Assessment of the AO classification of intracapsular fractures of the proximal femur

Chris M. Blundell, Martyn J. Parker, Glyn A. Pryor,
James Hopkinson-Woolley, Sachin S. Bhonsle

From Peterborough District Hospital, England

There are a number of classification systems for intracapsular fractures of the proximal femur, but none has been shown to be practical with satisfactory reproducibility and accurate predictive value. We have investigated the AO classification and evaluated intra- and interobserver accuracy and its value in predicting treatment and outcome.

We found it to have very poor intra- and interobserver reliability and to be of limited predictive use for the outcome of treatment. A simplified system in which the subdivisions were allocated to one of three groups of undisplaced, displaced and basal fractures was found to be of value. We conclude that this is the only division which is appropriate for these fractures and that the AO system for intracapsular fractures is too complicated and should not be used.

J Bone Joint Surg [Br] 1998;80-B:679-83.

Received 17 October 1997; Accepted 19 January 1998

There has been no comprehensive, reliable and universally accepted method of classification of intracapsular fractures of the proximal femur since Waldenström¹ described a simple division into stable (undisplaced) and unstable (displaced) fractures. The most commonly used method for the classification of these fractures is probably that of Garden,² who divided them into four groups using the degree of displacement or impaction, but a number of reports have shown this system to have a low interobserver reliability.³⁻⁶ A radiological study found that the degree of displacement between fractures graded as Garden III or IV was not significant.⁷ A clinical study also found that the difference

in the rates of fracture healing for Garden-III and Garden-IV fractures treated by internal fixation was insufficient to justify the separate grades.⁸ The value of this system is therefore doubtful.

Pauwels' classification⁹ which is based on the angle of the fracture line with the femoral shaft has been questioned and has consistently been reported to be of little value in determining the risk of nonunion or of avascular necrosis.¹⁰⁻¹³ The level of the fracture has been used as the basis of other systems,^{14,15} but most studies have found it to have no correlation with the incidence of complications of fracture healing^{13,16} although there have been some contradictory reports. Barnes et al¹⁷ found an increased risk of nonunion for the more proximally based fractures while Fielding et al¹¹ reported the opposite.

Any useful system should consider the severity of the injury, be a guide to treatment and facilitate the comparison of results. It must also have good intra- and interobserver reliability. The most recently devised system is part of the AO universal classification for the whole skeleton (Fig. 1).¹⁸ To investigate its value for intracapsular fractures of the proximal femur, we considered its reliability and reproducibility by investigating intra- and interobserver error and then its accuracy in predicting the method of treatment and the incidence of the complications of fracture healing.

Patients and Methods

The radiographs used for the assessment of inter- and intraobserver variation were selected from the database of the Peterborough Hip Fracture Project as being representative of all the possible types of intracapsular fracture. They had previously been classified using the AO system at the time of admission to hospital.

To determine the intra- and interobserver reliability each of the five observers (CMB, MJP, GAP, JHW and SSB) independently assessed the anteroposterior (AP) and lateral radiographs of 71 intracapsular fractures and classified them. They were not allowed to see how the fracture had subsequently been treated or if complications of healing had occurred. Each observer had reference to the AO classification (Fig. 1) and was also given access to the relevant instructional CD-ROM while the fractures were

C. M. Blundell, FRCS, Orthopaedic Specialist Registrar
M. J. Parker, FRCS, Orthopaedic Research Fellow
J. Hopkinson-Woolley, FRCS, Orthopaedic Specialist Registrar
S. S. Bhonsle, MS Orth, Senior House Officer
G. A. Pryor, MS, FRCS, Consultant Orthopaedic Surgeon
Peterborough District Hospital, Thorpe Road, Peterborough PE3 6DA, UK.

Correspondence should be sent to Mr C. M. Blundell at 44 Christchurch Road, Norwich, Norfolk NR2 3NE, UK.

©1998 British Editorial Society of Bone and Joint Surgery
0301-620X/98/48476 \$2.00

Fig. 1

The AO classification of intracapsular proximal femoral fractures.

being classified; they were given as much time as required to evaluate the radiographs accurately.

