


Does Computed Tomography Improve Reproducibility in the Classification of Transtrochanteric Fractures?*

A tomografia computadorizada melhora a reprodutibilidade na classificação das fraturas transtrocanterianas?

Murilo Alexandre¹ Giancarlo Cavalli Polesello^{1,2}  Edio Cavassani Neto¹
Nayra Deise dos Anjos Rabelo¹ Marcelo Cavalheiro de Queiroz^{1,2} Walter Ricioli Junior^{1,2}

¹ Hip Group, Orthopedics and Traumatology Department, Faculdade de Ciências Médicas da Santa Casa de São Paulo (FCMSCSP), Irmandade Santa Casa de Misericórdia, São Paulo, SP, Brazil

² School of Medical Sciences, Santa Casa de São Paulo, São Paulo, SP, Brazil

Address for correspondence Giancarlo Cavalli Polesello, MD, PhD, Rua Dr. Cesário Mota Júnior, 112, Bairro Vila Buarque, São Paulo, SP, 01221-020, Brazil (e-mail: dot.quadril@hotmail.com).

Rev Bras Ortop 2019;54:361–367.

Abstract

With the aging of the population, there was a significant increase in the prevalence of hip fractures, with high mortality rates, sequelae and expenses. Understanding the fracture profile and classifying it correctly is critical to define the appropriate treatment. Several radiographic classifications have been developed for transtrochanteric fractures, such as Tronzo, Evans-Jensen, AO and Boyd-Griffin, but their reproducibility is not always satisfactory. The present review aimed to elucidate whether the addition of computed tomography (CT) implies a greater reproducibility than simple radiography in the classification of transtrochanteric fractures, and whether this is a better examination to identify the fracture trait. A search was conducted in the PubMed, Lilacs, Scielo and Cochrane databases between July 2016 and June 2017, limited to the last 15 years. All retrospective, prospective and systematic reviews articles published in the English language, with evaluation of men and/or women, were considered for review. We have excluded case reports, studies that evaluated tomography or radiographs in isolation, and duplicate studies. The research presented 112 articles, of which 5 contemplated the proposed criteria. Reproducibility for the classification of transtrochanteric fractures presented variable results and was influenced by factors such as the type of classification, the use of the simplified or complete classification, the specialty of the evaluator, his experience, and the methodology proposed by the works. There are indications that there is benefit for the use of CT, especially for fractures considered unstable, but its use as a tool to ensure better reproducibility (intra- and interobserver) remains controversial and needs further studies.

Keywords

- hip fractures/ classification
- validation studies
- tomography x-ray computed
- radiography

Introduction

Transtrochanteric fractures are those between the base of the femoral neck up to 2.5 cm distal to the minor trochanter.¹ Although they may occur in young people after high-energy

* Work performed at the Hip Group, Orthopedics and Traumatology Department, Irmandade da Santa Casa de Misericórdia de São Paulo, São Paulo, SP, Brazil.

 Giancarlo Cavalli Polesello's ORCID is <https://orcid.org/0000-0002-95248871>.

received
March 6, 2018
accepted
October 30, 2018

DOI <https://doi.org/10.1055/s-0039-1693045>.
ISSN 0102-3616.

Copyright © 2019 by Sociedade Brasileira de Ortopedia e Traumatologia. Published by Thieme Revinter Publicações Ltda, Rio de Janeiro, Brazil

License terms



Resumo

Palavras-chave

- fraturas do quadril/ classificação
- estudos de validação
- tomografia computadorizada por raios x
- radiografia

