



Original article

The reliability of a Garden, AO and simple II stage classifications for intracapsular hip fractures



Povilas Masionis*, Valentinas Uvarovas, Giedrius Mazarevičius, Kiril Popov, Šarūnas Venckus, Karolis Baužys, Narūnas Porvaneckas

Clinic of Rheumatology, Traumatology Orthopaedic and Reconstructive Surgery, Centre of Orthopaedics and Traumatology, Medical Faculty, Vilnius University, Republican Vilnius University Hospital, Šiltnamių g. 29, LT-04130 Vilnius, Lithuania

ARTICLE INFO

Article history:

Received 2 August 2017

Accepted 14 November 2018

Keywords:

Femoral neck

Hip fracture

Garden

AO

ABSTRACT

Introduction: Successful treatment starts by accurate classification of pathology, but there is no conclusive, reliable and universally accepted method for classification of intracapsular femoral neck fractures. As a perfect classification should have high intra- and interobserver agreement, this study aims to access reliability of three classification systems: Garden, AO and simple II stage classification.

Materials and methods: Four orthopaedic trauma surgeons (two of them professors) and two senior orthopaedic trauma residents were invited to evaluate 136 blinded anterior-posterior and lateral X-rays of patients with femoral neck fractures. Observers classified fractures according to IV stage Garden, AO and simple II stage classifications. The exercise was repeated after one month on same but randomised X-rays. Cohen kappa was used to measure inter- and intraobserver agreement. Fleiss kappa was used to access multi-rater agreement.

Results: AO classification showed an overall agreement of 0.22 (fair agreement). Garden classification had overall reliability slightly higher than AO, but matching same fair agreement group (0.33). II stage classification provided the highest estimates: from 0.35 (fair agreement) to 0.83 (almost perfect agreement) and multi-rater agreement of 0.50 (moderate agreement). There was seen no difference in intra- and interobserver agreement between observer groups (professors, trauma surgeons and trauma residents).

Discussion: All three classification systems showed equal adoption among differently experienced observer groups. Despite this finding, IV stage Garden and AO classifications should be avoided in clinical use because of poor reproducibility. Only simple II stage classification showed sufficient intra- and interobserver reliability.

Level of evidence: IV, Retrospective study.

© 2018 Elsevier Masson SAS. All rights reserved.

1. Introduction

The management of femoral neck fractures (FNF) was 19th of the top 20 most expensive diagnoses for Medicare in 2013 in the US and it is expected to increase as the incidence is believed to be doubled by 2050 [1–9]. Successful treatment starts by accurate classification of pathology, but there is no conclusive, reliable and universally accepted method for classification of intracapsular FNF [10].

Firstly, intracapsular FNF were classified by Waldenström to stable (undisplaced) and unstable (displaced) [11]. In 1935, Pauwels

introduced classification based on the angle of the fracture line with the femoral shaft [12]. Although, it was reported to have low reliability and little value in predicting avascular necrosis and nonunion [13,14]. Garden developed possibly the most widely used classification system for FNF in 1961 [15]. Recently, this method suffered from criticism for questionable reliability and nonunion prediction [16–18]. The most recent and probably the most complex classification tool was developed as a part of the AO universal fracture classification system [10,19]. Furthermore, in current literature, there are some proofs that primary stable and unstable classification is possibly the best tool to classify intracapsular FNF [10,17,18].

A perfect classification for FNF should have a high intra- and interobserver agreement, predict possible complications, facilitate the comparison of results and suggest the most suitable implant which could provide good patient outcomes and cost-effectiveness

* Corresponding author.

E-mail address: povilasmasionis@yahoo.com (P. Masionis).

