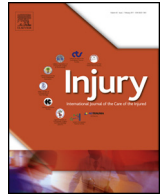




Contents lists available at ScienceDirect

Injury

journal homepage: www.elsevier.com/locate/injury



Reliability of the classification of proximal femur fractures: Does clinical experience matter?

Tom J. Crijns^a, Stein J. Janssen^b, Jacob T. Davis^c, David Ring^{a,*}, Hugo B. Sanchez^d,
the Science of Variation Group

Althausen Peter, Amini Michael H., Appleton Paul, Babis George C., Babst Reto H., Ballas Efstathios G., Barquet Antonio, Begue Thierry, Bishop Julius, Borris Lars C., Buckley Richard, Chesser Tim, Choudhari Pradeep, Cornell Charles, Crist Brett D., DeCoster Thomas A., Elias Nelson, Frihagen Frede, Garnavos Christos, Giordano Vincenzo, Haverlag R., Havlicek Tomo, Hurwit Shep, Ibrahim Edward F., Iyer Vishwanath M., Jenkinson Richard, Jeray Kyle, Kabir Koroush, Kanakaris Nikolaos K., Klostermann Cyrus, Kreder Hans J., Kreis B.E., Kristan Anze, Lygdas P., McGraw Iain, Mica Ladislav, Mirck B., Moreta-Suarez Jesus, Morgan Steven J., Nikolaou Vassilios S., Omara Timothy, Pesantez Rodrigo, Pirpiris Marinis, Poelhekke L.M.S.J., Pountos Ippokratis, Prayson Michael, Quell M., Rodríguez-Roiz Juan M., Satora Wojciech, Schandelmaier Peter, Schepers T., Short Nicholas L., Smith Raymond M., Spoor A.B., Stojkowska Pemovska Emilija, Swiontkowski Marc, Taitzman Lisa, Tosounidis Theodoros, Tyllianakis Minos, Van bergen C.J.A., Van de Sande M.A.J., Van Helden S.H., Verbeek Diederik O., Wascher Daniel C., Weil Yoram

^a Department of Surgery and Perioperative Care, Dell Medical School, Health Discovery Building 6.706, 1701 Trinity St., Austin, TX 78723, USA

^b Department of General Surgery, OLVG, Oosterpark 9, 1091 AC Amsterdam, The Netherlands

^c Department of Orthopaedic Surgery, JPS Health Network, 1500 S. Main St, Fort Worth, TX 76104, USA

^d Department of Orthopaedic Surgery, Acclaim Physician Group, Ben Hogan Center, 800 5th Ave, Suite 400, Fort Worth, TX 76104, USA

ARTICLE INFO

Article history:

Accepted 21 February 2018

Keywords:

AO/OTA classification
Interobserver agreement
Proximal femur fractures

ABSTRACT

Background: Radiographic fracture classification helps with research on prognosis and treatment. AO/OTA classification into fracture type has shown to be reliable, but further classification of fractures into subgroups reduces the interobserver agreement and takes a considerable amount of practice and experience in order to master.

Questions/purposes: We assessed: (1) differences between more and less experienced trauma surgeons based on hip fractures treated per year, years of experience, and the percentage of their time dedicated to trauma, (2) differences in the interobserver agreement between classification into fracture type, group, and subgroup, and (3) differences in the interobserver agreement when assessing fracture stability compared to classifying fractures into type, group and subgroup.

Methods: This study used the Science of Variation Group to measure factors associated with variation in interobserver agreement on classification of proximal femur fractures according to the AO/OTA classification on radiographs. We selected 30 anteroposterior radiographs from 1061 patients aged 55 years or older with an isolated fracture of the proximal femur, with a spectrum of fracture types proportional to the full database. To measure the interobserver agreement the Fleiss' kappa was

* Corresponding author at: Department of Surgery and Perioperative Care, Dell Medical School, Health Discovery Building 6.706, 1701 Trinity St., Austin, TX 78723, USA.
E-mail addresses: tom.crijns@austin.utexas.edu (T.J. Crijns), stein.janssen@austin.utexas.edu (S.J. Janssen), jdavis07@jpshealth.org (J.T. Davis), david.ring@austin.utexas.edu (D. Ring), hsanchez@jpshealth.org (H.B. Sanchez).

<https://doi.org/10.1016/j.injury.2018.02.023>

0020-1383/© 2018 Elsevier Ltd. All rights reserved.

determined and bootstrapping (resamples = 1000) was used to calculate the standard error, z statistic, and 95% confidence intervals. We compared the Kappa values of surgeons with more experience to less experienced surgeons.

Results: There were no statistically significant differences in the Kappa values on each classification level (type, group, subgroup) between more and less experienced surgeons. When all surgeons were combined into one group, the interobserver reliability was the greatest for classifying the fractures into type (kappa, 0.90; 95% CI, 0.83 to 0.97; $p < 0.001$), reflecting almost perfect agreement. When comparing the kappa values between classes (type, group, subgroup), we found statistically significant differences between each class. Substantial agreement was found in the clinically relevant groups stable/unstable trochanteric, displaced/non-displaced femoral neck, and femoral head fractures (kappa, 0.60; 95% CI, 0.53 to 0.67, $p < 0.001$).