Com o envelhecimento populacional, houve um aumento significativo da prevalência das fraturas do quadril, com alto índice de mortalidade, de sequelas, e alto custo. Compreender o perfil da fratura e classificá-la de forma correta é fundamental para definir o tratamento adequado. Diversas classificações radiográficas foram desenvolvidas para as fraturas transtrocanterianas, tais como as de Tronzo, de Evans-Jensen, de Boyd-Griffin e AO, porém sua reprodutibilidade nem sempre é satisfatória. O presente trabalho objetivou analisar se o acréscimo da tomografia computadorizada (TC) implica em maior reprodutibilidade do que a radiografia simples na classificação das fraturas transtrocanterianas e se esta é melhor para a identificação do traço de fratura. Foi realizada uma pesquisa nas bases de dados PubMed, Lilacs, Scielo e Cochrane entre julho de 2016 e junho de 2017, limitada aos últimos 15 anos. Todos os trabalhos retrospectivos, prospectivos e revisões sistemáticas publicados na língua inglesa, com avaliação de homens e/ou de mulheres, foram considerados para a revisão. Foram excluídos relatos de casos, estudos que avaliaram de forma isolada a TC ou radiografias e estudos duplicados. A pesquisa apresentou 112 artigos, dos quais 5 preencheram os critérios propostos. A reprodutibilidade para a classificação das fraturas transtrocanterianas apresentou resultados variáveis e influenciados por fatores como o tipo de classificação, o uso da classificação simplificada ou completa, a especialidade do avaliador, a experiência e a metodologia proposta pelos trabalhos. Há indícios de que há algum benefício para o uso da TC, sobretudo para fraturas consideradas instáveis, porém sua utilização como ferramenta para garantir uma melhor reprodutibilidade (intra- e interobservador) ainda permanece controversa e carece de mais estudos.

trauma, elderly people falling to the ground are the most affected.² Populational aging resulted in a significant increase in the prevalence of osteoporosis and of hip fractures, leading to higher mortality rates, sequelae, and associated costs³⁻⁷ when compared with the prevalence and the indexes of other fractures.^{3,8-10}

The diagnosis of transtrochanteric fractures is based on anamnesis and clinical evaluation; in an elderly individual, a low-energy trauma causes functional incapacity, shortening, and external rotation of the lower limb.¹¹ Radiographic evaluation in anteroposterior (AP) (with lower limbs in internal rotation and traction) and lateral (L) views of the hip confirms the diagnosis, as well as the characteristics of the fracture.¹¹

A good understanding of the fracture type and its correct classification are critical to define the most appropriate osteosynthesis procedure.¹² Thus, to be adequate for the daily clinical practice, the classification system must be simple, easy to apply, and present good intra- and interobserver reproducibility.¹³ In order to achieve this objective, several radiographic classifications have been developed, such as those by Boyd et al,¹⁴ by Tronzo,⁸ by Jensen,¹⁵ by Evans,¹⁶ and the AO classification.¹ However, in some cases, classificatory divergences impair intra- and interobserver reproducibility.¹⁷⁻¹⁹ These cases may require diagnostic methods for better identification and agreement to allow a good choice of treatment.^{2,17,18,20,21}

Computed tomography (CT) is a more advanced radiological technique that generates more detailed and complete images, which are useful for the evaluation of complex and articular fractures, such as those of the tibial pilon,^{22,24} of the tibial plateau,^{24,25} of the distal humeral,²⁶ and of the calca-

neus.²⁷ Following this principle, the use of a tomographic classification was proposed for transtrochanteric fractures, as suggested by Nakano.²⁸

The availability of CT and its use led to the possibility of better intra- and interobserver reproducibility, as well as of a greater diagnostic precision for routine practice.^{19,29-31} However, the literature is still controversial. In some studies, this technique is deemed dispensable, whereas others considered it superior when compared with radiography.^{28,32,33}

Therefore, we have performed a systematic review to determine if the addition of CT presents superior intra- and/or interobserver reproducibility compared with fracture classifications based on plain radiographies alone.

Materials and Methods

A query was performed at the Pubmed, Lilacs, Scielo and Cochrane databases between July 2016 and June 2017, limited to the last 15 years. The following descriptors were used: (*intertrochanteric OR hip*) AND *fracture* AND (*classification OR Tronzo OR AO OR Evans Jensen OR Boyd Griffin*) AND (*reproducibility OR validation studies OR reliability*) AND (*tomography OR CT*) AND (*x-ray OR radiography*). At the initial query, all papers presenting two or more terms in the title and an abstract were included for review.

After the review of the papers, references were crossed to ensure the inclusion of all potential studies. All retrospective, prospective, and systematic English-language review articles with male and/or female patients comparing the reproducibility of radiographic and tomographic analysis of

transtrochanteric fracture classifications were considered for evaluation. We excluded case reports, studies assessing CT or radiography in isolation, duplicate studies, and diagnostic studies of occult or stress fractures.