They were instructed not to discuss the classification system among themselves until after the study had been completed to avoid increasing the observer correlation above that normally expected. They submitted their classifications without keeping a record and these were then held in sealed envelopes until the end of the study. Eight to ten weeks after the initial assessment the radiographs were presented in a different order for the observers to reclassify them. Three sets of radiographs could not be classified for a second time as these had been required for clinical use

and could not be obtained, leaving 68 fractures for reclassification.

To investigate the predictive value of the AO classification in outcome and treatment, we evaluated a larger series of radiographs. All proximal femoral fractures admitted to Peterborough District Hospital from November 1994 to March 1997 were classified (by one observer, MJP) using among others, the AO classification system. The study group of 415 consecutive patients all had had acute intracapsular fractures which had been treated surgically with follow-up of at least six months from injury. The method of surgical treatment was recorded and for those treated by

internal fixation, the incidence of nonunion was determined. This was defined as displacement of the fracture necessitating revision surgery or failure of the fracture to be united radiologically one year from injury.

For statistical analysis we used the kappa test as introduced by Cohen to determine the level of reliability. This has been interpreted by Altman as "the chance-corrected proportional agreement".¹⁹ Kappa is a coefficient of agreement which has a value varying from +1 representing perfect agreement, through 0 representing an agreement no better than chance, to -1, absolute disagreement. There are no definitions of the level of agreement which is acceptable but the guidelines of Svanholm et al²⁰ are that values of 0.75 or more represent excellent agreement, 0.5 to 0.75 good, and 0.5 or less poor agreement. The kappa coefficient depends on the prevalence of the category in question. Thus, we must be aware that a highly prevalent category may result in a raised level of expected agreement; this, however, is inherent in classification systems.

Results

Table I gives the results of the kappa statistical analysis of the inter- and intraobserver variation of the AO classification. The range of values for interobserver reliability which compared the first classifications between observers yielded kappa values of from 0.21 to 0.51 with an overall mean of 0.30. Thus, on all but one occasion these represented poor reliability. Intraobserver analysis gave kappa values ranging from poor (0.33) to good (0.61), but with a mean of 0.5, again representing overall poor reliability.

The 68 fractures were then reassigned using a simplified system into undisplaced (B1.1, B1.2 and B1.3), basal

(B2.1) and displaced (B2.2, B2.3, B3.1, B3.2 and B3.3) to determine whether there was an improvement in the intra- and interobserver reliability. The results are shown in Table II. Interobserver reliability was much improved; all but one observer had a mean reliability coefficient greater than 0.75, i.e. excellent agreement. The results for this simplified system also gave excellent intraobserver reliability (kappa values 0.77 to 0.96, mean 0.85).

The results of the second part of the study assessing the predictive value of the AO system, with the larger group of 415 patients, are shown in Table III. This also gives the rates of nonunion for those fractures treated by internal fixation.

Discussion

The AO classification of intracapsular fractures of the hip is at present the most complex method devised. It is based on a combination of the level of the fracture, the degree of displacement and the angle of the fracture line. It includes basicervical or basal fractures within the intracapsular system, although these have previously been described as extracapsular fractures.²¹⁻²³ This lack of agreement occurs because the fracture line crosses the anatomical boundaries of the hip capsule. Our study shows that for basal fractures the treatment and prognosis are similar to those of an extracapsular fracture; basal fractures should therefore be classified either separately or as an extracapsular injury.

Because of the poor reproducibility of the AO system of classification, a simplified method was designed taking into account the inter- and intraobserver reliability, the method of treatment and the rate of union. B1 (undisplaced) fractures were predominantly fixed using parallel screws (in 72

Table I. Kappa values of agreement for inter- and intraobserver variation for the AO classification

	AO classification					Intraobserver
	Interobserver					
	CB	MJP	JHW	GAP	SSB	
CB	X	0.3	0.28	0.21	0.34	0.33
MJP		X	0.24	0.29	0.28	0.61
JHW			X	0.22	0.51	0.56
GAP				X	0.3	0.37
SSB					X	0.61
Mean	0.28	0.28	0.31	0.25	0.36	0.5

Table II. Kappa values of inter- and intraobserver variation for the simplified classification

