Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Computed tomography reduces treatment uncertainty without improving classification reliability in distal radius fractures

Mehmet Akan

akanmehmet9@gmail.com

Şişli Etfal Eğitim ve Araştırma Hastanesi

Güngör Alibakan

Şişli Etfal Eğitim ve Araştırma Hastanesi

Mehmet Akif Köse

Şişli Etfal Eğitim ve Araştırma Hastanesi

Işın Kızılkan

Şişli Etfal Eğitim ve Araştırma Hastanesi

Mehmet Alp Yüksel

Şişli Etfal Eğitim ve Araştırma Hastanesi

Osman Tuğrul Eren

Şişli Etfal Eğitim ve Araştırma Hastanesi

Research Article

Keywords: Distal radius fracture, Computed tomography, Fracture classification, Observer variation, Clinical decision-making

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11400094/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

To determine whether adding computed tomography (CT) to plain radiographs reduces uncertainty in surgical decision-making for distal radius fractures and improves the reliability of four classification systems.

Methods

Fifty acute displaced distal radius fractures were assessed by four orthopaedic surgeons, first with radiographs alone and, after a four-week washout, with radiographs and CT. At each session, observers classified fractures using the AO/OTA, Fernandez, Frykman and Universal systems and recorded a treatment decision (operative, conservative or undecided); for undecided cases, they selected the factor most likely to resolve their uncertainty. Agreement was assessed with kappa statistics, and changes between sessions with clustered McNemar tests.

Results

Undecided responses decreased from 29.0% with radiographs alone to 3.5% with CT, and operative decisions increased from 55.5% to 73.0% (both $p < 0.001$). Interobserver agreement on treatment remained fair (Fleiss' κ, 0.303 vs. 0.294). Interobserver reliability of classification was slight to moderate on radiographs (κ, 0.184–0.461) and did not improve with CT (κ, 0.169–0.386; differences, − 0.075 to + 0.021), although CT changed the assigned category in 23.0–51.0% of assessments and reclassified one third of apparently extra-articular fractures as articular. Radiological parameters were the leading cause of indecision with radiographs alone (65.5%) but were not cited for any of the seven assessments that remained undecided after CT.

Conclusion

CT did not make fracture classification more reproducible or bring surgeons to the same treatment choice, but it substantially reduced individual decisional uncertainty. Once imaging uncertainty was resolved, residual uncertainty related to patient factors rather than radiological findings.

Introduction

Distal radius fractures (DRFs) are among the most common fractures in orthopaedic practice. In a nationwide register study of 23,394 adult fractures, surgical treatment rates ranged from 18% for extra-articular to 48% for complete articular fractures, illustrating the link between fracture pattern and treatment [1].

Classification systems aim to describe fracture morphology, guide treatment, and standardise communication and research. The most widely used systems for DRFs are the AO/OTA classification, which distinguishes extra-articular, partial articular and complete articular types with further groups; the mechanism-based Fernandez classification; the Frykman classification, which also accounts for distal radioulnar joint involvement; and the Universal (Cooney) classification, based on displacement and articular involvement [2, 3]. On plain radiographs, their interobserver reliability is generally fair to moderate and declines as classification complexity increases [4–7].

CT is frequently obtained for displaced DRFs because radiographs may miss intra-articular involvement [8, 9]. However, adding CT has not consistently improved the interobserver reliability of classification [10–12], suggesting that limited reproducibility lies in the classification systems rather than in the imaging modality. Most of these studies measured agreement on fracture categories. Cone-beam CT has been shown to change the suggested treatment [13], but whether CT helps the surgeon commit to a treatment decision, which is the purpose for which it is usually requested, has received little attention.

This study aimed to evaluate the effect of CT on uncertainty in surgical decision-making for DRFs, to compare the inter- and intraobserver reliability of four classification systems with radiographs alone and with CT, and to identify the factors underlying uncertainty that persists after CT. We hypothesised that CT would reduce decisional uncertainty without significantly improving classification reliability, and that uncertainty persisting after CT would relate to patient-specific rather than imaging factors.

Methods

Study design

This single-centre study used retrospectively collected imaging data, with observer assessments performed according to a predefined protocol. The study was approved by the institutional clinical research ethics committee and conducted in accordance with the Declaration of Helsinki. Informed consent was waived because only anonymised images were analysed retrospectively.

