Subcapital hip fractures: the Garden classification should be replaced, not collapsed

Lijkele Beimers;* Hans J. Kreder, MD, MPH;†‡ Gregory K. Berry, MD;§ David J.G. Stephen, MD;†
Emil H. Schemitsch, MD;¶ Michael D. McKee, MD;¶ Susan Jaglal, PhD**

Objectives: To evaluate rater agreement for a simple 2-category classification of subcapital hip fractures versus the 4-category Garden classification and to determine the effect of clinician experience on the level of agreement. **Setting:** Sunnybrook and Women's Health Sciences Centre, Toronto, a level 1 trauma centre. **Method:** Eleven raters, with varying levels of clinical experience (3 fellowship-trained orthopedic surgeons, 4 clinical fellows and 4 residents), classified 34 pairs of anteroposterior and lateral radiographs of patients with subcapital hip fractures according to whether the fracture was stable (the fragments move as a unit) or unstable (the fragments move independently), and according to Garden's original 4-category classification. The exercise was repeated 1 month later. The radiographs were obtained from a fracture database to represent a wide spectrum of injury severity. **Outcome measures:** The level of agreement beyond chance, quantified by use of the SAV statistic of O'Connell and Dobson. **Results:** The most experienced raters demonstrated almost perfect inter- and intrarater agreement with respect to stable and unstable fractures (SAV > 0.80). The raters demonstrated only fair agreement for the Garden classification (mean SAV 0.64). Even junior clinicians demonstrated substantial agreement regarding fracture stability, with much lower scores for the Garden classification. Collapsing the Garden classification responses into 2 categories (stages I and II v. III and IV) was not synonymous with rater categorization of stable versus unstable. **Conclusion:** The Garden classification for subcapital hip fractures is unreliable and should be abandoned in favour of categorizing fractures as stable versus unstable.

Objectifs : Évaluer la mesure dans laquelle les évaluateurs s'entendent sur une classification simple à deux catégories des fractures sous-capitales de la hanche par rapport à la classification de Garden à quatre catégories et déterminer l'effet de l'expérience du clinicien sur le niveau de l'entente. **Contexte :** Sunnybrook and Women's Health Sciences Centre, Toronto, centre de traumatologie de niveau 1. **Méthode :** Onze évaluateurs dont l'expérience clinique variait (trois chirurgiens orthopédiques formés en fellowship, quatre chercheurs cliniciens et quatre résidents) ont classé 34 paires de radiographies antéropostérieures et latérales de patients ayant subi une fracture sous-capitale de la hanche, en fonction de la stabilité (fragments bougeant d'un seul bloc) ou de l'instabilité (fragments bougeant indépendamment) de la fracture, et selon la classification originale de Garden à quatre catégories. On a répété l'exercice un mois plus tard. On a tiré les radiographies d'une base de données sur les fractures afin qu'elles représentent un vaste éventail de gravités. **Mesures de résultats :** Le niveau d'entente au-delà du hasard, quantifié par l'utilisation de la statistique VRPO d'O'Connell et Dobson. **Résultats :** Les évaluateurs les plus chevronnés ont montré une uniformité presque parfaite entre évaluateurs et par rapport à eux-mêmes en ce qui concerne les fractures stables et instables (VRPO > 0,80). Ils ont montré une entente moyenne seulement en ce qui concerne la classification de Garden (VRPO moyen, 0,64). Même les cliniciens débutants ont montré une entente importante en ce qui concerne la stabilité des fractures et beaucoup moindre dans le cas de la classification de Garden. La réduction en deux caté-

From the *Faculty of Medical Sciences, University of Groningen, The Netherlands, the †Division of Orthopedic Surgery and the ‡Department of Health Policy Management & Evaluation, Sunnybrook and Women's College Health Sciences Centre, Toronto, the §Division of Orthopedic Surgery, Hôpital du Sacré-Coeur de Montréal, Montreal, Que., the ¶Division of Orthopedic Surgery, St. Michael's Hospital, Toronto, and the **Department of Physical Therapy, University of Toronto, Toronto, Ont.