The following data were extracted from the included studies: name of the first author, year of publication, country, study design, study objective, sample size, number of evaluators, as well as their specialties, classification used, method of analysis, and results. These data were tabulated in Microsoft Excel version 2016 (Microsoft Corporation, Redmond, WA, USA) and analyzed descriptively. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) recommendation for systematic reviews was followed.³⁴

The reproducibility of the studies was interpreted according to the kappa coefficient, based on the guidelines proposed by Landis et al: values from 0.00 to 0.20 indicate bad reproducibility; from 0.21 to 0.40, reasonable; from 0.41 to 0.60, moderate; from 0.61 to 0.80, strong; and from 0.81 to 1.00, excellent reproducibility.³⁵

Results

Initially, the query resulted in 112 papers; after applying the inclusion and exclusion criteria, 5 papers were obtained. Published between 2003 and 2017, these papers present a

minimum of 30 and a maximum of 110 patients, with a mean age ranging from 59 to 85 years old. Due to the restricted number of papers, conflicts of interests were not considered as exclusion criteria (► Fig. 1).

All of the papers used the AO classification system, while two of them employed the Evans-Jensen classification, and only one used the Boyd Griffin system (as evidenced in table 1).

Three papers concluded that the use of CT has benefits for the correct classification of transtrochanteric fractures with good reproducibility, especially in the identification of comminution and of unstable traits.^{29,30,33} In contrast, two papers indicated that CT does not present better classification reproducibility when compared with plain radiography^{1,33} (► Table 1).

Discussion

The reproducibility of the classification of transtrochanteric fractures had variable results and was influenced by several factors, such as classification system, sample size, use of the AO classification in its simplified or complete form, the specialty of the evaluators (radiologist versus orthopedist), the experience of the evaluators (staff professionals versus residents), and the methodology proposed by each work. Thus, the discussion was based on questions about the various factors that could influence these results.