Patients

The hospital archive was retrospectively screened for eligible patients. Patients were eligible if they were aged 18 years or older and had an acute displaced DRF with anteroposterior and lateral radiographs and CT obtained at initial presentation. Displacement was defined as intra-articular step-off > 2 mm, dorsal angulation > 5° or radial shortening > 3 mm [14]. Pathological and open fractures, previous ipsilateral DRF or wrist surgery, and concomitant carpal fractures were excluded. Eligibility was determined by an orthopaedic surgeon not involved in the observer group, and 50 patients meeting these criteria were included.

Imaging

Standard anteroposterior and lateral wrist radiographs and CT examinations with 1-mm slice thickness and axial, coronal and sagittal multiplanar reconstructions were retrieved from the picture archiving and communication system and anonymised before evaluation.

Observers and assessment procedure

Four orthopaedic and traumatology specialists served as independent observers. Before the study, all received a standardised briefing document with written definitions and representative images for each classification system. Observers were blinded to demographic and clinical information so that classification and treatment decisions reflected only the imaging available at each session, independently of patient factors.

Two sessions were separated by a four-week washout period to minimise recall bias. At the first session (T0), all 50 cases were presented in the same order with radiographs only. At the second session (T4), radiographs and CT were presented in a separate computer-generated random order for each observer to minimise order effects. At both sessions, each observer independently recorded the AO/OTA classification at group level (A2–C3, from which type level was derived), the Fernandez (types 1–5), Frykman (types 1–8) and Universal (types 1–4) classifications, and a treatment decision: operative, conservative or undecided. For undecided cases, observers selected the factor most likely to resolve their uncertainty: (A) patient age, general condition or comorbidities; (B) radiological parameters; or (C) functional demands and expectations. These categories reflect the principal determinants of treatment choice in DRFs [15, 16]. Results are reported for observer-case assessments (four observers × 50 cases = 200 per session) and, where stated, at case level.

Outcomes

The study had two predefined outcome domains. Decision-related outcomes were the distribution of treatment decisions, the proportion of undecided responses, interobserver agreement on treatment, and the factors cited for indecision. Classification-related outcomes were interobserver reliability at T0 and T4, intraobserver agreement between sessions, and changes in the assigned category after CT.

Statistical analysis

Because each case was rated by four observers, observer-case assessments are clustered within cases. Pooled comparisons of paired binary outcomes between T0 and T4 (undecided, operative and conservative responses) therefore used the clustered McNemar test with cases as clusters [17]; for individual observers, the exact McNemar test was used. Factors cited for indecision were reported descriptively because of the small number of undecided assessments at T4.

Interobserver agreement on treatment and on each classification system was assessed with Fleiss' kappa with asymptotic 95% confidence intervals (CIs), and intraobserver agreement between T0 and T4 with Cohen's kappa, with 95% CIs calculated from the asymptotic standard error. Unweighted kappa was used throughout because the categories of several systems, such as the mechanism-based Fernandez classification, are not ordered. Because both sessions involved the same cases and observers, the T0 and T4 kappa values are not independent; they were therefore compared descriptively, on the basis of the size of the difference and the overlap of their CIs, rather than with tests that assume independent samples. Kappa values were interpreted according to Landis and Koch (≤ 0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, > 0.80 almost perfect), and complete agreement was defined as identical ratings by all four observers.

Changes in the assigned category between T0 and T4 were reported as proportions of observer-case assessments. For systems with ordered categories (AO/OTA, Universal and Frykman), the balance between upgrades and downgrades was compared with the clustered McNemar test; direction was not assessed for the mechanism-based Fernandez classification. The sample of 50 cases and four observers is comparable to previous reliability studies of similar design [6, 10]. All analyses were performed with SPSS version 27.0 (IBM Corp., Armonk, NY, USA), and statistical significance was set at $p < 0.05$.

Results

All classification and decision fields were complete for all 400 observer-case assessments.

Surgical decision-making

The proportion of undecided responses decreased from 29.0% at T0 to 3.5% at T4 ($p < 0.001$), with a significant reduction in three of the four observers; the reduction for Observer 2 was not significant ($p = 0.063$) (Table 1). No decided assessment became undecided after CT. Of the 51 undecided assessments resolved by CT, 30 became operative and 21 conservative, and 21 of the 142 assessments decided at T0 (14.8%) changed between operative and conservative. Overall, operative decisions increased from 55.5% to 73.0% ($p < 0.001$) and conservative decisions from 15.5% to 23.5% ($p = 0.003$). Complete agreement among the four observers on treatment increased from 34.0% to 52.0% of cases, whereas chance-corrected interobserver agreement remained fair ($\kappa = 0.303$ vs. 0.294).

