

Article

Computed Tomography Does Not Improve Intra- and Interobserver Agreement of Hertel Radiographic Prognostic Criteria

Paulo Ottoni di Tullio¹, Vincenzo Giordano^{1,*} , William Dias Belangero², Robinson Esteves Pires³ , Felipe Serrão de Souza¹, Pedro José Labronici⁴, Caio Zamboni⁵, Felipe Malzac¹, Paulo Santoro Belangero⁶, Roberto Yukio Ikemoto⁷, Sergio Rowinski⁸ and Hilton Augusto Koch⁹

- ¹ Serviço de Ortopedia e Traumatologia Prof. Nova Monteiro—Hospital Municipal Miguel Couto, Rio de Janeiro 22430-160, Brazil
 - ² Departamento de Ortopedia, Reumatologia e Traumatologia—Universidade Estadual de Campinas (UNICAMP), Campinas 13083-970, Brazil
 - ³ Departamento do Aparelho Locomotor—Universidade Federal de Minas Gerais (UFMG), Belo Horizonte 31270-901, Brazil
 - ⁴ Departamento de Cirurgia Geral e Especializada—Universidade Federal Fluminense (UFF), Niteroi 24220-900, Brazil
 - ⁵ Departamento de Ortopedia—Santa Casa de São Paulo, São Paulo 01221-020, Brazil
 - ⁶ Departamento de Ortopedia e Traumatologia—Escola Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), São Paulo 04021-001, Brazil
 - ⁷ Grupo de Ombro e Cotovelo—Faculdade de Medicina do ABC, Santo André 09060-870, Brazil
 - ⁸ SUORT—Clínica Integrada, São Paulo 01258-010, Brazil
 - ⁹ Departamento de Radiologia—Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro 21941-901, Brazil
- * Correspondence: v_giordano@me.com; Tel.: +55-(21)-99751-6859



Citation: Tullio, P.O.d.; Giordano, V.; Belangero, W.D.; Pires, R.E.; de Souza, F.S.; Labronici, P.J.; Zamboni, C.; Malzac, F.; Belangero, P.S.; Ikemoto, R.Y.; et al. Computed Tomography Does Not Improve Intra- and Interobserver Agreement of Hertel Radiographic Prognostic Criteria. *Medicina* **2022**, *58*, 1489. <https://doi.org/10.3390/medicina58101489>

Academic Editors: Ivo Dumić-Čule and Tomislav Čengić

Received: 17 September 2022

Accepted: 17 October 2022

Published: 19 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Background and Objectives:* Proximal humerus fractures are the second most frequent site of avascular necrosis (AVN), occurring in up to 16% of cases. The Hertel criteria have been used as a reference for the prediction of humerus head ischemia. However, these are based solely on the use of radiographs, which can make interpretation extremely difficult due to several reasons, such as the overlapping fragments, severity of the injury, and noncompliant acute pain patients. The objectives of the study were to evaluate the role of computed tomography (CT) in the interpretation of the Hertel criteria and to evaluate the intra- and interobserver agreement of orthopedic surgeons, comparing their area of expertise. *Materials and Methods:* The radiographs and CT scans of 20 skeletally mature patients who had fractures of the proximal humerus were converted to jpeg and mov, respectively. All images were evaluated by eight orthopedic surgeons (four trauma surgeons and four shoulder surgeons) in two different occasions. The intra- and interobserver agreement was assessed by using the Kappa coefficient. The level of significance was 5%. *Results:* There was a weak-to-moderate intraobserver agreement ($\kappa < 0.59$) for all examiners. Only the medial metaphyseal hinge greater than 2 mm was identified by 87.5% of evaluators both in the radiographic and CT examinations in the two rounds of the study ($p < 0.05$). There was no significant interobserver agreement ($\kappa < 0.19$), as it occurred only in some moments of the second round of evaluation. *Conclusions:* The prognostic criteria for humeral head ischemia evaluated in this study showed weak intra- and interobserver agreement in both the radiographic and tomographic evaluation. CT did not help surgeons in the primary interpretation of Hertel prognostic criteria used in this study when compared to the radiographic examination.

Keywords: proximal humerus fracture; humerus head necrosis; avascular necrosis risk factors; posttraumatic avascular necrosis; Hertel criteria

1. Introduction

Fractures of the proximal humerus represent 4 to 5% of all fractures, being the third most common in the human body [1]. Approximately 85% of patients can be treated non-surgically, particularly older patients with fractures involving the surgical neck of the humerus [2]. However, fractures considered more complicated require surgical treatment, especially in younger patients. Indeed, about 13 to 16% of all fractures of the proximal humerus are in three, four, or more parts, including the humerus head, greater tuberosity, lesser tuberosity, and shaft, and present unacceptable displacements [3]. Although a good functional outcome has been reported both after non-surgical treatment and after internal fixation [4], several postoperative complications have been described as a result of proximal humerus fractures, including avascular necrosis (AVN) of the humeral head [5].

Regardless of the type of treatment, proximal humerus fractures are the second most frequent site of AVN, occurring in up to 16% of cases [6]. The main risk factors for AVN are a greater number of fragments, head-split fracture, short segment of the humeral calcar, rupture of the medial hinge, displaced tuberosities, glenohumeral fracture/dislocation, and significant angular displacement of the head [7,8]. Other factors, such as the surgical approach and poor anatomical reduction, have also been implicated as risk factors [8]. Hertel et al. [7] radiographically evaluated 100 fractures of the proximal humerus in 98 patients who underwent internal fixation over a period of four years. These authors observed that good predictors of ischemia were posteromedial metaphyseal extension of the head less than 8.0 mm, medial hinge rupture (>2.0 mm), and fracture patterns involving the anatomical neck. The combination of these three factors led to a positive predictive value of 97% for the development of AVN of the humeral head.

Shortly after its publication, other studies from the same group observed high intra- and interobserver reliability of the Hertel classification, providing a more adequate description of proximal humerus fractures compared to the systems described by the Neer system and the AO group (Arbeitsgemeinschaft für Osteosynthesefragen) [9,10]. However, there are some potential confounding factors in the study by Hertel et al. [7], such as the use of radiographs alone and the adoption of deltopectoral approach in all surgical procedures. In particular, the use of radiographs can make interpretation extremely difficult due to the overlapping fragments, severity of the injury, and noncompliant acute pain patients, making it impossible to perform all the shoulder standard views and not allowing substantial agreement [11].