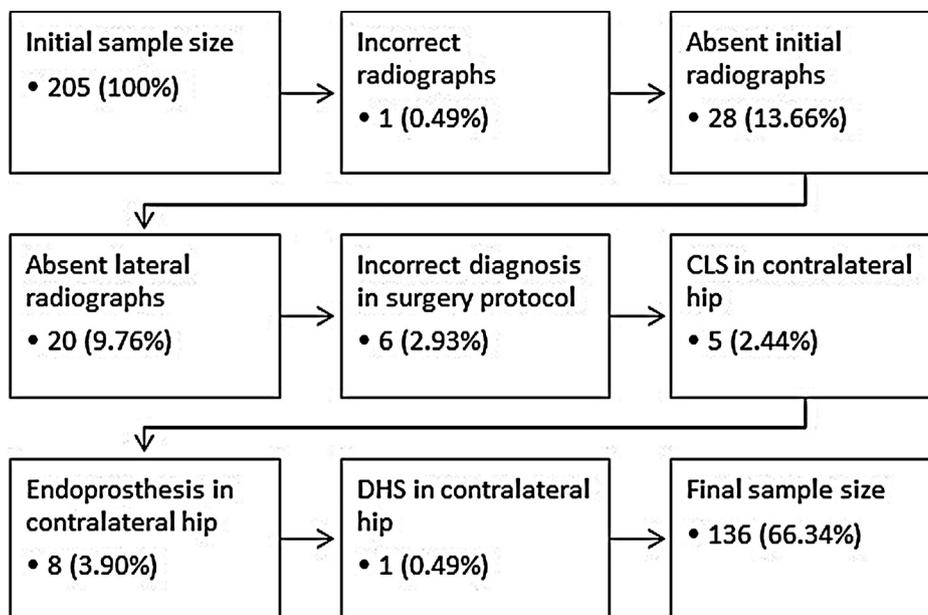


Fig. 1. Data analysis. CLS - cancellous lag screws, DHS - dynamic hip screw.

[17,20]. This study aims to assess reliability of three classification systems: Garden, AO and simple II stage classifications.

2. Materials and methods

2.1. Patients

From 2015 the first of the September to 2016 the first of the September 206 patients who were admitted to a single institution for FNF, were retrospectively selected for this study. The exclusion criteria were incorrect or absent radiographs and comorbidities which could influence final decision and are seen in radiographs: osteoarthritis of the hip joint, the hip prosthesis in contralateral side, previous FNF of the opposite side, hip joint dysplasia, pelvic fracture (Fig. 1). The final sample size was 136 patients of which 101 (74%) were females and 35 (26%) males with the average age of 76 ± 14 years and in 67 (49%) cases affected right side. In the radiographs, the name of the patient, patients and radiographs identification number were shaded to make it impossible to identify the patient (Fig. 2). Afterwards, radiographs were numbered.



Fig. 2. An example of blinded radiograph. 69 years old woman with right femoral neck fracture treated by cancellous lag screws. All observers suggested hip arthroplasty in first evaluation. In the second assessment, 5 out of 6 observers had chosen hip arthroplasty and one - cancellous lag screws as the preferred treatment method.

The quality of all radiographs was representative and initial decision on treatment option had been based on them.

3. Methods

Four orthopaedic trauma surgeons (two of them professors) and two senior orthopaedic trauma residents were invited to take part in the observer's group and evaluate 136 anterior-posterior and lateral X-rays. They were not allowed to see how the fractures were treated and if any complications occurred. All the participants were familiar with the Garden and AO classifications. Each observer was provided with a hyperlink to his personal folder on the virtual cloud (OneDrive, Microsoft corp.) which was shared by the main investigator. In the folder, there were X-rays of the patients and an overview with schemes and definitions of Garden, AO classifications and stable, unstable fractures (Fig. 3). The stable fracture was defined as having normal or near normal alignment and 2 fragments are expected to move as a unit under minimal force. Unstable fractures had opposite properties compared to stable. Each observer was asked to classify fractures independently according to Garden, AO and simple II stage classifications. The observers were provided as much time as needed for accurate evaluation. After the assessment had been done, the main investigator unshared the folders. One month after the initial assessment was done, same observers were provided with personal folders on the virtual cloud again and asked to reclassify the same randomised X-rays for the same criteria.