Table 1
Treatment decisions with radiographs alone (T0) and radiographs plus computed tomography (T4)

	T0	T4	χ^2	p
Treatment decision, all observer-case assessments (n = 200), n (%)				
Operative[a]	111 (55.5)	146 (73.0)	20.08	< 0.001
Conservative[a]	31 (15.5)	47 (23.5)	8.53	0.003
Undecided[a]	58 (29.0)	7 (3.5)	26.81	< 0.001
Undecided responses by observer (n = 50 each), n				
Observer 1[b]	14	0		< 0.001
Observer 2[b]	10	5		0.063
Observer 3[b]	12	2		0.002
Observer 4[b]	22	0		< 0.001
Factor most likely to resolve uncertainty, n (% of undecided)				
Radiological parameters	38 (65.5)	0 (0)		
Age, general condition or comorbidities	17 (29.3)	5 (71.4)		
Functional demands and expectations	3 (5.2)	2 (28.6)		
Interobserver agreement on treatment				
Fleiss' κ (95% CI)[c]	0.303 (0.219–0.388)	0.294 (0.193–0.394)		
Complete agreement among four observers, % of cases	34.0	52.0		

[a]Clustered McNemar test with cases as clusters.

[b]Exact McNemar test. No assessment changed from decided at T0 to undecided at T4.

[c]Asymptotic 95% confidence interval.

CI, confidence interval; T0, radiograph-only assessment; T4, radiograph-plus-computed-tomography assessment

Among the 58 undecided assessments at T0, radiological parameters were cited as the factor most likely to resolve uncertainty in 65.5%, patient age, general condition or comorbidities in 29.3%, and

functional demands in 5.2%. Of the seven assessments still undecided at T4, five (71.4%) cited age, general condition or comorbidities and two (28.6%) functional demands; none cited radiological parameters.

Classification reliability

Interobserver reliability with radiographs alone ranged from slight for the AO/OTA group level to moderate for the Universal classification (Fig. 1). Adding CT did not materially change kappa values: differences between sessions ranged from − 0.075 for the Universal to + 0.021 for the Frykman classification, and the CIs overlapped for all systems. Complete agreement was highest for the Universal classification (48.0% at T0 and 46.0% at T4) and lowest for the AO/OTA group level (8.0% and 10.0%) and Frykman classification (6.0% and 12.0%).

Intraobserver agreement between T0 and T4 ranged from slight to substantial (κ, 0.128–0.725) and varied between observers (Table 2). Agreement was moderate to substantial for Observers 2 and 4, fair to moderate for Observer 3 and slight to moderate for Observer 1.

Table 2
Intraobserver agreement between radiograph-only (T0) and radiograph-plus-computed-tomography (T4) assessments, by classification system and observer

Classification	Observer 1	Observer 2	Observer 3	Observer 4
AO/OTA type	0.407 (0.20–0.61)	0.725 (0.54–0.91)	0.548 (0.34–0.76)	0.635 (0.45–0.82)
AO/OTA group	0.298 (0.15–0.45)	0.560 (0.39–0.73)	0.275 (0.12–0.43)	0.406 (0.25–0.56)
Universal	0.246 (0.05–0.44)	0.590 (0.38–0.80)	0.388 (0.19–0.58)	0.661 (0.48–0.84)
Fernandez	0.172 (0.02–0.32)	0.504 (0.35–0.66)	0.302 (0.11–0.49)	0.467 (0.29–0.65)
Frykman	0.128 (0.02–0.24)	0.575 (0.42–0.73)	0.319 (0.16–0.48)	0.521 (0.36–0.68)

Values are unweighted Cohen's κ with 95% confidence intervals (κ ± 1.96 × asymptotic standard error). Interpretation according to Landis and Koch: ≤0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, > 0.80 almost perfect. All values were statistically significant ($p < 0.05$).