We hypothesized that the adoption of computed tomography (CT) images, including three-dimensional reconstruction (3D CT), to radiographs of the proximal humerus in patients with fractures increases the reproducibility of the Hertel criteria, thus improving intra- and interobserver agreement among orthopedic surgeons. The objective of the study was to evaluate the role of CT in the interpretation of the Hertel criteria, using the radiographic evaluation of the same cases as a standard to identify the criteria described by these authors as prognostics for AVN of the humeral head.

2. Materials and Methods

Radiographs and CT scans of 20 skeletally mature patients who had fractures of the proximal humerus treated at the Institution between January and December 2020 were selected. Patients were chosen at random. We included patients older than 18 years, with a confirmed diagnosis of proximal humerus fractures, and who signed the Informed Consent form. We excluded patients with missing demographic data; inaccurate imaging exams; and presenting a pathological fracture or previous fracture history, surgical history, congenital deformities, or degenerative changes at the same region.

Patients' privacy and security during the acquisition, storage, and transmission of their medical information were protected. The identity of the patients was not revealed. All patients were surgically treated and had standard preoperative radiographic and CT studies. The age ranged from 52 to 70 years, with a mean of 59.4 (SD \pm 6.21882) years. Fifteen (75%) patients were female, and the right side was affected in 12 (60%) cases. The

main injury mechanism was a fall to the ground in 17 (85.0%) cases (Table 1). The imaging evaluation comprised a true AP (Grashey), scapular Y, and axillary view [12]. In patients with limited abduction of the glenohumeral joint due to pain or joint incongruity, the modified axillary view was performed [13]. The CT evaluation comprised 5 mm axial, coronal, and sagittal slices and a 3D reconstruction. Both radiographic and tomographic images were kept anonymous for the purpose of this study.

Table 1. Patient demographics and STANDARD assessment (*n* = 20).

Case	Age (Years)	Gender	Side	Mechanism of Trauma	Neer (Parts)	Hertel					
						X-ray C1	X-ray C2	X-ray C3	CT C1	CT C2	CT C3
1	58	M	R	MCA	III (H, GT, S)	P	A	A	P	A	A
2	53	F	R	Fall to the ground	II (GT)	A	A	A	A	A	A
3	52	F	R	Fall to the ground	III (H, GT, S)	P	P	A	P	P	A
4	68	M	R	Fall to the ground	II (LT)	P	P	A	P	P	A
5	57	F	L	Fall to the ground	IV	A	A	I	A	A	A
6	65	F	L	Fall to the ground	IV	I	A	A	A	A	A
7	55	F	L	Fall to the ground	IV	A	A	I	A	A	P
8	64	F	R	Fall to the ground	IV	P	P	A	P	P	A
9	67	F	R	Fall to the ground	IV	A	A	I	A	A	P
10	59	M	L	MCA	III (H, GT, S)	P	A	A	P	A	A
11	53	M	R	MVA	III (H, GT, S)	P	A	A	P	A	A
12	56	F	L	Fall to the ground	II (H)	P	A	A	P	P	A
13	52	F	R	Running over	III (H, GT, S)	P	P	A	P	P	A
14	70	F	R	Fall to the ground	IV	P	A	A	P	A	A
15	66	F	L	Fall to the ground	II (LT)	P	P	A	P	P	A
16	58	F	R	Fall to the ground	III (H, GT, S)	I	A	A	A	A	A
17	54	F	L	Fall to the ground	II (LT)	P	P	A	A	A	A
18	69	M	R	Fall to the ground	II (H)	P	A	A	P	A	A
19	53	F	L	Fall to the ground	II (H)	A	A	A	A	A	A
20	59	F	R	Fall to the ground	III (H, GT, S)	A	A	A	A	A	A

Source: SOT-Nova, HMMC, 2022. Legends: M—male; F—female; L—left; R—right; II—two parts; III—three parts; IV—four or more parts; H—head; GT—greater tuberosity; LT—lesser tuberosity; S—shaft; X-ray—radiography; CT—computed tomography; C1—Criterion A; C2—Criterion B; C3—Criterion C; P—present; A—absent; I—inconclusive; MCA: motorcycle accident.

Fractures were classified radiographically, according to the Neer system [12,14]. In addition, three of the Hertel criteria were evaluated on radiographs and on CT (Figure 1). All images were initially evaluated independently by two consultants who were fellowship-trained in orthopedic trauma and with more than 10 years of experience. In cases of disagreement, a senior consultant (with more than 20 years of experience in orthopedic trauma) evaluated the images. There were only 2 cases of disagreement, Cases 10 and 11. Both were further classified after the third evaluator as Neer 3-parts (head, greater tuberosity, and shaft), with medial metaphyseal extension < 8 mm observed on both the X-ray and CT. The standard (namely STANDARD) was considered when an absolute agreement was encountered between at least two evaluators.

The Standards for Reporting of Diagnostic Accuracy (STARD) questionnaire was used to determine surgeons’ perception of their accuracy in identifying each of the three criteria [15,16]. The selection was evidence-based whenever possible; therefore, the “inconclusive” category was added. Thus, both the evaluators defined as STANDARD and the respondents invited to participate in the study were asked to answer whether each of these three criteria was “present”, “absent” or “inconclusive” (cannot be evaluated) in the radiographic and tomographic examinations.

The radiographic images of the 20 cases were extracted from the DICOM (Digital Imaging and Communications in Medicine) disk, which was developed by the American College of Radiology (ACR) and the National Electrical Manufacturers Association (NEMA—Arlington, VA, USA); converted to jpeg (Joint Photographic Experts Group) with 1000 × 1000 pixels and 300 DPI resolution; and stored case by case in individual folders, from 01 to 20 [17]. In the same way, the tomographic images of the 20 cases were extracted from the DICOM disk, recorded as individual frames, and saved in the mov (Multimedia

Container File) format, a file name extension for the QuickTime multimedia file format (Apple Inc., Cupertino, CA, USA), with a resolution of 1228 × 657 ppi [18]. Figures 2–4 illustrate some of the cases used in the study.