Conclusions: This study adds to a growing body of evidence that relatively simple distinctions are more reliable and that this is independent of surgeon experience.

© 2018 Elsevier Ltd. All rights reserved.

Introduction

Proximal femur fractures are common among the elderly and notorious for their high rates of morbidity and mortality [1–3]. Radiographic fracture classification helps with clinical decision making, communication, and research on prognosis and treatment. There is evidence that the AO/OTA Classification of Fractures and Dislocations is more reliable for classification of trochanteric fractures of the proximal femur into fracture types and groups than other classification systems (e.g. Evans, Kyle, and Boyd) [4,5]. Further classification of fractures into subgroups reduces the interobserver agreement [5] and takes a considerable amount of practice and experience in order to master. We are interested whether the reliability of the AO/OTA classification system of proximal femur fractures differs when comparing more experienced to less experienced observers. In addition, we are interested in differentiating stable versus unstable and displaced versus non-displaced fractures due to its treatment implications.

We therefore assessed: (1) the difference in interobserver agreement between more and less experienced trauma surgeons based on: hip fractures treated per year, years of experience, and the percentage of their time dedicated to trauma, (2) the difference in interobserver agreement based on classification type (trochanter, femoral neck, and head), group, and subgroup, and (3) the difference in interobserver agreement when assessing fracture stability and displacement.

Materials and methods

Study design and setting

This cross-sectional survey study was exempt from review by our institutional review board. We selected 30 anteroposterior radiographs from 1061 patients aged 55 years or older with an isolated fracture of the proximal femur who presented to a single Level 2 trauma center between December 2010 and September 2013. Fractures were classified by three orthopedic surgeons who came to a consensus opinion of the classification using the AO/OTA fracture classification [6] and included: 31-A 1.1 (6.7%), 31-A1.2 (10%), 31-A1.3 (3.3%), 31-A2.1 (10%), 31-A2.2 (13%), 31-A2.3 (6.7%), 31-A3.1 (3.3%), 31-A3.3 (3.3%), 31-B1 (6.7%), 31-B2 (33%), and 31-B3 (3.3%).

Using a random number generator, we selected a patient in the database and from that patient onward, we consecutively added radiographs of each of the fracture subgroups until we had a representative number of each fracture type up to 30 cases in total. Once a certain fracture type had reached a number proportional to the distribution of that fracture in the database additional images of that fracture type were skipped to limit spectrum bias.

We used SurveyMonkey (Palo Alto, CA, USA), an online survey design tool to develop the survey including the 30 cases. Anteroposterior radiographs were shown and participants were asked: “What type of fracture is shown on the radiograph?” Each question had fifteen answer choices to a subgroup level (Table 2). Members of the Science Of Variation Group (SOVG) were sent an invitation to participate in this study. The SOVG has a list of emails of several hundred orthopaedic, trauma, and plastic surgeons, but only a subset participates regularly in the surveys, and even regular participants only respond to surveys in their region of expertise, so it is not possible to measure a meaningful response rate.

Surgeon characteristics

Sixty-five surgeons completed the survey, 62 (94%) men and the majority residing in Europe (60%) and North America (29%) (Table 1). The surgeons had a median experience of 18 years, with an interquartile range from 11 to 27 years. The median of hip fractures per year that was treated by these surgeons was 45, with an interquartile range from 25 to 75.

Statistical analysis

To measure the interobserver agreement the Fleiss' kappa was determined and bootstrapping (resamples = 1000) was used to calculate the standard error, z statistic, and 95% confidence intervals. The Kappa value reflects the degree to which observers agree with one another, taking into account the chance that observers agree at random. Perfect agreement is represented by a Kappa of 1, while agreement solely by chance would equate to a Kappa of 0. The classification of Landis and Koch [7] for categorical data is used frequently when interpreting Kappa values: a Kappa between 0.01 and 0.20 represents slight agreement, a value between 0.21 and 0.40 represents fair agreement, a value between 0.41 and 0.60 represents moderate agreement, between 0.61 and 0.80 represents substantial agreement, and greater than 0.81 represents almost perfect agreement.

We compared the Kappa values of surgeons with more experience (≥ 80 of their percentage of work dedicated to trauma; ≥ 17 years in practice; ≥ 50 hip fractures treated per year) to less experienced surgeons (< 80 of their percentage of work dedicated to trauma; < 17 years in practice; < 50 hip fractures treated per year). The definition of more or less experience was defined by splitting the continuous variables (years in practice and hip fractures treated per year) on the median (Table 1) and for the categorical variable, percentage of work dedicated to trauma, we strived to create two groups that were similar in the number of observers to maximize statistical power.

Download English Version:

<https://daneshyari.com/en/article/8718681>

Download Persian Version:

<https://daneshyari.com/article/8718681>

[Daneshyari.com](https://daneshyari.com)