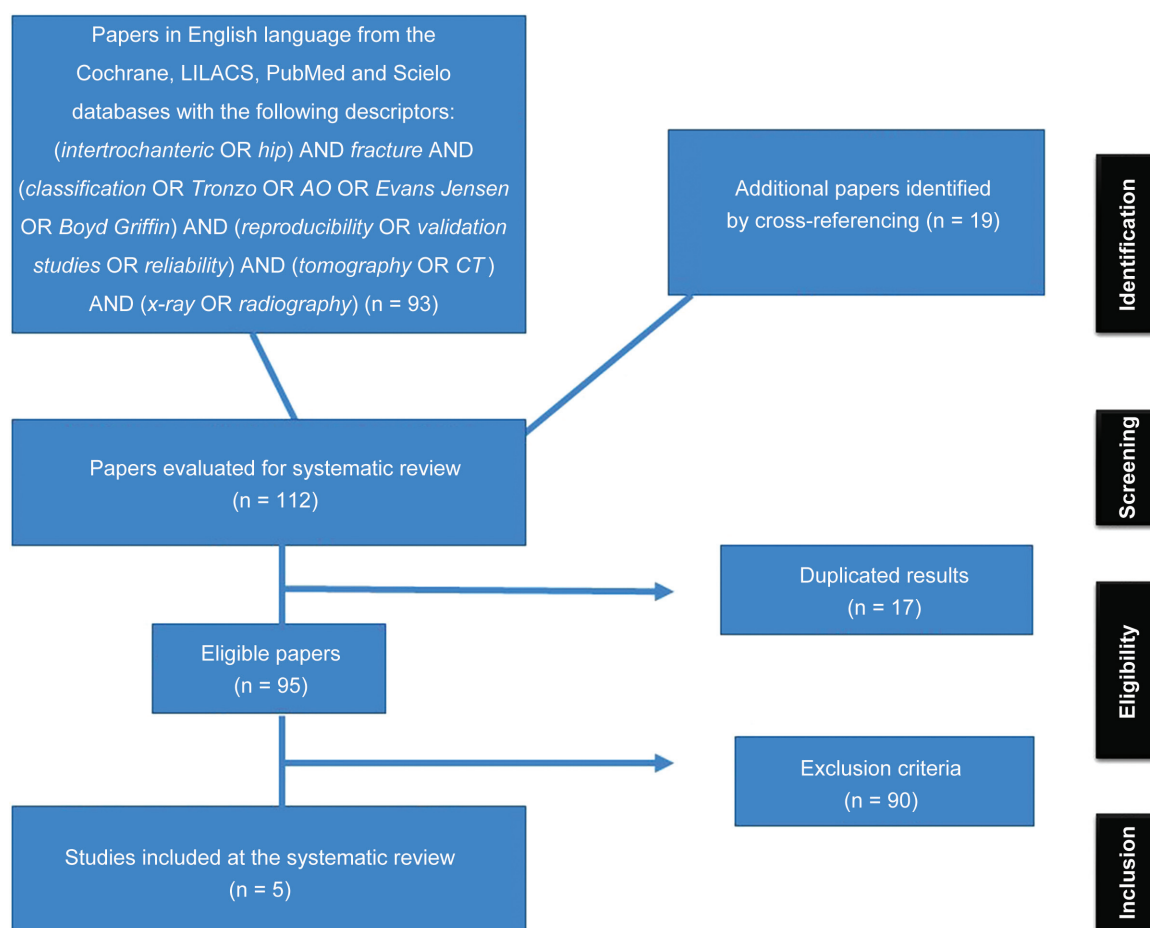


Fig. 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol for literature review.

Table 1 Systematic review result

Articles	Sample	Evaluators	Classification	Index	Result
Chapman CB et al, 2003. ³²	61 patients	3 orthopedists and 2 radiologists	AO ¹ e Evans-Jensen	kappa	No reproducibility improvement
Cavaignac E et al, 2013. ²⁰	53 patients	1 radiology resident and 1 radiologist	AO ¹ e Evans-Jensen	kappa	No reproducibility improvement
Mihir Thanvi et al, 2013. ²⁹	31 patients	2 orthopedists and 2 radiologists	AO ¹ e Boyd Griffing	kappa	Reproducibility improvement
Isida R et al, 2015. ³⁰	110 patients	1 orthopedist and 1 radiologist	AO ¹	kappa	Reproducibility improvement
Van Embden D et al, 2016. ³³	30 patients	4 orthopedists, 5 radiologists and 2 orthopedics residents	AO ²	kappa	Reproducibility improvement

1. Complete AO classification (with subgroups)

2. Simplified AO classification (without subgroups)

Which classification has better reproducibility?

The simplified AO classification had better reproducibility compared with other classifications, such as the complete AO classification and its subgroups, the Evans-Jensen and Boyd Griffin classifications,^{20,29,33} both for tomography and radiography. These results are probably due to the lower number of information in the simplified AO classification, which would make both their comprehension and memorization easier when compared with the complete classification.

When analyzing CT scans, Ito et al³⁶ obtained a 75% concordance with the complete AO classification, and a 65.9% concordance with the Evans-Jensen classification. Van Embden D. et al.³³ did not observe a higher reproducibility for the simplified AO classification when adding CT to the radiographic study, obtaining kappa values of 0.70 (strong) not using CT, and of 0.68 (strong) using CT images. Cavaignac E. et al.²⁰ observed a reasonable reproducibility with the addition of CT to the complete AO classification (0.28 to 0.33), whereas the reproducibility of the Evans-Jensen classification was moderate (0.50) with only radiography, and reasonable (0.35) using radiography associated with CT.

Isida R. et al.³⁰ found a strong reproducibility using CT and the complete AO classification (kappa value of 0.78), with a 100% positive predictive value, a 79% negative predictive value, and 95% specificity for comminution presence. When analyzed only by radiographs, the sensitivity was 48%, and the negative predictive value was 29%.

Does radiographic and tomographic standardization influence reproducibility?

Although CT has a well-documented importance in complex and intra-articular fracture detection,²²⁻²⁷ the results of transtrochanteric fractures may have been influenced by the divergences in the radiological techniques used.