Correspondence to: Dr. Hans J. Kreder, Division of Orthopedic Surgery, M4N 3M5, Sunnybrook and Women's College Health Sciences Centre, 2075 Bayview Ave., Toronto ON M4N 3M5; fax 416 480-5886; kreder@rogers.com

gories (stades I et II c. III et IV) des réponses fondées sur la classification de Garden n'était pas synonyme des catégories stables ou instables établies par les évaluateurs. **Conclusion :** La classification de Garden des fractures sous-capitales de la hanche n'est pas fiable et il faudrait la remplacer par une catégorisation fonction de la stabilité ou l'instabilité de la fracture.

Numerous classification systems have been proposed for fractures of the femoral neck. The Garden classification represents the most popular system currently used in clinical practice. This 4-stage system is based on the degree of displacement of the fracture as seen on the anteroposterior (AP) radiograph of the hip.[1]

Although the Garden classification is intended to aid clinical decision-making, recent studies have shown that trained clinicians could not reliably differentiate among the 4 stages.[2-5] Several authors have demonstrated better rater agreement when the Garden classification was collapsed into 2 categories.[4,5] Garden stages I and II fractures were considered "undisplaced" and Garden stages III and IV fractures were considered "displaced." Strict application of the Garden classification, however, can lead to confusion between the "displaced" and "undisplaced" categories since the trabecular pattern of the acetabulum and femoral head are lined up in both type II and type IV fractures on the AP view (use of the lateral film is not part of the original Garden classification). Unfortunately, including the lateral radiograph to assess the stage of the fracture does not necessarily improve the observer variation of the Garden classification.[5]

Whether or not the fracture fragments are completely separated and move independently (fracture stability) represents an important injury parameter for making treatment decisions. The purpose of this study was to evaluate the chance-corrected rater agreement of a simple 2-category classification, with explicit definition of stable and unstable subcapital hip fractures. A secondary objective was to compare the chance-corrected rater agreement of this 2-category classification and the original Garden classification, and to determine the effect of clinician experience on the level of agreement for both classifications.

Materials and methods

Thirty-four patients with subcapital hip fractures were chosen from a fracture database such that a full spectrum of injury was represented. The original preoperative AP and lateral radiographs that had been used for clinical decision-making were obtained, and each corresponding pair was numbered sequentially with all patient identifiers removed. Three attending staff members (fellowship-trained orthopedic surgeons), 4 fellows and 4 residents were asked to independently classify the 34 pairs of AP and lateral radiographs according to whether the fracture was stable or unstable. Prior to rating, definitions of stable and unstable subcapital hip fractures were given to all raters. Stable fractures were defined as having some continuity across the fracture site (i.e., fracture impaction) such that the 2 fragments would be expected to move as a unit with minimal force (Figs. 1 and 2). Unstable fractures were defined as having no continuity across the fracture site such that the 2 fragments would be expected to move independently with minimal force (Figs. 3 and 4). The raters were then asked to classify each fracture according to the Garden classification. The classification according to the original definitions proposed by Garden was reviewed prior to rating.[1] Each rater then reclassified the same radiographs 4 weeks later according to both classification schemes.

Statistics

The level of agreement beyond chance was quantified using the SAV statistic of O'Connell and Dobson.[6] We used the unweighted SAV statistic, which allows a direct comparison of classification systems with differing numbers of categories (i.e., the 4 categories of the Garden classification v. the 2 categories for the stable–unstable classification). The Koch and Landis criteria for interpretation of statistical scores were used to interpret the SAV values. Values between zero and 0.20 represent slight agreement, 0.21 and 0.40 fair agreement, 0.41 and 0.60 moderate agreement, 0.61 and 0.80 substantial agreement, and above 0.81 almost perfect agreement.[7]

Results

Attending staff surgeons showed almost perfect interobserver agreement for the 2-category classification (SAV = 0.81), compared with only fair agreement for the Garden classification (SAV = 0.33). A similar trend was noted for the SAV scores of residents and fellows (Table 1).