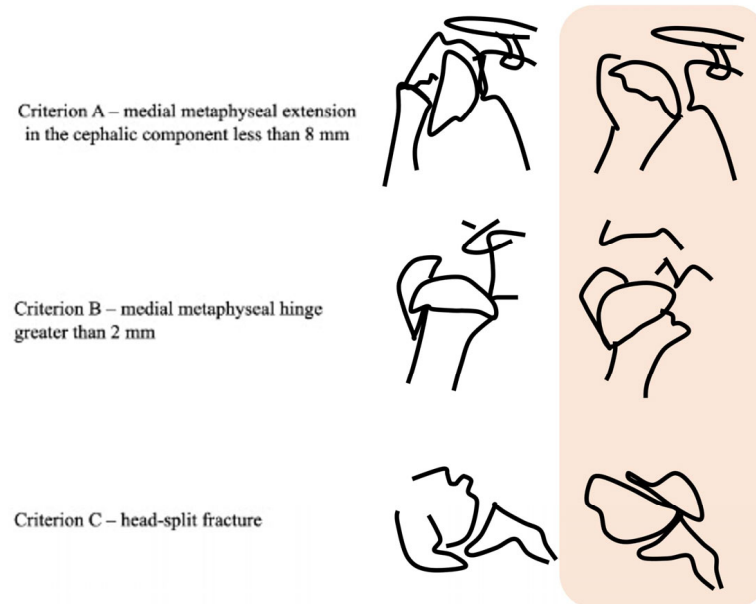


Figure 1. The three Hertel criteria adopted in the study. Criterion A represents the medial metaphyseal extension < 8 mm; Criterion B represents the medial metaphyseal hinge < 2 mm; and Criterion C represents the humerus head-split fracture. (Image produced on computer and from the author’s personal archive VG).

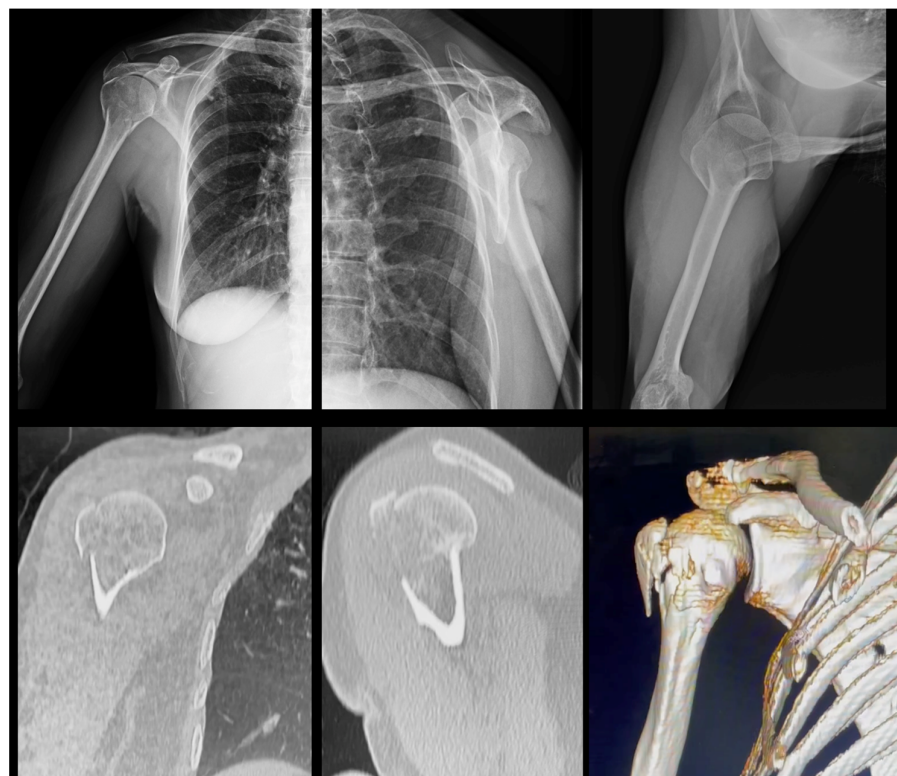


Figure 2. Radiographic and CT images of a 53-year-old female patient with a right proximal humerus fracture (Case 2). Note that none of Hertel’s prognostic criteria is present in either of the two imaging exams. The fracture was classified as Neer 2-parts (greater tuberosity).

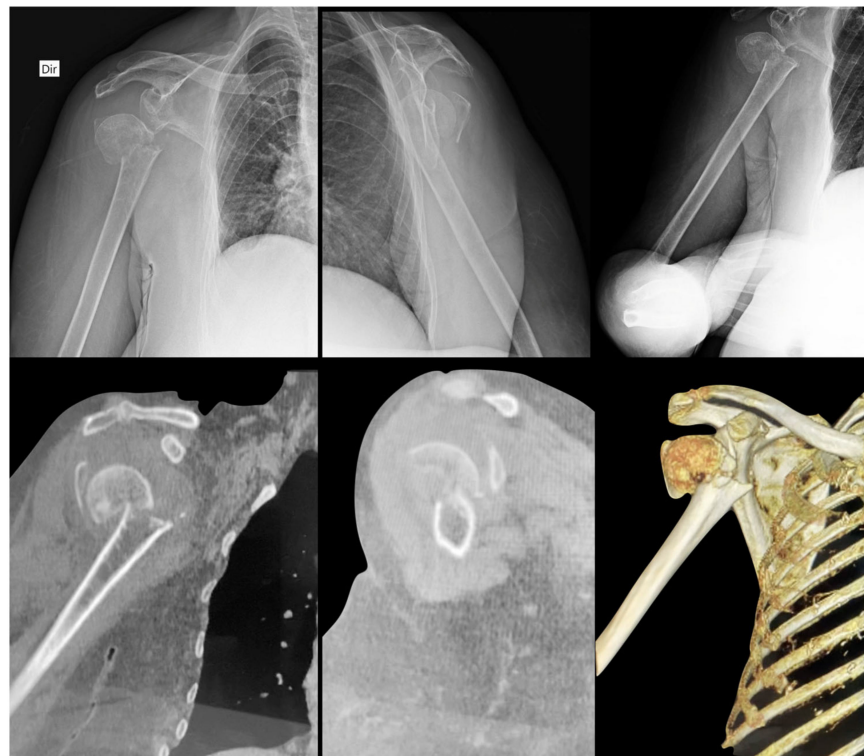


Figure 3. Radiographic and CT images of a 52-year-old female patient with a right proximal humerus fracture (Case 3). Note that the medial metaphyseal extension in the cephalic component less than 8 mm (Criterion A) and the medial metaphyseal hinge greater than 2 mm (Criterion B) are present and easily identified in both imaging exams. The fracture was classified as Neer 3-parts (head, greater tuberosity, and shaft).