4. Statistical analysis

Statistical analysis was performed by calculating Cohen kappa value for inter- and intraobserver reliability as observed agreement. To calculate the multi-rater kappa for the interobserver agreement we used the method of Fleiss. We used the interpretation of the level of agreement by the guidelines of Landis and Koch: less than 0.00 poor reliability, 0.00 to 0.20 slight reliability, 0.21 to 0.40 fair reliability, 0.41 to 0.60 moderate reliability, 0.61 to 0.80 substantial agreement and 0.81 to 1.00 almost perfect agreement [21]. Statistical analysis was performed using SPSS v21.0 software. For additional calculations, Microsoft Excel 2016 was used.

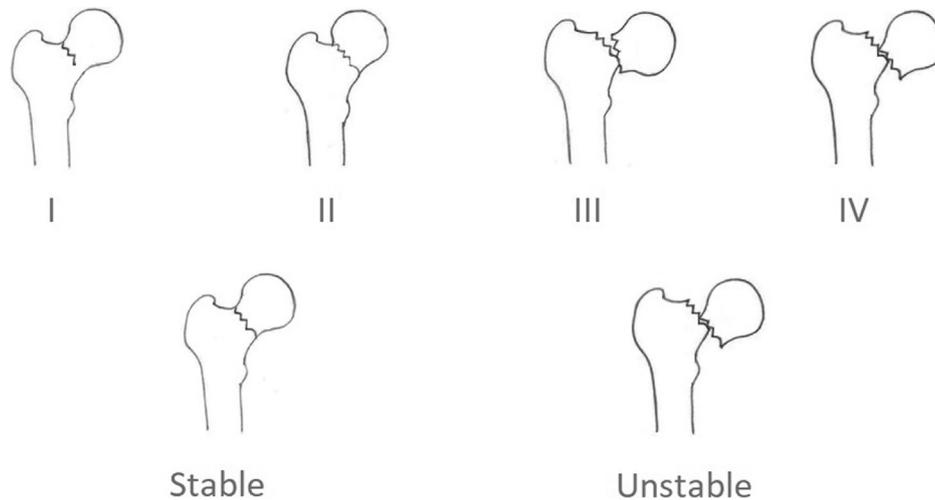


Fig. 3. Scheme used in assistance for observers. Garden classification above and simple II stage classification below.

5. Results

Cross table for patients on AO, Garden and simplified classification is presented in Table 1. Values for Cohen kappa of intra- and interobserver variation of AO, Garden, simplified classification and treatment method based on X-rays are presented in Tables 2–4 respectively. The range of values of interobserver reliability for AO classification was from 0.11 (slight agreement) to 0.43 (moderate agreement) with Fleiss kappa for overall agreement of 0.22 (fair agreement). Garden classification showed reliability from fair (0.21) to moderate (0.41) and overall reliability slightly higher than AO, but matching same fair agreement group (0.33). The II stage classification had the highest reliability: from 0.35 (fair agreement) to 0.83 (almost perfect agreement) and overall Fleiss kappa of 0.50 which represents moderate agreement.

Furthermore, there was seen no difference in intra- and interobserver agreement between observer groups (professors, trauma surgeons and trauma residents) (Table 5). All three groups matched fair agreement for AO classification and fair agreement for Garden classification. Professors and residents reached substantial agreement in II stage classification and trauma surgeons showed moderate agreement.

6. Discussion

The most important finding of the present study was lower interobserver agreement of AO and Garden classification than compared to simple II stage. In order to be a useful classification, there should be at least moderate rater consistency which was shown only by II stage method. Our findings favour opinion that the first two systems should be avoided in clinical use.

This is not the first report to propose II stage classification instead of AO and Garden for FNF. In literature, the interobserver agreement of Garden classification varies from fair to moderate and Blundell et al. found AO system to have a fair agreement [10,16,17]. Furthermore, the same studies showed from moderate to almost perfect agreement for II stage classification. Van Embden et al. found interobserver reliability of 0.31 for Garden and 0.52 for stable/unstable classification, which is in agreement with our results of 0.33 and 0.50 respectively. In both studies, observers were provided with overview and figures for Garden classification. Although, Van Embed et al. tried to mimic clinical situation and did not provide any additional assistance in II stage classification but we provided observes with the definition and schemes of the stable/unstable constept (Fig. 3). Despite, the different assistance

Table 1 Patients cross table for Garden, simple two stage and AO classifications and observer's decisions. Second observation is presented in brackets.