Intraobserver agreement for the 2-category classification with definition of stable and unstable was also very high. Attending staff surgeons had a SAV score of 0.83 (nearly perfect) for the 2-category classification, compared with 0.65 (substantial agreement) for the Garden classification. SAV scores for intraobserver agreement for residents and fellows were also higher for the 2-category classification than for the original Garden classification (Table 2).

More-experienced clinicians overall showed higher levels of intra- and interobserver agreement than junior clinicians for both classifications.

In 10 (1.3%) instances fractures categorized as "stable" were not classified as Garden stage I or II, and in 38 (5.1%) instances fractures categorized as "unstable" were not classified as Garden stage III or IV.

Discussion

The clinical usefulness of the Garden classification has been questioned. Authors have found poor observer agreement for the complete 4-stage classification.[2–5] Others have noted similar avascular necrosis and nonunion rates when grouping stages I and II fractures together and stages III and IV fractures together.[8–10] Eliasson and associates,[11] interpreting both AP and lateral projections of the hip radiographs, found that the anatomical displacement in stages I and II was almost equal and that the displacement of stage III did not differ from that of stage IV fractures.

The findings that 1.3% of stable fractures and 5.1% of unstable fractures were not categorized as Garden stage I or II or Garden stage III or IV, respectively, suggests that a simple collapse of the Garden classification into 2 groups is not the same as a separate classification of stable versus unstable fractures using the definitions outlined.

Several other classification systems

FIG. 1. Anteroposterior radiograph of a subcapital hip fracture typically classified as stable.

FIG. 2. Lateral radiograph of a subcapital hip fracture typically classified as stable.

FIG. 3. Anteroposterior radiograph of a subcapital hip fracture typically classified as unstable.
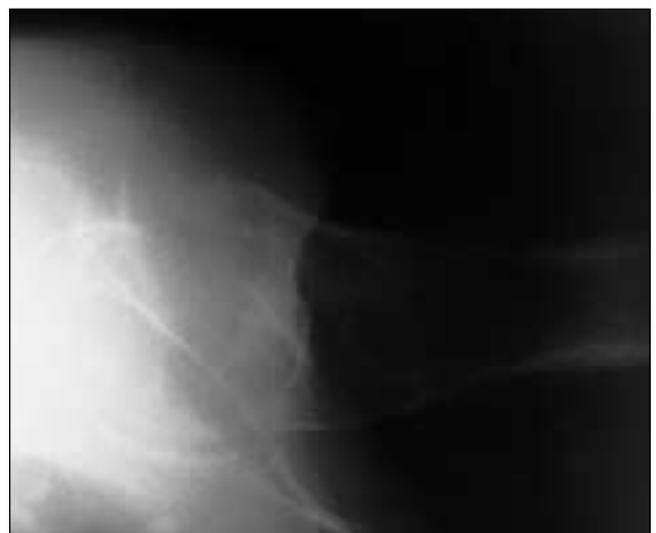

FIG. 4. Lateral radiograph of a subcapital hip fracture typically classified as unstable.

have been proposed for femoral neck fractures. The Pauwels classification divides fractures into 3 types based on the direction of the fracture line across the femoral neck. Pauwels[12] suggested that the more vertical the shear angle, the higher was the incidence of nonunion. Unfortunately, studies showed that neither the Pauwels angle nor the classification had any predictive value for the rate of nonunion or aseptic necrosis, although there have been some contradictory reports.[13,14] In addition, the fracture line varies in obliquity with rotation of the distal fragment and position of the X-ray beam, leading to misclassification.

Blundell and associates[15] investigated the AO classification of subcapital hip fractures and found that intra- and interobserver consistency was very poor and that the predictive value with respect to complications and outcome was limited. De Boeck[16] also investigated the AO classification. Of the 5 hips classified by 10

colleagues using this classification, no fracture was classified identically.