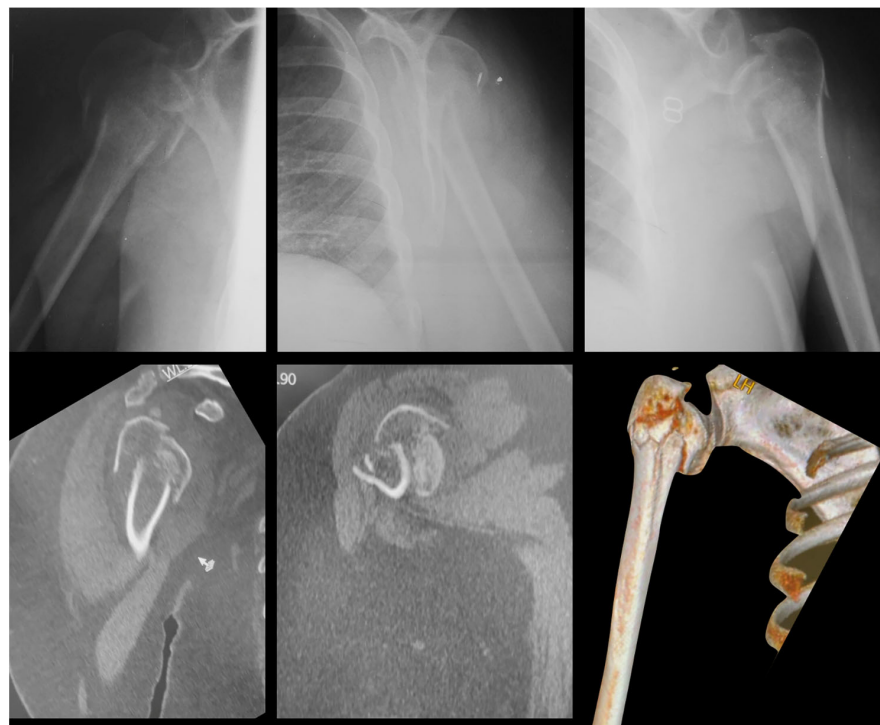


Figure 4. Radiographic and CT images of a 67-year-old female patient with a right proximal humerus fracture (Case 9). Note that it is possible to observe the split-head fracture of the head (Criterion C) only on CT, especially in 3D reconstruction. The fracture was classified as Neer 4-parts.

The images were sent by email to eight board-certified orthopedic surgeons with more than 10 years of experience in the treatment of proximal humerus fractures (Observers 1–8), in 20 individual folders, along with instructions on how to respond (Appendix A). All possible data that could identify the patients, such as name, initials, and date of birth, were removed from the exams so that their confidentiality was fully preserved. Four surgeons with fellowship in orthopedic trauma (namely TRAUMA 1 to TRAUMA 4) and four surgeons with fellowship in shoulder and elbow (namely, SHOULDER 1 to SHOULDER 4) were invited. Surgeons were asked to determine whether each of the Hertel prognostic criteria used in the study was “present”, “absent”, or “inconclusive” on radiographic and CT imaging. Respondents received the illustration of the prognostic criteria used in the study. They were asked to start with the radiographic images and not store images on their computers. Images were evaluated on two occasions, separated by 15 days—Round 1 (R1) and Round 2 (R2). For R2, the 20 folders were randomized and sent back to respondents by email, with the same guidelines as in R1.

The inferential analysis consisted of the Kappa coefficient (κ) to assess intraobserver agreement in the criteria (positive, negative, and inconclusive) according to radiographic and tomographic images [19]. Landis and Koch [19] suggest the following interpretation: $\kappa < 0.19$ —no/very poor agreement; $0.20 < \kappa < 0.39$ —weak agreement; $0.40 < \kappa < 0.59$ —moderate agreement; $0.60 < \kappa < 0.79$ —good agreement; and $\kappa \geq 0.80$ —very good (excellent) agreement. The interobserver agreement was defined by the level of concordance (positive/negative/inconclusive) in the cases divided by the total number of cases ($n = 20$) and multiplied by 100 (% of concordant cases). The interobserver agreement of the eight surgeons and the STANDARD was provided by four percentages of concordant cases (positive/negative/inconclusive/general). The significance level adopted was 5%. Statistical analysis was performed by using SPSS version 26 software (IBM, New York, NY, USA).

3. Results

3.1. Intraobserver Agreement

There was a weak-to-moderate intraobserver agreement for all examiners. For analytical consistency, the percentage of agreement was judged useful only when the Kappa coefficient was significant (Supplementary Materials Table S1). Only the medial metaphyseal hinge greater than 2 mm (Criterion B) was identified by all evaluators both in the radiographic and CT examinations in the two rounds of the study ($p < 0.05$), except by the evaluator SHOULDER 4. Both Criteria A and C were identified significantly less frequently in the two rounds of the study, with no difference between radiographic and tomographic evaluations ($p > 0.05$). The SHOULDER 4 evaluator showed no intraobserver agreement for any of the criteria evaluated in R1 and R2.

3.2. Interobserver Agreement

The percentage of “inconclusive” concordant cases by CT in R1 and R2 was not processed due to improper operation (division by zero), as there were no “inconclusive” cases by CT observed according to the STANDARD. There was practically no significant agreement with the STANDARD, as it occurred only in some moments of R2. Interobserver agreement was seen as weak in both R1 and R2. Despite this, it was observed that there is strong agreement between the imaging methods used in the study. There was significant agreement at a level of 5% for the criteria between the radiographic and tomographic exams of moderate-to-strong degree for almost all evaluators, regardless of the moment (R1 and R2). However, it was observed that the radiographic evaluation presented a higher number of the “inconclusive” category than the CT, regardless of the criterion (Criteria A, B, and C) and the moment (R1 and R2), although there was no statistical significance. Supplementary Materials Tables S2 and S3 provide the percentage of concordant cases in R1 and R2, respectively.