	Simplified classification					Intraobserver
	Interobserver					
	CB	MJP	JHW	GAP	SSB	
CB	X	0.61	0.58	0.63	0.64	0.77
MJP		X	0.82	0.82	0.82	0.85
JHW			X	0.82	0.84	0.91
GAP				X	0.91	0.77
SSB					X	0.96
Mean	0.61	0.77	0.76	0.8	0.8	0.85

Table III. Method of treatment and incidence of nonunion for 415 fractures classified by the AO system, by number and *percentage*

AO type	Treatment*					Nonunion
	Screws	SHS	Hemi	THR	Conservative	
Undisplaced						
B1.1	46	1	1	0	1	4/47 (8.5)
B1.2	19	0	1	0	2	1/19 (5.3)
B1.3	7	0	0	1	0	0/7 --
Basal						
B2.1	0	13	3	1	0	0/13 --
Displaced						
B2.2	3	0	2	0	0	1/3 (33.3)
B2.3	2	1	2	0	0	1/3 (33.3)
B3.1	32	0	46	0	1	10/32 (31.3)
B3.2	43	1	28	2	0	8/44 (18.2)
B3.3	46	2	106	0	2	14/48 (29.2)

* SHS, sliding hip screw; Hemi, hemiarthroplasty; THR, total hip replacement

out of 79 cases) and had a rate of nonunion of less than 9%. B2.1 fractures behaved as a separate group in which the method of treatment was by a sliding hip screw in 13 out of 17 cases. There were no cases of nonunion within this group. The method of treatment of the remaining fracture types was variable with 130 cases fixed with screws and the remaining 184 having a hemiarthroplasty. The nonunion rate for these was clearly higher at 26% (Table III).

The results of review with the simplified classification showed that for intraobserver analysis the mean kappa value was 0.85, which is equivalent to excellent reliability, and for interobserver reliability the mean kappa value was 0.75. All but one observer had a mean coefficient of interobserver reliability greater than 0.75, i.e. excellent agreement.

Our study shows that the AO classification of intracapsular hip fractures has an extremely poor intra- and interobserver reliability. The system is too complex and detailed to allow observers to be confident in their assessment. The simplified system in which fractures are grouped as either displaced, undisplaced or basal, gives excellent intra- and interobserver reliability. It could be suggested that reducing the number of possible options into which a fracture is classified would in itself yield a more reliable system, but this is not necessarily the case. For example, Sidor et al²⁴ found that reducing the number of classification categories for the modified Neer system for proximal humeral fractures did not increase reliability.

Basal fractures (31B2.1) were considered separately in our modified system because they require a different method of treatment. They were found to have a high degree of both intra- and interobserver reliability. There was agreement on a B2.1 grade in 49 out of 60 (82%) separate classifications in six hip fractures; this included three sets of radiographs in which all observers agreed on the classification of B2.1 at both the initial and subsequent classifications. Thus, the assignment of a B2.1 classification was made with a high degree of confidence. This level of agreement only occurred in one other set of radiographs (a B1.2 fracture) out of the total of 68 hip fractures.

The only other reports to address the value of the AO classification of hip fractures found it to be of no value. Newey, Ricketts and Roberts²⁵ studied whether the method of treatment used was related to the AO classification and found that classification of proximal femoral fractures into groups 31.A and 31.B (broadly intracapsular and extracapsular) did predict treatment, but that further subdivision was unhelpful. A small study of five hip fractures graded by ten observers found that no fracture was classified identically by all observers.²⁶

It has previously been suggested that intracapsular fractures should be classified simply as either displaced or undisplaced.²⁷ Our study supports this in that the simplified classification has a much higher rate of inter- and intraobserver reliability than the more complex AO classification system, and predicts the method of treatment and the rate of nonunion more effectively. Because of the high degree of both intra- and interobserver variation which we have demonstrated, the use of the AO classification of intracapsular fractures of the proximal femur as it stands should be abandoned. A simplified system based on the three main groups of the AO system would be of value.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