AO/OTA, AO Foundation/Orthopaedic Trauma Association

Changes in fracture classification

CT changed the assigned category in 23.0% (AO/OTA type) to 51.0% (Frykman) of assessments (Table 3). Upgrades significantly outnumbered downgrades for the AO/OTA type and group and the Frykman classification, but not for the Universal classification ($p = 0.081$). At case level, at least one observer upgraded the AO/OTA type in 22 of 50 cases, whereas no case was upgraded by all four

observers. Of the 71 assessments classified as extra-articular (AO/OTA type A) on radiographs, 23 (32.4%) were reclassified as articular after CT. With the Universal classification, 26 of 77 (33.8%) extra-articular assessments were reclassified as articular and 11 of 16 (68.8%) nondisplaced articular assessments as displaced.

Table 3
Change in fracture classification after the addition of computed tomography, by observer-case assessment (n = 200)

Classification	Upgrade (T4 > T0), n (%)	Downgrade (T4 < T0), n (%)	Any change, n (%)	χ^{2a}	p
AO/OTA type (A–C)	35 (17.5)	11 (5.5)	46 (23.0)	7.38	0.007
AO/OTA group (A2–C3)	74 (37.0)	26 (13.0)	100 (50.0)	12.52	< 0.001
Universal (1–4)	37 (18.5)	21 (10.5)	58 (29.0)	3.05	0.081
Fernandez (1–5)[b]	–	–	91 (45.5)	–	–
Frykman (1–8)	75 (37.5)	27 (13.5)	102 (51.0)	14.40	< 0.001

[a]Clustered McNemar test comparing upgrades with downgrades, with cases as clusters (50 cases, four observers each).

[b]Direction not assessed because Fernandez types are mechanism-based and unordered.

AO/OTA, AO Foundation/Orthopaedic Trauma Association; T0, radiograph-only assessment; T4, radiograph-plus-computed-tomography assessment

Discussion

The main finding of this study is that adding CT to radiographs reduced the proportion of DRFs for which surgeons could not reach a treatment decision from 29.0% to 3.5%, yet did not bring surgeons closer to the same decision. CT also did not improve the interobserver reliability of any of the four classification systems, although it changed the assigned category in up to half of assessments. When uncertainty persisted after CT, observers attributed it to patient-related rather than radiological factors.

Classification alone is an insufficient guide to treatment in DRFs [15]. The dissociation between individual certainty and interobserver agreement observed here has parallels in previous work. When surgeons were given numerical radiographic measurements instead of radiographs, agreement on treatment decreased, leading the authors to conclude that surgeons disagree about the indication for surgery rather than about the interpretation of the image [18]. Computer-generated measurements improved agreement only among senior surgeons [19], and post-reduction radiographs did not improve agreement, treatment choice varying with surgeon-related factors [20]. Eriksson et al. reported that cone-beam CT changed the suggested treatment in 38% of observations, in both directions [13]. In the present

study, 14.8% of decided assessments changed after CT, and the main effect was the resolution of indecision; because CT was added to rather than substituted for radiographs and observers could record indecision, its effect on decisional certainty could be quantified directly. Raw agreement on treatment increased after CT because decisions converged towards operative treatment, but chance-corrected agreement did not change. The significant increase in operative decisions suggests that CT, by revealing articular involvement, shifts treatment towards surgery; whether this improves outcomes warrants prospective evaluation.

The factors cited for indecision also shifted after CT. With radiographs alone, uncertainty was mainly radiological; after CT, the few remaining undecided assessments cited age, general condition or functional demands. Observers were deliberately blinded to patient characteristics so that decisions would reflect imaging alone. Once imaging uncertainty had been resolved, the remaining uncertainty therefore concerned patient factors, indicating that these, rather than further imaging, determine treatment in residual uncertain cases. This is consistent with the literature: there is no agreed definition of instability in DRFs [21], radiographic alignment predicts patient-reported outcome only modestly and mainly in younger patients [22], and age and bone quality predict secondary displacement [16]. These findings support decision-making that integrates fracture morphology with age, bone quality and functional demand [15].

No DRF classification system has achieved more than moderate interobserver agreement on plain radiographs, and the present kappa values are in keeping with previous studies [6, 7, 10]. Agreement at the AO/OTA type level was more than twice that at group level, a pattern also reported for AO/OTA subgroups [4, 5]; simpler levels are more reproducible but carry less morphological detail for surgical planning. Higher agreement has been reported with two observers and a joint calibration session [23]; the use of four observers and a written briefing alone may partly explain the lower values here. Consistent with previous reports, CT did not improve reliability [10, 11], and agreement remains fair even when CT is the primary modality [12]. Eriksson et al., in contrast, found that agreement on the AO classification improved from fair to moderate with cone-beam CT, particularly for dorsally comminuted fractures [13]; the higher resolution of cone-beam CT and its assessment independently of radiographs may explain this difference. Overall, the limitation appears to lie in overlapping category definitions rather than in image quality [3].