4. Discussion

Overall, it was observed that intra- and interobserver agreement was weak in both R1 and R2. Analyzing individually each of the prognostic criteria used in this study, it was observed that only the presence of a medial metaphyseal hinge greater than 2 mm (Criterion B) was identified by all evaluators both in the radiographic examination and in the CT in the two rounds of the study as statistically significant, meaning that there is no evidence supporting our hypothesis. There was practically no significant agreement with the STANDARD, occurring only in some moments of R2. Nevertheless, there was a strong agreement between the imaging methods used in the study for almost all the evaluators, regardless of the moment (R1 and R2), with the CT evaluation showing a lower number of the “inconclusive” responses than the radiography evaluation, regardless of the criterion (A, B, and C) and moment (R1 and R2).

It is known that posttraumatic osteonecrosis of the proximal humerus is the most common complication after fractures in this anatomic region, occurring in up to 16% of patients, and represents a problem for both the patient and the surgeon [6,8,20]. In this context, surgeons must be able to assess the risk factors for AVN of the humeral head to influence not only decision-making, but also to guide them regarding the prognosis in relation to this complication [13,14,21]. Hertel et al. [7] described some predictors of humeral head ischemia after fracture; however, little is discussed about the difficulty of interpreting these criteria in the presence of numerous factors that potentially make it difficult to evaluate the radiographic examination.

The guidelines recommend that radiographic screening should be the first imaging investigation in the emergency department [22]. However, a high rate of suboptimal shoulder radiographs has been identified, particularly in AP and axillary views, resulting in increased workload, increased radiation to patients, inconvenience and decreased patient satisfaction, and increased risk of incorrect or missed diagnoses [23]. The variability in interpretation and the questioned reliability of this test have led some authors to independently assess its effectiveness. Martínez-Sola et al. [24] found a low-to-moderate degree of interobserver agreement, using a single AP radiograph of the shoulder, denoting the difficulty of interpretation by orthopedic surgeons of various levels of experience of some of the currently most used classifications for fractures of the proximal humerus. Likewise, Iordens et al. [25] found weak intra- and interobserver agreement by using radiographs of the proximal end of the humerus for the Hertel criteria. Most likely because of this difficulty in radiographic evaluation, which can be understood as a confusing diagnostic and prognostic factor, some authors observed that the Hertel criteria were not sufficient to determine a greater chance of progression to osteonecrosis of the humeral head [11]. Analyzing specifically the three prognostic criteria evaluated in this study, we can see that the presence of a medial metaphyseal hinge greater than 2 mm (Criterion B) was identified in a statistically significant way by all evaluators, except the evaluator SHOULDER 4. Interestingly, from this evaluator, we observed that there was an almost symmetrical pattern of inversion of his responses between R1 and R2, thus leading us to believe that an attention bias may have occurred during the responses. Thereby, what was considered to be a bad prognostic factor in R1 was considered to be a good prognostic factor in R2, and vice versa.

It is known that there is a strong correlation between AVN of the humeral head and medial metaphyseal hinge [7,20,26]. Hertel et al. [7] found an accuracy of 0.79 for ischemia when the medial hinge was interrupted by more than 2 mm. Solberg et al. [20] noted the occurrence of osteonecrosis in all patients in whom the medial hinge was initially less than 2 mm in length. However, these authors [20] were unable to identify whether medial hinge extension or a history of dislocation was the specific cause of osteonecrosis. Humeral head AVN has been reported in up to 33% of patients, and late surgery (>48 h) appears to be an important prognostic factor for ischemia [27]. However, while a history of dislocation associated with a proximal humerus fracture could be expected to lead to osteonecrosis from capsular injury and direct vascular damage, this is not fully supported in the literature [8,26,28–31]. Interestingly and antagonistically, Neviasser et al. [29], in a series

of 34 patients treated with open reduction and internal fixation of fractures of the proximal humerus, showed that the posteromedial hinge length was not an accurate parameter to predict the risk of osteonecrosis of the humeral head, although it has been shown to be useful for surgical planning, especially as it involves an important support region (calcar) for internal fixation, greatly reducing the incidence of cut-out and/or cut-through [31].

The prognostic ability of the other two criteria evaluated in this study (medial metaphyseal extension in the cephalic component less than 8 mm—Criterion A; and fracture by partition of the head—Criterion C) is even more controversial and, probably, is mainly due to the inability to fully assess the proximal humerus morphology in a fractured segment with radiographs alone, as pointed out before. In this sense, the use of CT gains importance, as it allows the surgeon and the radiologist to perform the reconstruction of the proximal humerus in different planes, with slices of small sizes and in a three-dimensional perspective; however, in the light of current knowledge, there is no evidence of that 3D CT is superior to 2D CT [32]. Campochiaro et al. [11] suggested that all fractures involving the calcar should be studied with CT, as an accurate assessment of the fracture in three planes is necessary. In this study, only the presence of a medial metaphyseal hinge greater than 2 mm (Criterion B) was identified in a statistically significant way by all evaluators in the CT, except for the SHOULDER 4 evaluator; meanwhile, the other two criteria showed little uniformity between evaluators. There was a trend toward greater identification of Criteria A and C by TRAUMA surgeons on CT; however, this was not statistically significant.

