Observer variation in the Lauge-Hansen classification of ankle fractures

Precision improved by instruction

Sten Rasmussen[1], Poul Verner Madsen[1] and Kim Bennicke[2]

We investigated whether instruction can decrease the variation in the classification of ankle fractures. The radiographs of 100 consecutive ankle fractures were twice assessed by 8 observers, using the Lauge-Hansen classification system. 1 group of 4 observers was given special instruction between 2 assessments. Without instruction the interobserver variation was identical at the first and second assessments. After instruction, the agreement increased from 0.66 to 0.75. These results show that the variation in the Lauge-Hansen classification system can be reduced by instruction.

Departments of [1]Orthopedics and [2]Radiology, Hjørring Hospital, DK-9000 Hjørring, Denmark
Correspondence: Dr. Sten Rasmussen, Kong Hans Gade 11, DK-9000 Aalborg, Denmark. Tel +45–98 16 19 14
Submitted 93-02-10. Accepted 93-06-21

According to Yde (1980), the classification of Lauge-Hansen (1954) gives a very exact description of ankle fractures. However, a high interobserver variation was found by Lindsjö (1985) and Nielsen et al. (1990). This could be caused by insufficient knowledge of the classification system. We conducted this study to determine whether this observer variation can be reduced by instruction.

Material and methods

The radiographs of 100 consecutive ankle fractures were collected. The series did not include cases with open physes, direct fractures, pathological fractures or preexisting ankle deformation. No special criteria were set about the quality of the radiographs, as they had been accepted as of sufficient quality to form the basis of treatment. Each set of the radiographs was classifiable according to the Lauge-Hansen system (1942, 1954).

The observers were 2 consultant radiologists, 2 senior house officer radiologists, 2 senior registrar orthopedic surgeons, and 2 senior house officer orthopedic surgeons. The observers were randomly assigned to 2 groups, A and B, with one at each level of education.

Each observer was asked to classify the fractures according to the Lauge-Hansen system with the help of special forms. The classes were supination-adduction fracture stage SA 1–2, supination-eversion fracture stage SE 1–4, pronation-abduction fracture stage PA 1–3, pronation-eversion fracture stage PE 1–4 and pronation-dorsal-flexion stage PD 1–2. After 2 weeks Group B had a 2-hour instruction in the Lauge-Hansen classification. After 4 weeks, the same radiographs were again assessed by the 8 observers, but in a new and random order.

Kappa statistics (Siegel and Castellan 1988) were used to measure interobserver and intraobserver variations. The proportion of times that the observers agree P(A) and the proportion of times when the observers agreed by chance P(E) were calculated. The Kappa coefficient of agreement is the ratio of the proportions of time that the observers agreed (corrected for chance agreement) to the maximum proportion of times when the observers could agree (corrected for chance agreement). The statistic z was used to test the significance of Kappa and 95 percent confidence limit of Kappa was determined.

Results

From 3.5 to 6.3 percent of the ankle fractures were classified as SA, 66–69 percent as SE, and 25–29 percent as P (Table 1). At the second assessment, the classification by the group of observers who received instruction (B2) differed from the other group (A2).

At the first assessment, the interobserver variation was equal in Groups A1 and B1 (Table 2). At the sec-

Table 1. The distribution of 100 ankle fractures classified twice according to Lauge-Hansen by 2 groups (A and B) of each 4 observers. Group B received instruction between first and second assessments

	A1	B1	A2	B2
SA1	5	7	10	9
SA2	19	9	15	13
SE1	2	1	2	0
SE2	65	70	41	61
SE3	31	32	27	39
SE4	177	164	203	164
PA1	6	10	4	2
PA2	3	3	9	2
PA3	33	39	38	25
PE1	13	21	6	17
PE2	2	1	0	0
PE3	17	18	13	35
PE4	26	21	23	32
PD1	0	0	5	0
PD2	1	4	4	1

Table 2. Observer agreement P(A), agreement by chance P(E), kappa coefficient K and statistics in the 2 groups of observers at the first (A1 and B1) and second (A2 and B2) assessments. Group B received instruction between first and second assessments

	P(A)	P(E)	K	95% confidence limits of K	Statistic *z* of K
A1	0.63	0.25	0.51	0.45–0.57	16.7
B1	0.61	0.22	0.50	0.45–0.55	18.0
A2	0.66	0.30	0.53	0.46–0.60	14.4
B2	0.75	0.22	0.67	0.62–0.73	24.4

Table 3. Intraobserver agreement P(A), agreement by chance P(E) and statistics. Group B received instruction between first and second assessments

	P(A)	P(E)	K	95% confidence limits of K	Statistic *z* of K
A, consultant	0.82	0.29	0.75	0.66–0.87	11.8
A, registrar	0.66	0.28	0.53	0.40–0.65	8.36
A, SHO	0.52	0.27	0.34	0.22–0.46	5.62
A, SHO	0.59	0.24	0.46	0.35–0.57	8.36
B, consultant	0.53	0.23	0.39	0.29–0.50	7.27
B, registrar	0.61	0.25	0.48	0.37–0.59	8.44
B, SHO	0.59	0.24	0.46	0.34–0.57	7.98
B, SHO	0.53	0.17	0.43	0.34–0.52	9.40

ond assessment, the kappa coefficient in Group B2 after instruction was higher than in Group A2. The intraobserver variation is shown in Table 3. The kappa coefficient for the observers who received instruction (mean 0.44) was slightly lower than for the other observers (mean 0.52).

Discussion

The Lauge-Hansen classification system of ankle fractures consists of 15 well-defined stages in 5 classes. Yde (1980) showed in 488 ankle fractures that 99 percent could be classified according to this system. Only Stage 2 PA and PE fractures were not fully covered by Lauge-Hansen. Lindsjö (1985) compared 6 patient materials classified according to Lauge-Hansen and found a high variation in the frequency of the different fracture types. The reason for this disagreement was not clear, but different selection criteria could play a role.

In our study, the distribution of fracture types was comparable to Yde's (1980) study. The observed variation at the first assessment in this study was not better than in the study by Nielsen et al. (1990). However, instruction reduced the observed variation. A classification system is useful only if it can help in decision-making. The Lauge-Hansen system has to a certain extent fulfilled this demand. This study shows the benefit of instruction.

Acknowledgements

The financial support of Forskningslegatet ved Hjørring Sygehus, stiftet af Niels Jensen, and Nordjyllands Amts Forskningsråd for the studies presented in this paper is greatly appreciated.

References

Lauge-Hansen N. Fractures of the ankle. III. Genetic roentgenologic diagnosis of the fractures of the ankle. Am J Roentgenol 1954; 71: 456–71.

Lauge-Hansen N, Ankelbrud I. Genetisk diagnose og reposition. Thesis, University of Copenhagen, Copenhagen, Denmark 1942.

Lindsjö U. Classification of ankle fractures: the Lauge-Hansen or AO system?. Clin Orthop 1985; 199: 12–6.

Nielsen J Ø, Dons-Jensen H, Sørensen H T. Lauge-Hansen classification of malleolar fractures. An assessment of the reproducibility in 118 cases. Acta Orthop Scand 1990; 61 (5): 385–7.

Siegel S, Castellan N J. (Eds.). Nonparametric statistics for the behavioral sciences. McGraw Hill, New York 1988.

Yde J. The Lauge-Hansen classification of malleolar fractures. Acta Orthop Scand 1980; 51 (1): 181–92.