Our study shows that raters can reliably distinguish stable from unstable subcapital hip fractures using the definitions as outlined. Although agreement increases with experience, even junior clinicians can achieve moderate agreement for the simple 2-category classification, with stable and unstable definition. In contrast, agreement was only fair for the Garden classification due to different interpretations of the fracture radiographs by the raters.

For research purposes, correct classification may be necessary for selecting patients to enter prospective trials or for stratification purposes. Owing to poor agreement, the original Garden classification would not suit this purpose.

Conclusions

In order to be clinically useful, a classification should have at least moderate rater consistency. Given the relatively poor performance of the original 4-stage Garden classification, we recommend that it be abandoned in favour of a 2-stage classification into stable versus unstable for subcapital hip fractures, given the definitions as noted. For research purposes, classification of subcapital hip fractures should be undertaken by experienced clinicians in order to maximize agreement. Further studies will be needed to evaluate the relationship between patient outcome and stable versus unstable subcapital hip fractures.

Table 1

SAV Scores Showing Interobserver Consistency for Both the Stable Versus Unstable and Garden Classifications

	Classification	
Level of clinical experience	Stable v. unstable	Garden
Attending staff	0.81	0.33
Fellows	0.56	0.43
Residents	0.63	0.36
Mean	0.67	0.38

Table 2

SAV Scores Showing Intraobserver Consistency for Both the Stable Versus Unstable and Garden Classifications

	Classification	
Level of clinical experience	Stable v. unstable	Garden
Attending staff	0.83	0.65
Fellows	0.77	0.72
Residents	0.68	0.54
Mean	0.76	0.64

References

1. Garden RS. Low-angle fixation in fractures of the femoral neck. *J Bone Joint Surg Br* 1961;43:647-61

2. Frandsen PA, Andersen E, Madsen F, Skjødt T. Garden's classification of femoral neck fractures. An assessment of inter-observer variation. *J Bone Joint Surg Br* 1988;70(4):588-90.

3. McCabe J, O'Farrell D, O'Byrne J, O'Brien T. Observer variation in the radiographic classification of femoral neck fractures [abstract]. *J Bone Joint Surg Br* 1994;76 (Suppl II–III):120.

4. Scavenius M, Ibsen A, Rønnebech 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.

5. Thomsen NO, Jensen CM, Skovgaard N, Pedersen MS, Pallesen P, Soe-Nielsen NH, et al. Observer variation in the radiographic classification of fractures of the neck of the femur using Garden's system [see comment]. *Int Orthop* 1996;20(5):326-9. Comment in: *Int Orthop* 1997;21(6):416.

6. O'Connell DL, Dobson AJ. General observer-agreement measures on individual subjects and groups of subjects. *Biometrics* 1984;40:973-83.

7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-74.

8. Barnes R, Brown JT, Garden RS, Nicoll EA. Subcapital fractures of the femur. A prospective review. *J Bone Joint Surg Br* 1976;58(1):2-24.

9. Hogh J, Jensen J, Lauritzen J. Dislocated femoral neck fractures. A follow-up study of 98 cases treated by multiple AO (ASIF) cancellous bone screws. *Acta Orthop Scand* 1982;53(2):245-9.

10. Parker MJ. Garden grading of intracapsular fractures: meaningful or misleading? *Injury* 1993;24(4):241-2.

11. Eliasson P, Hansson LI, Karrholm J. Displacement in femoral neck fractures. A numerical analysis of 200 fractures. *Acta Orthop Scand* 1988;59(4):361-4.

12. Pauwels F. *Der Schenkelhalsbruch. Ein mechanisches Problem.* Stuttgart:: F. Enke; 1935.

13. Boyd HB, Salvatore JE. Acute fracture of the femoral neck: internal fixation or prosthesis? *J Bone Joint Surg Am* 1964;46: 1066-8.

14. Parker MJ, Dynan Y. Is Pauwels classification still valid? [review] *Injury* 1998;29 (7):521-3.

15. 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(4):679-83.

16. De Boeck H. Classification of hip fractures. *Acta Orthop Belg* 1994;60(Suppl 1): 106-9.