Radiographs underestimate articular involvement: sigmoid notch involvement is frequently missed, and CT reveals malalignment in fractures that appear adequately reduced on radiographs [8, 9]. Accordingly, CT changed the assigned category in a quarter to half of assessments, with upgrades predominating, as previously reported [10, 13]. One third of fractures that appeared extra-articular on radiographs were reclassified as articular after CT, and two thirds of apparently nondisplaced articular fractures as displaced, suggesting that CT is informative both when radiographs suggest a simple pattern and when they underestimate articular displacement. Intraobserver agreement between sessions reflected both observer consistency and the effect of CT; the observer with the lowest agreement also reclassified most often after CT, indicating that the contribution of CT varies between observers. Observer-dependent

variation has also been reported for radiographic measurements and can lead to different treatment decisions [24, 25].

Clinically, these findings suggest that CT is most useful when the surgeon remains undecided on radiographs, particularly when radiographs suggest an extra-articular or nondisplaced articular fracture. CT should not, however, be expected to standardise treatment between surgeons; this would require agreed indications that incorporate patient factors.

This study has limitations. First, observers were blinded to clinical information by design, so decisions represent imaging-based assessments rather than complete clinical decisions. Second, the observers worked at the same institution and had comparable experience, which may limit generalisability. Third, the intraobserver analysis compared different imaging conditions and therefore does not represent test-retest reliability. Finally, there was no reference standard such as intraoperative findings, and the small number of undecided assessments at T4 precluded formal comparison of determining factors.

Conclusion

Adding CT to plain radiographs does not improve the reproducibility of DRF classification or interobserver agreement on treatment, but substantially reduces individual uncertainty in surgical decision-making. Uncertainty persisting after CT relates to patient-specific rather than radiological factors. CT is therefore better regarded as a tool for clarifying the individual surgeon's decision than for achieving consensus between surgeons.

Declarations

Author Contribution

M.A., M.A.Y. and O.T.E. designed the study and wrote the main manuscript text. M.A., G.A., I.K. and M.A.K. performed the observer assessments. M.A.K. and I.K. contributed to data collection and interpretation. M.A. and M.A.Y. performed the statistical analysis. G.A. and O.T.E. critically revised the manuscript. All authors reviewed and approved the final manuscript.