Cavaignac E. et al.²⁰ and Shen et al³¹ analyzed AP and L hip radiographs in only 56% and 68% of the cases, respectively. This is an inadequate method, since Koval et al¹¹ previously demonstrated that the radiographic evaluation of transtrochanteric fractures should follow a pattern, obtaining two

radiographic views (AP and L) in all cases. It is undeniable that radiographic examinations are limited in trauma rooms and in patients with pain and difficulty in positioning, but the creation of protocols for test standardization is of paramount importance.¹²

The same occurs with CT, despite some standardization in its performance. A technique-related bias was observed by Cavaignac E et al.,²⁰ who used only axial CT sections, whereas Van Embden D. et al.³³ used three tomographic views (coronal, sagittal and axial), but no 3D reconstruction. An additional factor is the lack of specification of the technique employed in the other studies, such as the number of radiographs, radiographic and CT scan views, patient positioning, and the use of 3D reconstruction models.

It can be inferred that the limitations of the tests and inadequate execution may have impaired the evaluation and generated a bias in the reproducibility results found.

Which classification subgroup is subject to greater variation? Does the use of computed tomography influence the choice of the treatment?

With the simplified AO classification, unstable fractures (A3) examined by CT presented higher reproducibility than those analyzed by plain radiographies.³⁰

The study by Van Embden D. et al.³³ corroborates these results when CT is added to the analysis. In this study, type A3 fractures presented lower concordance and a greater number of surgical implant modifications when compared with implants chosen based only on plain radiographic evaluations³³ (►Figs. 2-3 to 4). It is worth mentioning that type A3 fractures present the highest complication rates, such as pseudoarthrosis and fixation failure in up to 32% of the cases.³³

Although type A3 fractures are subject to the largest changes in their classification, Van Embden D. et al.³³ showed that types A1 and A2 fractures also had a change in the choice of implants after performing CT scans in 11 of their 30 cases.³³ In addition, Shen et al³¹ showed that CT use resulted in a better understanding of the fracture, which promoted a statistically significant reduction in surgical time for intramedullary nailing placement, regardless of the classification.



Fig. 2 Anteroposterior and lateral views of the hip, AO classification 31 A1.



Fig. 3 Anteroposterior and lateral views of the hip, AO classification 31 A2.



Fig. 4 Anteroposterior and lateral views of the hip, AO classification 31 A3.

Therefore, adding CT scans to unstable fractures workup is plausible because it leads to a better understanding of the fracture trait, facilitates the proper choice of the implant, and is economically less costly compared with the expenses associated with increased surgical time and/or complications. However, the question is whether this examination would also not be beneficial for simple, stable fractures, due to the possibility of diagnostic inaccuracies and, consequently, to wrong choices of implants and to the occurrence of complications.

Does the specialty and/or experience of the evaluators influence reproducibility values?

One of the difficulties in evaluating these papers is the divergence in the type and level of the specialty of the evaluators. Two studies allowed the comparison between radiologists and orthopedists. Chapman CB. et al.³² demonstrated a strong reproducibility for radiologists (0.67) and a moderate reproducibility for orthopedists (0.57) in the complete AO and in the Evans-Jensen classifications. Mihir Thanvi et al.²⁹ presented a lower reproducibility among radiologists compared with orthopedists, both for the simplified and complete OA classifications using radiographic, tomographic, and tomographic analysis with 3D reconstruction.

Furthermore, as exemplified by Isida R. et al.,³⁰ there is a bias in reproducibility evaluation that is related to the specialty of the evaluators. Orthopedists achieved moderate reproducibility (0.45), whereas radiologists presented excellent reproducibility (0.94), but the former evaluated exclusively radiographies, and the latter only analyzed CT scans. So it does not allow us to conclude that the specialty influences reproducibility.

Another variable that could influence reproducibility is the experience of the evaluators. In this sense, there are studies with resident physicians presenting classification errors and lower reproducibility, especially in unstable fractures,^{20,33} as well as studies in which evaluation by less experienced professionals did not compromise the quality of the classification compared with more experienced personnel.²⁹

Van Embden D. et al.³³ supported this latter claim by demonstrating that residents achieved similar reproducibility rates as both trauma surgeons and radiologists. However, Cavaignac E. et al.,²⁰ while agreeing that less experienced evaluators do not compromise the study, demonstrate excellent interobserver reproducibility (0.85) for senior evaluators and strong reproducibility (0.79) for junior evaluators when CT was added.