This study has some limitations. Although not evaluated in the study, there are other factors that may represent risk factors for AVN of the humeral head—such as the surgical approach and the quality of reduction—and generate confusion in its epidemiological estimate, which is illustrated by the widely varying rates of posttraumatic osteonecrosis between studies [8]. In addition, the reported rate of humeral head AVN depends, among other risk factors already mentioned, on the duration of patient follow-up, the intensity of the resulting symptoms, and the imaging methods used to diagnose its presence [4–6,33,34]. In addition, the study design did not include the follow-up of patients, thus making it impossible to observe which patients developed AVN of the humeral head and to correlate this complication with the findings of both the evaluators and the STANDARD. This information would be important to assess the accuracy of the prognostic criteria investigated in the study and to understand the lack of interobserver agreement with the STANDARD. Moreover, the use of images in JPEG format is biased due to zoom or brightness adjustments compared to dedicated software for viewing radiographs. Furthermore, more dedicated software for the MOV file would allow all CT planes to be viewed simultaneously, and this would theoretically increase accuracy and orientation when viewing a CT scan. The adoption of these measures could theoretically increase intraobserver agreement, which can be tested in future studies. Finally, the objective was to evaluate the role of both two- and three-dimensional CT in the interpretation of the three criteria described by Hertel, using the radiographic evaluation of the same cases as a standard. Unlike other studies [25,35–38], our study did not seek to assess whether there was a difference between 2D and 3D tomography. In the current study, we showed some usefulness for both the 2D and 3D tomography to help surgeons in the primary interpretation of Hertel prognostic criteria, with moderate intraobserver agreement, without, however, statistical difference when compared to the radiographic examination. Recent studies have shown that the use of other imaging modalities, such as segmented 3D CT [39] and 3D printed models [40,41], adds value to the understanding of the morphology of the proximal humerus fracture, which was not evaluated in the current study.

Some strengths of the study deserve to be highlighted. The main one was to show that none of Hertel prognostic criteria evaluated in this study present interobserver agreement, regardless of whether the assessment is performed by using radiographs, CT, or both. Other authors have already observed the same, including employing more modern imaging methods than those used in this study [11,25,39,41]. The main limitation of the study by Hertel et al. [7] was the assessment of the fracture pattern by a single observer, which

obviously lacks objectivity and reproducibility. The prognostic value of the Hertel criteria for decision-making has been questioned in particular because it has been reported that the humeral head can survive even in the initial absence of proximal flow from the anterior and posterior branches of the humeral circumflex artery [42]. Thus, our findings (and those of others) indicate that Hertel prognostic predictors should be used cautiously and sparingly when defining the eventual risk of AVN of the humeral head due to their low interobserver agreement.

5. Conclusions

The prognostic criteria for humeral head ischemia evaluated in this study showed weak intra- and interobserver agreement in both radiographic and tomographic evaluation. CT did not help surgeons in the primary interpretation of Hertel prognostic criteria used in this study when compared to radiographic examination.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/medicina58101489/s1>. Table S1. Distribution of criterion A, B, and C by radiographic and CT images between rounds R1 and R2. Table S2. Interobserver analysis in relation to STANDARD: percentage of cases in agreement with prognostic criteria A, B, and C by radiography and CT of the evaluators in round R1. Table S3. Interobserver analysis in relation to STANDARD: percentage of cases in agreement with prognostic criteria A, B, and C by radiography and CT of the evaluators in round R2.

Author Contributions: Conceptualization, V.G. and P.O.d.T.; methodology, V.G. and P.O.d.T.; software, P.O.d.T.; validation, V.G., P.O.d.T., H.A.K. and W.D.B.; formal analysis, V.G., P.O.d.T. and H.A.K.; investigation, P.O.d.T. and V.G.; resources, P.O.d.T.; data curation, V.G., P.O.d.T. and H.A.K.; writing—original draft preparation, V.G.; writing—review and editing, P.O.d.T., V.G. and H.A.K.; visualization, P.O.d.T., V.G., H.A.K., W.D.B., R.E.P., F.S.d.S., P.J.L., C.Z., F.M., P.S.B., R.Y.I. and S.R.; supervision, V.G.; project administration, P.O.d.T.; funding acquisition, P.O.d.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Hospital Municipal Miguel Couto (protocol code 002/2021, 1 February 2021).

Informed Consent Statement: Patient consent was waived due to the lockdown imposed by the COVID-19 pandemic, which prevented patients from going to the hospital to sign the document.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to acknowledge Edson Marchiori, José Sérgio Franco, Antonio Carlos Pires Carvalho, and César Fontenelle from the Universidade Federal do Rio de Janeiro (UFRJ) for their invaluable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Individual Folders with Instructions on How to Respond the Survey

- This study seeks to evaluate Hertel's radiographic criteria using computed tomography (CT);
- We separated 20 cases of fractures of the proximal humerus, with X-rays and CT images, individualized in a separate folder;
- Randomization was performed, and the cases were numbered from 1 to 20;
- We ask that each examiner define which Hertel criteria are present by evaluating the radiographs and then their respective CT scans;
- At the end of the evaluation, each examiner will have made 40 evaluations. 2 for each case;
- In the folder with the files there is a photo explaining each of the 3 Hertel criteria, (reproduced from these author original article);

- We kindly ask you to return these results for each patient. Example:
Case 1:
X-ray 1
 - criterion A present/absent/cannot evaluate;
 - criterion B present/absent/cannot evaluate;
 - criterion C present/absent/cannot evaluate;
 CT 1
 - criterion A present/absent/cannot evaluate;
 - criterion B present/absent/cannot evaluate;
 - criterion C present/absent/cannot evaluate;
 - The annotation of these results can be done in the way that suits the examiner. Annotated on a sheet of paper, or in an Excell table, or in a .doc text;
 - After ending with the document, please send it to the emails: p_tullio@hotmail.com and v_giordano@me.com;
 - Very soon we will send the same cases again but with a new randomization for a second evaluation, allowing us to make an intra and interobserver agreement comparison.
 Thanks a lot in advance.
Best regards.
Paulo and Vincenzo