Garden classification						
Types	NP	VU	KP	GM	KB	SV
1	9(6)	3(1)	6(4)	12(5)	6(6)	4(1)
2	17(12)	15(6)	13(18)	8(12)	12(16)	17(14)
3	54(44)	45(65)	76(87)	19(22)	54(68)	74(88)
4	56(74)	73(64)	41(27)	97(97)	64(46)	41(33)
Simple two stage classification						
Stable	24(24)	26(17)	40(50)	16(17)	24(28)	15(18)
Unstable	112(112)	110(119)	96(86)	120(119)	112(108)	121(118)
AO classification						
B1.1	6(7)	7(3)	7(9)	11(11)	14(7)	8(8)
B1.2	8(7)	7(9)	8(9)	2(3)	0(14)	6(4)
B1.3	6(2)	1(5)	1(0)	1(2)	1(0)	1(2)
B2.1	9(5)	6(2)	4(3)	9(8)	8(8)	15(12)
B2.2	31(22)	4(2)	41(27)	24(1)	28(22)	35(42)
B2.3	13(15)	1(5)	8(2)	9(7)	39(20)	24(13)
B3.1	4(4)	4(4)	22(9)	2(5)	3(4)	11(9)
B3.2	19(12)	40(49)	21(49)	16(18)	19(32)	15(46)
B3.3	40(62)	66(57)	24(28)	62(81)	24(29)	21(0)

Table 2
Kappa values for inter- and intraobserver agreement for the AO classification.

	Interobserver						Intra-observer
	NP	VU	KP	GM	KB	SV	
NP	X						0.37
VU	0.25	X					0.47
KP	0.28	0.22	X				0.26
GM	0.29	0.43	0.21	X			0.48
KB	0.26	0.22	0.24	0.23	X		0.32
SV	0.25	0.11	0.20	0.15	0.25	X	0.29
Mean	0.27	0.25	0.23	0.26	0.24	0.19	0.37

Table 3
Kappa values for inter- and intraobserver agreement for the Garden classification.

	Interobserver						Intraobserver
	NP	VU	KP	GM	KB	SV	
NP	X						0.40
VU	0.39	X					0.40
KP	0.32	0.30	X				0.42
GM	0.27	0.39	0.21	X			0.57
KB	0.41	0.45	0.36	0.36	X		0.46
SV	0.35	0.33	0.36	0.23	0.38	X	0.41
Mean	0.33	0.37	0.31	0.29	0.39	0.33	0.44

Table 4
Kappa values for inter- and intraobserver agreement for II stage classification.

	Interobserver						Intraobserver
	NP	VU	KP	GM	KB	SV	
NP	X						0.65
VU	0.73	X					0.70
KP	0.47	0.40	X				0.47
GM	0.70	0.77	0.35	X			0.90
KB	0.73	0.76	0.45	0.68	X		0.72
SV	0.72	0.74	0.40	0.83	0.68	X	0.76
Mean	0.67	0.68	0.41	0.67	0.66	0.67	0.70

Table 5
Kappa values for inter and intraobserver agreement within observer groups.

	AO	Garden	II stage
Professors	0.28 ± 0.10	0.37 ± 0.05	0.68 ± 0.12
Trauma surgeons	0.27 ± 0.10	0.33 ± 0.10	0.56 ± 0.20
Residents	0.23 ± 0.06	0.36 ± 0.07	0.68 ± 0.13

the results are the same. In our opinion, the main reason is that terms stable/unstable are one of the most fundamental and generally accepted in trauma surgery and acquired by every surgeon in his early training. Furthermore, author had two groups of different skill observers (trauma surgeons and residents) and found no differences between them. The same pattern was observed in present study as we had three differently experienced observer groups (professors, trauma surgeons and residents) and there was found no difference between them. Despite equal adoption among the differently skilled observers, the reproducibility of AO and Garden classification is too low for the clinical use.

