



Contents lists available at ScienceDirect

Journal of Hand Surgery Global Online

journal homepage: www.JHSGO.org

Original Research

Inter- and Intraobserver Reliability for the Computed Tomography Scan Assessment of Union After Surgery for Scaphoid Fractures and Nonunion

Morten Kjær, MD, ^{*} Dimitar Ivanov Radev, MD, [†] Robert Gvozdenovic, MD, PhD ^{*,‡}^{*} Department of Orthopedic Surgery, Hand Surgery Unit, Herlev-Gentofte University Hospital of Copenhagen, Copenhagen, Denmark[†] Department of Radiology, Bispebjerg and Frederiksberg University Hospital of Copenhagen, Copenhagen, Denmark[‡] Faculty of Health and Medical Sciences, Institute for Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

ARTICLE INFO

Article history:

Received for publication August 9, 2024

Accepted in revised form August 11, 2024

Available online September 14, 2024

Key words:

Agreement

Reliability

Scaphoid fracture

Scaphoid nonunion

Union

Purpose: Assessment of scaphoid union after operative treatment might be associated with uncertainty, disagreement, and misinterpretation. Union is defined as consolidation on three of four X-ray views. Computed tomography (CT) scans are increasingly used to evaluate union of scaphoid fractures and nonunion. It allows multiplanar assessment of the scaphoid. Reliability in CT scan assessment of union after operative intervention is limited. We hypothesized that inter- and intraobserver reliability of the CT scan assessment of union after operative treatment for scaphoid fracture and nonunion are substantial and better in fracture cases.

Methods: An institutional search identified 230 patients with operative intervention. According to the sample size, we randomly selected 60 patients (30 fractures and 30 nonunion), with age above 18 years and CT scans 6–26 weeks after surgery, when uncertainty about healing status is most pronounced. Three observers evaluated the CT scans on two occasions 6 weeks apart. Observers classified the scaphoid into two categories: No/partial/full union and $>/<50\%$ union.

Results: Interrater agreement in no/partial/full union was overall substantial, substantial in fractures, and moderate for nonunion cases. Interobserver agreement for $>/<50\%$ union was overall moderate and moderate for both fractures and nonunion groups. The interrater agreement of no/partial/full union with CT scans ≥ 12 weeks after surgery was substantial versus fair for the cases below. Intrarater agreement was overall better in cases with CT scans ≥ 12 weeks postoperative versus earlier.

Conclusions: CT scans show substantial agreement in the assessment of no/partial/full union after operative intervention, although the presence of metal artifacts and bone grafts potentially impair visibility among observers. The agreement was better in scaphoid fractures. CT scan assessment of $>/<50\%$ union is with overall moderate agreement. Inter- and intrarater agreement is improved when CT scans are taken 12 weeks or later after operative treatment, blinded to the clinical information.

Level of Evidence: Diagnostic, III.

Copyright © 2024, THE AUTHORS. Published by Elsevier Inc. on behalf of The American Society for Surgery of the Hand.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

The scaphoid is a complex cylindrical and angulated bone with retrograde blood supply. The surface is primarily covered with cartilage and unites with primary healing; thus, external callus cannot be detected.¹ Assessment of union after surgery for scaphoid fractures and nonunion is often debatable between surgeons and

radiologists. Determination of the union relies on both clinical evaluation and radiographic imaging. Uncertain union status prolongs the period with supportive brace, and functional restrictions. X-ray in four different views are commonly used to evaluate union after surgery. Signs of union include trabecular crossing and are primarily defined by bone consolidation on at least three of four views.² It is readily available and affordable and involves a minimal radiation dose. However, inter- and intraobserver reliability is reported to be fair/moderate in the evaluation of scaphoid fractures treated conservatively.³ The shape of the scaphoid complicates

Corresponding author: Morten Kjær, MD, Herlev-Gentofte University Hospital of Copenhagen, Hospitalsvej 1. Opg. 17, 2900 Hellerup, Denmark.

E-mail address: Morten.kjaer.02@regionh.dk (M. Kjær).

<https://doi.org/10.1016/j.jhsg.2024.08.005>

2589-5141/Copyright © 2024, THE AUTHORS. Published by Elsevier Inc. on behalf of The American Society for Surgery of the Hand. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Table 1
Demographic Data of Patients with Postoperative CT Scan After Treatment for Scaphoid Fracture (n=30) and Scaphoid Nonunion (n=30)

Demographic Data	
Sex (M)	48/60
Age (mean)	32 y (range 18–78)
Fracture/nonunion location	
Proximal/waist/distal	15/43/2
Procedure for scaphoid fracture	
ORIF/percutaneous/arthroscopic	17/10/3
Procedure for scaphoid nonunion	
Arthroscopic/open bone graft reconstruction and fixation	13/17
Surgical delay from injury (mean)	
Fracture/nonunion	47/533 (days)

ORIF, open reduction internal fixation.