References

1. Jabran, A.; Peach, C.; Ren, L. Biomechanical analysis of plate systems for proximal humerus fractures: A systematic literature review. *Biomed. Eng. Online* **2018**, *17*, 47. [[CrossRef](#)] [[PubMed](#)]
2. Hageman, M.G.; Jayakumar, P.; King, J.D.; Guitton, T.G.; Doornberg, J.N.; Ring, D.; Science of Variation Group. The factors influencing the decision making of operative treatment for proximal humeral fractures. *J. Shoulder Elb. Surg.* **2015**, *24*, e21–e26. [[CrossRef](#)] [[PubMed](#)]
3. Keding, A.; Handoll, H.; Brealey, S.; Jefferson, L.; Hewitt, C.; Corbacho, B.; Torgerson, D.; Rangan, A. The impact of surgeon and patient treatment preferences in an orthopaedic trauma surgery trial. *Trials* **2019**, *20*, 570. [[CrossRef](#)] [[PubMed](#)]
4. Maccagnano, G.; Solarino, G.; Pesce, V.; Vicenti, G.; Coviello, M.; Nappi, V.S.; Giannico, O.V.; Notarnicola, A.; Moretti, B. Plate vs reverse shoulder arthroplasty for proximal humeral fractures: The psychological health influence the choice of device? *World J. Orthop* **2022**, *13*, 297–306. [[CrossRef](#)] [[PubMed](#)]
5. Belayneh, R.; Lott, A.; Haglin, J.; Konda, S.; Zuckerman, J.D.; Egol, K.A. Osteonecrosis after surgically repaired proximal humerus fractures is a predictor of poor outcomes. *J. Orthop. Trauma* **2018**, *32*, e387–e393. [[CrossRef](#)] [[PubMed](#)]
6. Rutherford, C.S.; Cofield, R.H. Osteonecrosis of the shoulder. *Orthop Trans.* **1987**, *11*, 239.
7. Hertel, R.; Hempfing, A.; Stiehler, M.; Leunig, M. Predictors of humeral head ischemia after intracapsular fracture of the proximal humerus. *J. Shoulder Elb. Surg.* **2004**, *13*, 427–433. [[CrossRef](#)]
8. Patel, S.; Colaco, H.B.; Elvey, M.E.; Lee, M.H. Post-traumatic osteonecrosis of the proximal humerus. *Injury* **2015**, *46*, 1878–1884. [[CrossRef](#)]
9. Majed, A.; Macleod, I.; Bull, A.M.; Zyto, K.; Resch, H.; Hertel, R.; Reilly, P.; Emery, R.J. Proximal humeral fracture classification systems revisited. *J. Shoulder Elb. Surg.* **2011**, *20*, 1125–1132. [[CrossRef](#)]
10. Sukthankar, A.V.; Leonello, D.T.; Hertel, R.W.; Ding, G.S.; Sandow, M.J. A comprehensive classification of proximal humeral fractures: HGLS system. *J. Shoulder Elb. Surg.* **2013**, *22*, e1–e6. [[CrossRef](#)]
11. Campochiaro, G.; Rebuzzi, M.; Baudi, P.; Catani, F. Complex proximal humerus fractures: Hertel's criteria reliability to predict head necrosis. *Musculoskelet. Surg.* **2015**, *99*, S9–S15. [[CrossRef](#)]
12. Neer, C.S., 2nd. Displaced proximal humeral fractures. I. Classification and evaluation. *J. Bone Joint Surg. Am.* **1970**, *52*, 1077–1089. [[CrossRef](#)]
13. Senna, L.F.; Pires, E.; Albuquerque, R. Modified axillary radiograph of the shoulder: A new position. *Rev. Bras. Ortop.* **2016**, *52*, 115–118. [[CrossRef](#)]
14. Neer, C.S., 2nd. Displaced proximal humeral fractures. II. Treatment of three-part and four-part displacement. *J. Bone Joint Surg. Am.* **1970**, *52*, 1090–1103. [[CrossRef](#)]
15. Bossuyt, P.M.; Reitsma, J.B.; Bruns, D.E.; Gatsonis, C.A.; Glasziou, P.P.; Irwig, L.M.; Lijmer, J.G.; Moher, D.; Rennie, D.; de Vet, H.C.; et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Clin. Radiol.* **2003**, *58*, 575–580. [[CrossRef](#)]