It is well known, that kappa values strongly depend on the number of modalities of classification investigated. It could be potential bias as, in present study, we had 9 level AO, 4 level Garden and 2 level simple classification. Garden and simple II stage methods could not be divided into upper and lower levels. Although, AO femoral neck fractures classification has 3 groups (B1, B2, B3) and each of them is consisted of 3 subgroups (B1.1, B1.2, B1.3, B2... etc.) having 9 modalities in total. Because of this feature, repeated analysis of AO classification data was performed after 9 subgroups were added to 3 groups from both observations. After

simplification lower boundary of inter-observer reliability classification improved from 0.11 (slight agreement) to 0.27 (fair agreement). Despite the improvement of upper boundary from 0.43 to 0.59 and Fleiss kappa for overall agreement from 0.22 to 0.39, it remained in same moderate and fair agreement groups respectively. After simplification kappa values improved but mainly remained within the same groups with the most probable reason that AO classification remains too detailed and complex tool for good reproducibility.

Some previous investigations named simple “stable/unstable” classification as the simplified Garden (Garden I-II as stable and Garden III-IV as unstable) [17]. Although, we would like to oppose this opinion as the observers identified 12.56% of Garden I - II fractures to be unstable and 6.53% of Garden III - IV as stable. Our data is in agreement with Beimers et al. who found 1.3% of stable and 5.1% of unstable were not categorized as Garden stage I-II or Garden stage III-IV respectively and stated that Garden classification should be replaced not collapsed [18]. This data suggests, that “stable/unstable” classification should not be used or named as collapsed IV stage Garden as II stages of “stable/unstable” do not correspond to relevant IV stages of Garden tool. Furthermore, historically, simple classification into stable and unstable FNF was firstly described by Walderström in 1924, before the development of Garden classification [11]. To eliminate further confusion in literature—simple “stable/unstable” classification should not be called or used as simplified Garden.

It has been previously suggested that FNF should be classified to stable, unstable and basicervical [10]. Anatomically basicervical fractures are intracapsular but clinical rationale of such

classification is doubtful. In retrospective and biomechanical studies, DHS with the derotational screw was found as the most suitable treatment option for these fractures [22,23]. In our opinion, basicervical FNF should not be used as the third stage as they are treated as extracapsular ones.

Authors are aware of limitations of the present study. First of all, it is a retrospective study and we did not evaluate patient outcomes. As a result, we do not have data on II stage classification relation to aseptic necrosis and nonunion rates. Furthermore, observers and patients were from the same hospital and some of the observers possibly took part in the definitive treatment of the patients.

Consent

Institution's ethic board approved the investigation. Written informed consent was obtained from the patients for publication of this study and any accompanying images.

Disclosure of interest

The authors declare that they have no competing interest.

Funding sources

No financing was received for the study.

Authors' contribution

Povilas Masionis designed the study, performed data analysis and drafted the manuscript. Narūnas Porvaneckas, Giedrius Mazarevičius, Valentinas Uvarovas, Šarūnas Venckus, Karolis Baužys, Kiril Popov took part as observers and drafted the manuscript.