References

1. Waldenström J. Fractures recentes du col femoral: traitement opératoire ou orthopédique. *J Chir* 1924;24:129.
2. Garden RS. Low-angle fixation in fractures of the femoral neck. *J Bone Joint Surg [Br]* 1961;43-B:647-63.
3. Frandsen PA, Andersen E, Madsen F, Skjodt T. Garden classification of femoral neck fractures: an assessment of inter-observer variation. *J Bone Joint Surg [Br]* 1988;70-B:588-90.
4. McCabe J, O'Farrell D, O'Byrne J, O'Brien T. Observer variation in the radiographic classification of femoral neck fractures. *J Bone Joint Surg [Br]* 1994;76-B:(Suppl II & III):120.
5. Scavenius M, Ibsen A, Rønnebeck J, Aagaard H. Inter- and intra-observer variation in the assessment of femoral neck fractures according to the Garden classification. *Acta Orthop Scand* 1996;67(Suppl 267):29.
6. Thomsen NO, Jensen CM, Skovgaard N, et al. Observer variation in the radiographic classification of fractures of the neck of the femur using Garden's system. *Int Orthop* 1996;20:326-9.

7. Eliasson P, Hansson LI, Kärrholm J. Displacement in femoral neck fractures: a numerical analysis of 200 fractures. *Acta Orthop Scand* 1988;59:361-4.
8. Parker MJ. Garden grading of intracapsular fractures: meaningful or misleading? *Injury* 1993;24:241-2.
9. Pauwels F. *Der schenkelhalsbruch ein mechanisches problem: Grundlagen des Heilungsvorganges Prognose und kausale Therapie*. Stuttgart: Ferdinand Enke Verlag, 1935.
10. Crawford HB. Impacted femoral neck fractures. *Clin Orthop* 1969; 66:90-3.
11. Fielding JW, Wilson SA, Ratzan S. A continuing end-result study of displaced intracapsular fractures of the neck of the femur treated with the Pugh nail. *J Bone Joint Surg [Am]* 1974;56-A:1464-72.
12. Massie WK. Treatment of femoral neck fractures emphasizing long term follow-up observations on aseptic necrosis. *Clin Orthop* 1973; 92:16-62.
13. Parker MJ. Prediction of fracture union after internal fixation of intracapsular femoral neck fractures. *Injury* 1994;25(Suppl 2):3-6.
14. Banks HH. Factors influencing the result in fractures of the femoral neck. *J Bone Joint Surg [Am]* 1962;44-A:931-64.
15. Tamai K, Azuma H, Kako K. A new anatomic classification of capital fragments in femoral neck fractures with the epiphyseal scar as a guide. *Clin Orthop* 1983;179:147-56.
16. Svenningsen S, Benum P, Nesse O, Furset OI. Internal fixation of femoral neck fractures: compression screw compared with nail plate fixation. *Acta Orthop Scand* 1984;55:423-9.
17. Barnes R, Brown JT, Garden RS, Nicoll EA. Subcapital fractures of the femur: a prospective review. *J Bone Joint Surg [Br]* 1976;58-B: 2-24.
18. Muller ME, Nazarian S, Koch P, Schatzker J. *The AO classification of fractures of long bones*. Berlin, etc: Springer-Verlag, 1990.
19. Altman DG. *Practical statistics for medical research*. Third ed. Chapman and Hall, 1995:403-9.
20. Svanholm H, Starklint H, Gundersen HJG, et al. Reproducibility of histomorphologic diagnoses with special reference to the kappa statistic. *APMIS* 1989;97:689-98.
21. Bohler L. *The treatment of fractures*. Vol. 11, 5th ed. New York, Grune & Stratton, 1957.
22. Harris LJ. Closed retrograde intramedullary nailing of peritrochanteric fractures of the femur with a new nail. *J Bone Joint Surg [Am]* 1980;62-A:1185-93.
23. Murray RC, Frew JF. Trochanteric fractures of the femur: a plea for conservative treatment. *J Bone Joint Surg [Br]* 1949;31-B:204-19.
24. Sidor ML, Zuckerman JD, Lyon T, et al. The Neer classification system for proximal humeral fractures: an assessment of interobserver reliability and intraobserver reproducibility. *J Bone Joint Surg [Am]* 1993;75-A:1745-50.
25. Newey ML, Ricketts D, Roberts L. The AO classification of long bone fractures: an early study of its use in clinical practice. *Injury* 1993;24:309-12.
26. De Boeck H. Classification of hip fractures. *Acta Orthop Belg* 1994;60 (Suppl 1):106-9.
27. Editorial. Reporting hip fractures in the elderly. *Acta Orthop Scand* 1988;59:359-60.