16. Giordano, V.; Gomes, A.F.; Amaral, N.P.; Albuquerque, R.P.; Pires, R.E. Preventing surgical complications: A survey on surgeons' perception of intra-articular malleolar screw misplacement in a cadaveric study. *Patient Saf. Surg.* **2011**, *5*, 24. [[CrossRef](#)]
17. Oladiran, O.; Gichoya, J.; Purkayastha, S. Conversion of JPG image into DICOM image format with one click tagging. In *Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management: Health and Safety*; Duffy, V., Ed.; Springer: Cham, Switzerland, 2017; pp. 61–70. [[CrossRef](#)]
18. El-Boghdadly, K.; Onwochei, D.N.; Millhoff, B.; Ahmad, I. The effect of virtual endoscopy on diagnostic accuracy and airway management strategies in patients with head and neck pathology: A prospective cohort study. *Can. J. Anaesth.* **2017**, *64*, 1101–1110. [[CrossRef](#)]
19. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [[CrossRef](#)]
20. Solberg, B.D.; Moon, C.N.; Franco, D.P.; Paiement, G.D. Surgical treatment of three and four-part proximal humeral fractures. *J. Bone Joint Surg. Am.* **2009**, *91*, 1689–1697. [[CrossRef](#)]
21. Leyshon, R.L. Closed treatment of fractures of the proximal humerus. *Acta Orthop. Scand.* **1984**, *55*, 48–51. [[CrossRef](#)]
22. Sandstrom, C.K.; Kennedy, S.A.; Gross, J.A. Acute shoulder trauma: What the surgeon wants to know. *Radiographics* **2015**, *35*, 475–492. [[CrossRef](#)]
23. Richards, B.; Riley, J.; Saithna, A. Improving the diagnostic quality and adequacy of shoulder radiographs in a District General Hospital. *BMJ Qual. Improv. Rep.* **2016**, *5*, u209855.w3501. [[CrossRef](#)]
24. Martínez-Sola, R.; León-Muñoz, V.J.; Najem-Rizk, A.N.; Soler-Vasco, B.; Arrieta-Martínez, C.J.; López-Sorroche, E.; Cárdenas-Grande, E.; Salmerón-Vélez, G.; Ruiz-Molina, J.Á.; Martínez-Martínez, F.; et al. 'Absolute' inter-observer classifications agreement for proximal humeral fractures with a single shoulder anteroposterior X-ray. *J. Orthop. Surg.* **2021**, *29*, 23094990211010520. [[CrossRef](#)] [[PubMed](#)]
25. Iordens, G.I.; Mahabier, K.C.; Buisman, F.E.; Schep, N.W.; Muradin, G.S.; Beenen, L.F.; Patka, P.; van Lieshout, E.M.; van Hartog, D. The reliability and reproducibility of the Hertel classification for comminuted proximal humeral fractures compared with the Neer classification. *J. Orthop. Sci.* **2016**, *21*, 596–602. [[CrossRef](#)]
26. Robinson, C.M.; Khan, L.A.; Akhtar, M.A. Treatment of anterior fracture-dislocations of the proximal humerus by open reduction and internal fixation. *J. Bone Joint Surg. Br.* **2006**, *88*, 502–508. [[CrossRef](#)]
27. Schnetzke, M.; Bockmeyer, J.; Loew, M.; Studier-Fischer, S.; Grütznert, P.A.; Guehring, T. Rate of avascular necrosis after fracture dislocations of the proximal humerus: Timing of surgery. *Obere Extremit.* **2018**, *13*, 273–278. [[CrossRef](#)]
28. Lee, C.K.; Hansen, H.R. Post-traumatic avascular necrosis of the humeral head in displaced proximal humeral fractures. *J. Trauma* **1981**, *21*, 788–791. [[CrossRef](#)]
29. Neviasser, A.S.; Hettrich, C.M.; Dines, J.S.; Lorich, D.G. Rate of avascular necrosis following proximal humerus fractures treated with a lateral locking plate and endosteal implant. *Arch. Orthop. Trauma Surg.* **2011**, *131*, 1617–1622. [[CrossRef](#)]
30. Trupka, A.; Wiedemann, E.; Ruchholtz, S.; Brunner, U.; Habermeyer, P.; Schweiberer, L. Dislozierte Mehrfragmentfrakturen des Humeruskopfes. Bedeutet die Luxation des Kopffragments eine Prognoseverschlechterung? *Unfallchirurg* **1997**, *100*, 105–110. [[CrossRef](#)]
31. Thompson, J.H.; Attum, B.; Rodriguez-Buitrago, A.; Yusi, K.; Cerejijo, C.; Obremskey, W.T. Open reduction and internal fixation with a locking plate via deltopectoral approach for the treatment of three and four-part and proximal humeral fractures. *JBJS Essent. Surg. Tech.* **2018**, *8*, e26. [[CrossRef](#)]
32. Berkes, M.B.; Dines, J.S.; Little, M.T.; Garner, M.R.; Shifflett, G.D.; Lazaro, L.E.; Wellman, D.S.; Dines, D.M.; Lorich, D.G. The impact of three-dimensional CT imaging on intraobserver and interobserver reliability of proximal humeral fracture classifications and treatment recommendations. *J. Bone Joint Surg. Am.* **2014**, *96*, 1281–1286. [[CrossRef](#)]
33. Greiner, S.; Kääh, M.J.; Haas, N.P.; Bail, H.J. Humeral head necrosis rate at mid-term follow-up after open reduction and angular stable plate fixation for proximal humeral fractures. *Injury* **2009**, *40*, 186–191. [[CrossRef](#)] [[PubMed](#)]
34. Sakai, T.; Sugano, N.; Nishii, T.; Hananouchi, T.; Yoshikawa, H. Extent of osteonecrosis on MRI predicts humeral head collapse. *Clin. Orthop. Relat. Res.* **2008**, *466*, 1074–1080. [[CrossRef](#)] [[PubMed](#)]
35. Bahrs, C.; Rolauffs, B.; Südkamp, N.P.; Schmal, H.; Eingartner, C.; Dietz, K.; Pereira, P.L.; Weise, K.; Lingenfelder, E.; Helwig, P. Indications for computed tomography (CT)-diagnostics in proximal humeral fractures: A comparative study of plain radiography and computed tomography. *BMC Musculoskelet. Disord.* **2009**, *10*, 33. [[CrossRef](#)]
36. Chelli, M.; Gasbarro, G.; Lavoué, V.; Gauci, M.O.; Raynier, J.L.; Trojani, C.; Boileau, P. The reliability of the Neer classification for proximal humerus fractures: A survey of orthopedic shoulder surgeons. *JSES Int.* **2022**, *6*, 331–337. [[CrossRef](#)]
37. Jia, X.; Chen, Y.; Qiang, M.; Zhang, K.; Li, H.; Jiang, Y.; Zhang, Y. Compared to X-ray, three-dimensional computed tomography measurement is a reproducible radiographic method for normal proximal humerus. *J. Orthop. Surg. Res.* **2016**, *11*, 82. [[CrossRef](#)]
38. Stirna, G.A.; Secundino, A.R.; Gonzalez, G.F.G.; Sola, W.C.; de Souza, G.A.L.; Dau, L. Inter/intra-observer evaluation between radiographs and tomographies for proximal humerus fracture. *Acta Ortop. Bras.* **2020**, *28*, 36–39. [[CrossRef](#)]
39. Dauwe, J.; Mys, K.; Putzeys, G.; Schader, J.F.; Richards, R.G.; Gueorguiev, B.; Varga, P.; Nijs, S. Advanced CT visualization improves the accuracy of orthopaedic trauma surgeons and residents in classifying proximal humeral fractures: A feasibility study. *Eur. J. Trauma Emerg. Surg.* **2020**, 1–7. [[CrossRef](#)]
40. Cocco, L.F.; Aihara, A.Y.; Lopes, F.P.P.L.; Werner, H.; Franciozi, C.E.; Dos Reis, F.B.; Luzo, M.V.M. Three-dimensional printing models increase inter-rater agreement for classification and treatment of proximal humerus fractures. *Patient Saf. Surg.* **2022**, *16*, 5. [[CrossRef](#)]

41. Puglisi, G.; Montemagno, M.; Denaro, R.; Condorelli, G.; Caruso, V.F.; Vescio, A.; Testa, G.; Pavone, V. 3D-printed models versus CT scan and X-rays imaging in the diagnostic evaluation of proximal humerus fractures: A triple-blind interobserver reliability comparison study. *Adv. Orthop.* **2022**, *2022*, 5863813. [[CrossRef](#)]
42. Lambert, S.M. Ischaemia, healing and outcomes in proximal humeral fractures. *EFORT Open Rev.* **2018**, *3*, 304–315. [[CrossRef](#)] [[PubMed](#)]