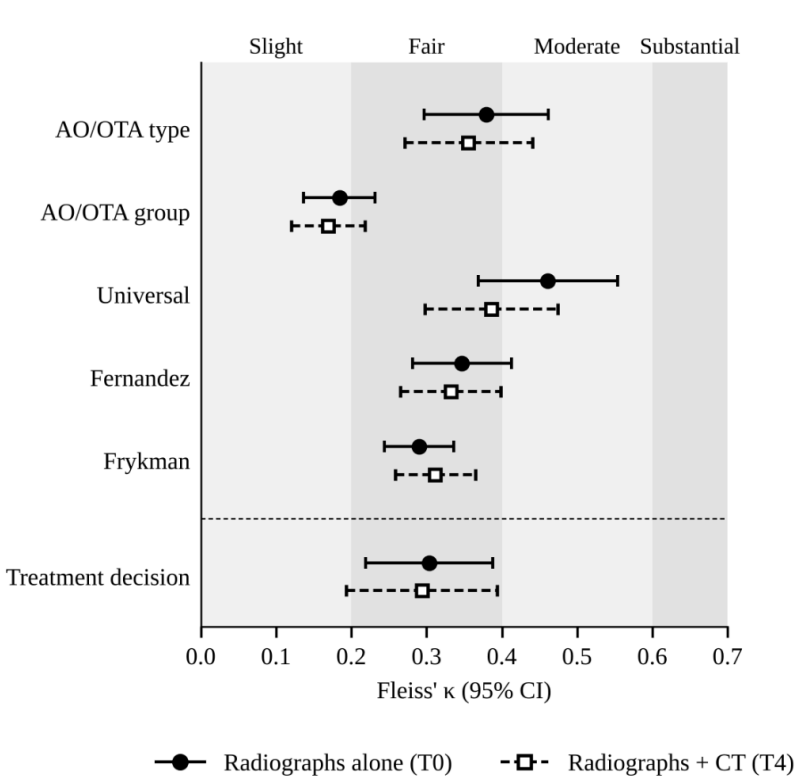
References

1. Rundgren J, Bojan A, Mellstrand Navarro C, Enocson A (2020) Epidemiology, classification, treatment and mortality of distal radius fractures in adults: an observational study of 23,394 fractures from the national Swedish fracture register. BMC Musculoskelet Disord 21:88. https://doi.org/10.1186/s12891-020-3097-8
2. Fernandez DL (2001) Distal radius fracture: the rationale of a classification. Chir Main 20:411–425. https://doi.org/10.1016/S1297-3203(01)00067-1

3. Shehovych A, Salar O, Meyer CER, Ford DJ (2016) Adult distal radius fractures classification systems: essential clinical knowledge or abstract memory testing? Ann R Coll Surg Engl 98:525–531. https://doi.org/10.1308/rcsann.2016.0237
4. Yinjie Y, Gen W, Hongbo W, Chongqing X, Fan Z, Yanqi F et al (2020) A retrospective evaluation of reliability and reproducibility of Arbeitsgemeinschaft für Osteosynthesefragen classification and Fernandez classification for distal radius fracture. Med (Baltim) 99:e18508. https://doi.org/10.1097/MD.0000000000018508
5. Jayakumar P, Teunis T, Giménez BB, Verstreken F, Di Mascio L, Jupiter JB (2017) AO distal radius fracture classification: global perspective on observer agreement. J Wrist Surg 6:46–53. https://doi.org/10.1055/s-0036-1587316
6. Carula BC, Giostri GS, Bebber F, Batista MFDS, da Silva FB, Novak EM (2022) Evaluation of intra- and interobserver reproducibility of the new AO/OTA classification for distal radius fractures compared with the Fernandez classification. Rev Bras Ortop (Sao Paulo) 57:917–923. https://doi.org/10.1055/s-0041-1740471
7. Küçük L, Kumbaracı M, Günay H, Karapınar L, Özdemir O (2013) Reliability and reproducibility of classifications for distal radius fractures. Acta Orthop Traumatol Turc 47:153–157. https://doi.org/10.3944/AOTT.2013.3038
8. Heo YM, Roh JY, Kim SB, Yi JW, Kim KK, Oh BH et al (2012) Evaluation of the sigmoid notch involvement in the intra-articular distal radius fractures: the efficacy of computed tomography compared with plain X-ray. Clin Orthop Surg 4:83–90. https://doi.org/10.4055/cios.2012.4.1.83
9. Dankelman LHM, Barvelink B, Verhofstad MHJ, Wijffels MME, Colaris JW (2024) Traditional radiography versus computed tomography to assess reduced distal radius fractures. Eur J Trauma Emerg Surg 50:2313–2321. https://doi.org/10.1007/s00068-024-02598-5
10. Kleinlugtenbelt YV, Groen SR, Ham SJ, Kloen P, Haverlag R, Simons MP et al (2017) Classification systems for distal radius fractures: does the reliability improve using additional computed tomography? Acta Orthop 88:681–687. https://doi.org/10.1080/17453674.2017.1338066
11. Arealis G, Galanopoulos I, Nikolaou VS, Lacon A, Ashwood N, Kitsis C (2014) Does the CT improve inter- and intra-observer agreement for the AO, Fernandez and Universal classification systems for distal radius fractures? Injury 45:1579–1584. https://doi.org/10.1016/j.injury.2014.06.017
12. Sharma M, Choudhury SR, Soundararajan R, Sheth R, Sinha A, Prakash M (2024) Reliability of distal radius fracture classification systems: a CT based study. Emerg Radiol 31:873–879. https://doi.org/10.1007/s10140-024-02294-2
13. Eriksson S, Rasool F, Dånmark I, Weber Lensing C, Werner J, Fornander L (2025) Cone-beam computed tomography influences the classification and choice of treatment for distal radius fractures. J Wrist Surg 14:333–341. https://doi.org/10.1055/s-0044-1782236
14. Kamal RN, Shapiro LM (2022) Practical application of the 2020 distal radius fracture AAOS/ASSH clinical practice guideline: a clinical case. J Am Acad Orthop Surg 30:e714–e720. https://doi.org/10.5435/JAAOS-D-21-01194