Limitations

There is a large variation in the number of evaluators in the review papers (between 2 and 11 individuals), and in their specialties (radiologists, orthopedists, and residents from both specialties), setting different degrees of experience and of technical knowledge.

Discrepancies in sample size were also an important factor (30 to 110 patients). This isolated factor could justify variations in kappa values, since the confidence interval for

interobserver agreement depends directly on the sample size and on the number of evaluators, which are inversely proportional.³⁷ Finally, the use of different methods and of difficult standardizations compromised the data analysis and the interpretation of the results.

Final Considerations

There is evidence that CT use is beneficial, especially for fractures considered unstable. Routine CT addition as a tool to ensure better reproducibility (both intra- and interobserver), as well as for decision-making in transtrochanteric fracture treatment, remains controversial. We believe that it is necessary to develop studies with a better level of evidence and equivalent methodologies to elucidate the benefits of CT as a tool for the classification of transtrochanteric fractures.

Conflicts of interests

The authors have no conflicts of interests to declare.

References

- Müller ME, Nazarian S, Koch P. Classification AO des fractures. Berlin: Springer-Verlag; 1987
- Pervez H, Parker MJ, Pryor GA, Lutchman L, Chirodian N. Classification of trochanteric fracture of the proximal femur: a study of the reliability of current systems. *Injury* 2002;33(08):713–715
- Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. *J Bone Miner Res* 2007;22(03):465–475
- Lenich A, Vester H, Nerlich M, Mayr E, Stöckle U, Füchtmeier B. Clinical comparison of the second and third generation of intramedullary devices for trochanteric fractures of the hip—Blade vs screw. *Injury* 2010;41(12):1292–1296
- Andruszkow H, Frink M, Frömke C, et al. Tip apex distance, hip screw placement, and neck shaft angle as potential risk factors for cut-out failure of hip screws after surgical treatment of intertrochanteric fractures. *Int Orthop* 2012;36(11):2347–2354
- Parker M, Johansen A. Hip fracture. *BMJ* 2006;333(7557):27–30
- Hip fracture: management. Guidance and guidelines | NICE [accessed in Feb 9, 2018]. Disponível: <https://www.nice.org.uk/guidance/cg124>
- Tronzo RG. Symposium on fractures of the hip. Special considerations in management. *Orthop Clin North Am* 1974;5(03):571–583
- Yang S, Liu Y, Yang T, Zou J, Yang H. Early Clinical Efficacy Comparison Study of Gamma3 Nail, Percutaneous Compression Plate (PCCP) and Femoral Head Replacement (FHR) Treatment on Senile Unstable Intertrochanteric Fractures. *J Invest Surg* 2018;31(02):130–135
- Shin YS, Chae JE, Kang TW, Han SB. Prospective randomized study comparing two cephalomedullary nails for elderly intertrochanteric fractures: Zimmer natural nail versus proximal femoral nail antirotation II. *Injury* 2017;48(07):1550–1557
- Koval KJ, Oh CK, Egol KA. Does a traction-internal rotation radiograph help to better evaluate fractures of the proximal femur? *Bull NYU Hosp Jt Dis* 2008;66(02):102–106
- Palm H, Jacobsen S, Sonne-Holm S, Gebuhr P; Hip Fracture Study Group. Integrity of the lateral femoral wall in intertrochanteric hip fractures: an important predictor of a reoperation. *J Bone Joint Surg Am* 2007;89(03):470–475
- Fung W, Jonsson A, Bühren V, Bhandari M. Classifying intertrochanteric fractures of the proximal femur: does experience matter? *Med Princ Pract* 2007;16(03):198–202