References

- [1] Roberts KC, Brox WT. AAOS clinical practice guideline: management of hip fractures in the elderly. *J Am Acad Orthop* 2015;23:138–40.
- [2] Florschütz AV, Langford JR, Haidukewych GJ, Koval KJ. Femoral neck fractures: current management. *J Orthop Trauma* 2015;29:121–9.
- [3] Xu D-F, Bi F-G, Ma C-Y, Wen Z-F, Cai X-Z. A systematic review of undisplaced femoral neck fracture treatments for patients over 65 years of age, with a focus on union rates and avascular necrosis. *J Orthop Surg* 2017;12:28.
- [4] Chiarello E, Tedesco G, Cadossi M, Capra P, Terrando S, Miti A, et al. Surgical prevention of femoral neck fractures in elderly osteoporotic patients. A literature review. *Clin Cases Miner Bone Metab* 2016;13:42–5.
- [5] Jiang J, Yang C, Lin Q, Yun X, Xia Y. Does arthroplasty provide better outcomes than internal fixation at mid- and long-term followup? A meta-analysis. *Clin Orthop* 2015;473:2672–9.
- [6] Wang F, Zhang H, Zhang Z, Ma C, Feng X. Comparison of bipolar hemiarthroplasty and total hip arthroplasty for displaced femoral neck fractures in the healthy elderly: a meta-analysis. *BMC Musculoskelet Disord* 2015;16:229.
- [7] Baker PN, Salar O, Ollivier BJ, Forward DP, Weerasuriya N, Moppett IK, et al. Evolution of the hip fracture population: time to consider the future? A retrospective observational analysis. *BMJ Open* 2014;4:e004405.
- [8] Thorngren K-G, Hommel A, Norrman PO, Thorngren J, Wingstrand H. Epidemiology of femoral neck fractures. *Injury* 2002;33:C1–7.
- [9] Torio CM, Moore BJ. National inpatient hospital costs: the most expensive conditions by payer. Agency for Healthcare Research and Quality 2013;Statistical Brief #204 <http://www.ncbi.nlm.nih.gov/books/NBK368492/>.
- [10] Blundell CM, Parker MJ, Pryor GA, Hopkinson-Woolley J, Bhonsle SS. Assessment of the AO classification of intracapsular fractures of the proximal femur. *J Bone Joint Surg Br* 1998;80:679–83.
- [11] Waldenström J. Fractures récentes du col femoral: traitement opératoire ou orthopédique. *J Chir* 1924;24:129.
- [12] Pauwels F. Der schenkelhalsbruch. ein mechanisches problem: grundlagen des heilungsvorganges, prognose und kausale therapie. Stuttgart: Ferdinand Enke Verlag 1935.
- [13] Turgut A, Kumbaracı M, Kalenderer Ö, İlyas G, Bacaksız T, Karapınar L. Is surgeons' experience important on intra- and inter-observer reliability of classifications used for adult femoral neck fracture? *Acta Orthop Traumatol Turc* 2016;50:601–5.
- [14] van Embden D, Roukema GR, Rhemrev SJ, Genelin F, Meylaerts SaG. The Pauwels classification for intracapsular hip fractures: is it reliable? *Injury* 2011;42:1238–40.
- [15] Garden RS. Low-angle fixation in fractures of the femoral neck. *Bone Joint J* 1961;43-B:647–63.
- [16] Aggarwal A, Singh M, Aggarwal AN, Bhatt S. Assessment of interobserver variation in Garden classification and management of fresh intracapsular femoral neck fracture in adults. *Chin J Traumatol* 2014;17:99–102.
- [17] Van Embden D, Rhemrev SJ, Genelin F, Meylaerts SaG, Roukema GR. The reliability of a simplified Garden classification for intracapsular hip fractures. *Orthop Traumatol Surg Res* 2012;98:405–8.
- [18] Beimers L, Kreder HJ, Berry GK, Stephen DJG, Schemitsch EH, McKee MD, et al. Subcapital hip fractures: the Garden classification should be replaced, not collapsed. *Can J Surg* 2002;45:411–4.
- [19] Müller ME, Nazarian S, Koch P, et al. The comprehensive classification of fractures of long bones. New York: Springer-Verlag; 1990.
- [20] Mahmoud SSS, Pearse EO, Smith TO, Hing CB. Outcomes of total hip arthroplasty, as a salvage procedure, following failed internal fixation of intracapsular fractures of the femoral neck: a systematic review and meta-analysis. *Bone Joint J* 2016;98-B:452–60.
- [21] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
- [22] Mallick A, Parker MJ. Basal fractures of the femoral neck: intra- or extracapsular. *Injury* 2004;35:989–93.
- [23] Massoud EIE. Fixation of basicervical and related fractures. *Int Orthop* 2010;34:577–82.