15. Nguyen SA, Dang AH, Tran DQ (2026) Distal radius fracture classifications in real life: reliability and how they change treatment. J Hand Surg Glob Online 8:100910. https://doi.org/10.1016/j.jhsg.2025.100910
16. Walenkamp MMJ, Aydin S, Mulders MAM, Goslings JC, Schep NWL (2016) Predictors of unstable distal radius fractures: a systematic review and meta-analysis. J Hand Surg Eur Vol 41:501–515. https://doi.org/10.1177/1753193415604795
17. Durkalski VL, Palesch YY, Lipsitz SR, Rust PF (2003) Analysis of clustered matched-pair data. Stat Med 22:2417–2428. https://doi.org/10.1002/sim.1438
18. Neuhaus V, Bot AG, Guitton TG, Ring DC (2015) Radiographs versus radiographic measurements in distal radius fractures. J Hand Microsurg 7:42–48. https://doi.org/10.1007/s12593-014-0164-0
19. Davidson A, Feldman G, Mosheiff R, Suna A, Joskowicz L, Weil YA (2024) Computer-generated radiographic measurements of distal radius fractures: does it help with decision making? J Hand Surg Am 49:796e1–796e7. https://doi.org/10.1016/j.jhsa.2022.09.015
20. Kalantar SH, Afshar A, Shariyate MJ, Nazarian A, Farhoud AR, Bagheri N et al (2025) The impact of post-reduction radiographs of distal radius fractures on treatment decisions among fellowship-trained orthopedic hand surgeons and general orthopedic surgeons. Arch Bone Jt Surg 13:643–648. https://doi.org/10.22038/ABJS.2025.81032.3696
21. Walenkamp MMJ, Vos LM, Strackee SD, Goslings JC, Schep NWL (2015) The unstable distal radius fracture—how do we define it? A systematic review. J Wrist Surg 4:307–316. https://doi.org/10.1055/s-0035-1556860
22. Mulders MAM, Detering R, Rikli DA, Rosenwasser MP, Goslings JC, Schep NWL (2018) Association between radiological and patient-reported outcome in adults with a displaced distal radius fracture: a systematic review and meta-analysis. J Hand Surg Am 43:710–719e5. https://doi.org/10.1016/j.jhsa.2018.05.003
23. Kaymakoglu M, Kolac UC, Bahsi A, Filibeli M, Turken MA, Dogan K et al (2025) Toward a new era in fracture imaging: zero echo time MRI vs CT in intra-articular distal radius fractures—a proof-of-concept study. Injury 56:112846. https://doi.org/10.1016/j.injury.2025.112846
24. O'Malley MP, Rodner C, Ritting A, Cote MP, Leger R, Stock H et al (2014) Radiographic interpretation of distal radius fractures: visual estimations versus digital measuring techniques. Hand (N Y) 9:488–493. https://doi.org/10.1007/s11552-014-9666-2
25. Suojärvi N, Sillat T, Lindfors N, Koskinen SK (2015) Radiographical measurements for distal intra-articular fractures of the radius using plain radiographs and cone beam computed tomography images. Skeletal Radiol 44:1769–1775. https://doi.org/10.1007/s00256-015-2231-4

Figures



	Radiographs	Radiographs + CT	Δκ
AO/OTA type	0.379 (0.296–0.462)	0.356 (0.271–0.441)	−0.023
AO/OTA group	0.184 (0.137–0.232)	0.169 (0.120–0.218)	−0.015
Universal	0.461 (0.369–0.554)	0.386 (0.298–0.474)	−0.075
Fernandez	0.347 (0.281–0.413)	0.332 (0.265–0.399)	−0.015
Frykman	0.290 (0.244–0.336)	0.311 (0.258–0.364)	+0.021
Treatment decision	0.303 (0.219–0.388)	0.294 (0.193–0.394)	−0.010

Figure 1

Interobserver agreement (Fleiss' κ with asymptotic 95% confidence intervals) for the four classification systems and for the treatment decision with radiographs alone (T0, filled circles and solid lines) and radiographs plus computed tomography (T4, open squares and dashed lines). Shaded bands indicate the Landis and Koch categories. Δκ, difference between T4 and T0; CT, computed tomography