- 14 Boyd H, Griffin L. Classification and treatment of trochanteric fracture. *Arch Surg* 1949;58(06):553–566
- 15 Jensen JS. Classification of trochanteric fractures. *Acta Orthop Scand* 1980;51(05):803–810
- 16 Evans EM. The treatment of trochanteric fractures of the femur. *J Bone Joint Surg Br* 1949;31B(02):190–203
- 17 Jin WJ, Dai LY, Cui YM, Zhou Q, Jiang LS, Lu H. Reliability of classification systems for intertrochanteric fractures of the proximal femur in experienced orthopaedic surgeons. *Injury* 2005;36(07):858–861
- 18 Schipper IB, Steyerberg EW, Castelein RM, van Vugt AB. Reliability of the AO/ASIF classification for pertrochanteric femoral fractures. *Acta Orthop Scand* 2001;72(01):36–41
- 19 Urrutia J, Zamora T, Besa P, Zamora M, Schweitzer D, Klaber I. Inter and intra-observer agreement evaluation of the AO and the Tronzo classification systems of fractures of the trochanteric area. *Injury* 2015;46(06):1054–1058
- 20 Cavaignac E, Lecoq M, Ponsot A, et al. CT scan does not improve the reproducibility of trochanteric fracture classification: a prospective observational study of 53 cases. *Orthop Traumatol Surg Res* 2013;99(01):46–51
- 21 van Embden D, Rhemrev SJ, Meylaerts SA, Roukema GR. The comparison of two classifications for trochanteric femur fractures: the AO/ASIF classification and the Jensen classification. *Injury* 2010;41(04):377–381
- 22 Zwipp H, Amlang M. [Treatment of fractures of the ankle in the elderly]. *Orthopade* 2014;43(04):332–338
- 23 Haapamaki VV, Kiuru MJ, Koskinen SK. Ankle and foot injuries: analysis of MDCT findings. *AJR Am J Roentgenol* 2004;183(03):615–622
- 24 Chan PS, Klimkiewicz JJ, Luchetti WT, et al. Impact of CT scan on treatment plan and fracture classification of tibial plateau fractures. *J Orthop Trauma* 1997;11(07):484–489
- 25 Mustonen AO, Koskinen SK, Kiuru MJ. Acute knee trauma: analysis of multidetector computed tomography findings and comparison with conventional radiography. *Acta Radiol* 2005;46(08):866–874
- 26 Doornberg J, Lindenhovius A, Kloen P, van Dijk CN, Zurakowski D, Ring D. Two and three-dimensional computed tomography for the classification and management of distal humeral fractures. Evaluation of reliability and diagnostic accuracy. *J Bone Joint Surg Am* 2006;88(08):1795–1801
- 27 Humphrey CA, Dirschl DR, Ellis TJ. Interobserver reliability of a CT-based fracture classification system. *J Orthop Trauma* 2005;19(09):616–622
- 28 Nakano T. Proposal for understanding trochanteric femoral fractures in the elderly and classification using 3DCT. *Orthopaedics (Glendale Calif)* 2006;19(05):39–45
- 29 Mihir Thanvi, Gupta AK, Goyal N, Agarwal R, Sharma C. Reliability of commonly used classification systems of trochanteric fractures of the proximal femur and evaluation of CT in improving the interobserver agreement. *Sch J App Med Sci* 2013;1(06):706–709
- 30 Isida R, Bariatsky V, Kern G, Dereudre G, Demondion X, Chantelot C. Prospective study of the reproducibility of X-rays and CT scans for assessing trochanteric fracture comminution in the elderly: a series of 110 cases. *Eur J Orthop Surg Traumatol* 2015;25(07):1165–1170
- 31 Shen J, Hu F, Zhang L, Tang P, Bi Z. Preoperative classification assessment reliability and influence on the length of intertrochanteric fracture operations. *Int Orthop* 2013;37(04):681–687
- 32 Chapman CB, Herrera MF, Binenbaum G, et al. Classification of intertrochanteric fractures with computed tomography: a study of intraobserver and interobserver variability and prognostic value. *Am J Orthop* 2003;32(09):443–449
- 33 van Embden D, Scheurkogel MM, Schipper IB, Rhemrev SJ, Meylaerts SA. The value of CT compared to radiographs in the classification and treatment plan of trochanteric fractures. *Arch Orthop Trauma Surg* 2016;136(08):1091–1097
- 34 Jüni P, Egger M. PRISMA reporting of systematic reviews and meta-analyses. *Lancet* 2009;374(9697):1221–1223
- 35 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(01):159–174
- 36 Ito R, Obara S, Atsumi T. Radiographic analysis of preoperative and postoperative 3dct images of trochanteric femoral fractures. *Showa Univ J Med Sci* 2014;26(04):271–282
- 37 Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol* 2012;65(07):778–784