsible for changes in the agreement scaling were only found in three groups. The radiology and orthopedic R1 group that climbed from the moderate coefficient to substantial (0.485 - 0.671 and 0.494 - 0.654

respectively), the orthopedic R3 group that climbed from substantial to almost perfect (0.770 - 0.836) and the radiology R3 group that fell from substantial to moderate (0.738 - 0.598). In the intraobserver classification we could see a high rate of agreement, since 14/36 presented substantial coefficient, 16/36 almost perfect, 4/36 moderate, 1/36 fair and 1/36 with 100%. In the group of orthopedists, five presented almost perfect coefficient (0.876 - 0.969), two substantial (0.701 - 0.704) and only one moderate (0.471).

DISCUSSION

In orthopedics classifications are extremely important and frequently used for treatment guidance, prognosis and case discussions. Hawkins' classification is the most widespread and used for talar neck fractures. Its importance as a prognostic factor and indication of treatment causes the need for a high rate of reproducibility and reliability, both intra- and interobserver. In analyzing the overall result, we inferred that Hawkins' classification presents a substantial interobserver result, without considering their experience, contact with this type of fracture and specialty (orthopedics or radiology). Such a confirmation was verified with the means within the satisfactory limits in the two evaluations.

Although there are no data in literature on the validity of Hawkins' classification for comparison of the different groups, it was observed in our survey that the evaluator's experience is crucial for better fracture-classification correlation, and that despite the biases, the degree of reliability of the classification is adequate for its use in the daily practice. This was demonstrated by the higher values, in general, according to experience, with progression from R1 to R3 in the two specialties, and better rates among the orthopedic graduates.

It was also observed with this study that the intraobserver classification varied only slightly, which means that this classification is reproducible, as 30 of the 36 evaluators remained within the coefficient ranges from substantial to almost perfect.

The professionals with less experience presented an improvement in the time between the two evaluations, probably due to the interest in and study of the classification, since they knew that they would be assessed for understanding of the topic again. The matter of length of studies can explain the greater agreement of the orthopedic R3 group and orthopedists specialized in feet, since the latter are familiar with the topic and come across it more frequently, and the former due to the studying necessary to acquire the title of specialist.

There were some decrease of agreement levels in the second evaluation, in the radiology R2 and R3 groups, orthopedic R2 group and orthopedists. Such an occurrence signifies deterioration in the groups in the subject studied, but may mean evolution of observers inside the group, causing different grades due to study of the classification. It would be a minor disagreement, which is believed to occur due to knowledge obtained in the interval between evaluations by only part of the observers from each group. Such a bias could be controlled by not informing the study participants of the existence of a second evaluation, which perhaps would not induce study geared towards the subject of many of the observers.

Since a classification serves as a prognosis and treatment guideline, it is extremely necessary for there to be agreement

among its users. According to the data encountered, we concluded that Hawkins' classification has achieved goals, facilitating understanding of the case, treatment guidelines, prognosis of the injury and discussions, as it presents satisfactory agreement among its observers.

CONCLUSION

In the majority of the groups the agreement of Hawkins' classification can be considered substantial, ranging between 0.6 and 0.8. It is only among the third-year orthopedic residents in the second evaluation that agreement can be considered perfect (0.836). Such a fact can be correlated with the studies undertaken by these residents prior to the examination to achieve the title of orthopedist. However, this contradicts the

agreement among orthopedic graduates, whose values were between 0.471 and 0.909, whereas high values are to be expected in both evaluations.

The result showed substantial agreement in general in the first and second evaluations, averaging 0.627 and 0.668 respectively. The data indicate a certain reliability of the classification, but that may vary a great deal according to the evaluator's experience, confirmed by the greater agreement, in general, pursuant to the years of activity in the area and studies undertaken.

ACKNOWLEDGMENT

To the direct and indirect participants of this study who made its performance possible.

REFERENCES

1. Sanders DW. Fractures of the talus. In: Bucholz RW, Heckman JD, editors. Rockwood & Green's fractures in adults. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2006. p.2249-93.
2. Heckman JD. Fraturas do talus. In: Bucholz RW, Heckman JD, editors. Rockwood & Green's fraturas em adultos. Tradução de Dr. Maurício Barreto de Castro 5a.ed. Barueri, SP: Manole; 2006. p.2091-132.
3. Hawkins LG. Fractures of the neck of the talus. J Bone Joint Surg Am. 1970;52(5):991-1002.
4. Everitt BS. The analysis of contingency tables. 2nd ed. London: Chapman & Hall; 1992. p.146-50.
5. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons; 1981.
6. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20(1):37-46.
7. Svanholm H, Starklint H, Gundersen HJ, Fabricius J, Barlebo H, Olsen S. Reproducibility of histomorphologic diagnoses with special reference to the kappa statistic. APMIS. 1989;97(8):689-98.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.