X-rays in capturing the fracture and nonunion site in the correct plane, whereas overlying structures often lead to misinterpretation.⁴ Furthermore, bone bridging is difficult, if not impossible, to differentiate from graft material when there is good compression at the nonunion site with a screw present. Therefore, the union is typically defined by the absence of clinical symptoms and adverse features on X-rays, such as a gap at the fracture/nonunion site, cystic changes, or signs of implant movement or loosening.²

Computed tomography (CT) scans are accurate and reliable in diagnosing a scaphoid fracture and displacement.^{5,6} Cystic changes and deformity are visualized in nonunion cases. CT scans are increasingly used routinely to evaluate union after operative treatment.^{7–9} It allows multiplanar reconstruction and assessment of the trabecular architecture at the fracture/nonunion site. Union is defined as $\geq 50\%$ bone bridging across the nonunion site, which in a biomechanical cadaver model has a load to failure compared to an intact scaphoid.¹⁰ The reliability of CT scans for assessing union in conservatively treated scaphoid fractures is reported to be superior to that of X-rays. However, after operative treatment, its reliability has been investigated in only one study.⁹ Evaluation of union in those cases is potentially complicated by metal artifacts; in nonunion cases, bone graft may be another disturbing factor in differentiating bone bridging from good compression at the nonunion site.

This study aims to report the reliability of the CT scan assessment of union after operative treatment of scaphoid fractures and nonunion. We hypothesized that inter- and intraobserver reliability of the CT scan assessment of union after operative treatment for scaphoid fracture and nonunion is substantial, with a superior agreement in fractures.

Materials and Methods

Approvals from the Danish National Committee on Health Research Ethics Danish (ID No. 2204737) and the Danish Data Protection Agency (Privacy) were obtained prior to the initiation of the study. We conducted an institutional search for patients with operative intervention from 2013 to 2021 for scaphoid fractures and for graft reconstruction with screw fixation for scaphoid nonunion. Inclusion criteria were age > 18 years, operative intervention for displaced scaphoid fractures, and graft reconstruction with screw fixation for scaphoid nonunion, both using a single headless compression screw. A total of 230 patients were identified. Of those, 153 patients were excluded because of the absence of CT scans or because their CT scans were obtained more than 26 weeks after the operative intervention. Additionally, patients were excluded if they had concomitant wrist injuries, if the intervention was a secondary surgery, or if the fixation method differed from the

inclusion criteria. The study aims to investigate agreement within the clinical window for detecting union, when the uncertainty about healing status is most pronounced. To reduce selection bias and align with sample size requirements, 17 patients who had CT scans performed before 2017, when CT scan was not standard practice, were randomly excluded.¹¹ This leaves a total of 30 patients who underwent screw fixation for scaphoid fracture and 30 patients who underwent bone grafting and screw fixation for scaphoid nonunion. An independent radiographer created two anonymized lists of serial CT scans in the axial plane, sagittal plane of the scaphoid axis, and coronal plane with slices of 1 mm in a random order between rating rounds. Three observers, including a resident hand surgeon (rater one), musculoskeletal radiologist (rater two), and a senior hand surgeon (rater three), evaluated the CT scans on two separate occasions that were 6 weeks apart, and the raters were blinded to all clinical information. Observers were allowed to scroll through the CT images without time restrictions. Observers were tasked to classify the scaphoid as united or not united based on the criteria of $>/<50\%$ bone bridging as reported by Singh et al.⁴ Another rating system was established to classify the operated scaphoid into three categories: no union, partial union, and full union. No union was defined with the absence of bone bridging, partial union as partial bone bridging (without giving the percentage), and full union as complete bone bridging across the entire fracture/nonunion site.

Statistical analysis

A sample size analysis was conducted according to the methodology of Donner et al.¹¹ We considered interobserver agreement of $k_0 = 0.8$ among three observers.^{5,12} We estimated the expected detection of healing as 50% , $\pi = 0.50$, in cases with CT scans within 6–26 weeks after surgery. The minimal acceptable limit was specified the agreement of $k_L = 0.60$, resulting in a sample size of 26 for each group.

Descriptive statistics are reported as median/mean with range. The multirater kappa measure was used to estimate the chance-corrected agreement among the observers in two categories no/partial/full union and $>/<50\%$ union, with subgroup kappa values for fractures and for the nonunion. We interpreted the values using the guidelines proposed by Landis and Koch.¹³ Values of 0.01–0.2 indicate slight agreement, 0.21–0.4 fair agreement, 0.41–0.6 moderate agreement, 0.61–0.8 substantial agreement, and 0.81–1.0 almost perfect agreement.

Results

The data set consisted of 60 cases, 30 with headless screw fixation for a scaphoid fracture and 30 with a graft reconstruction and headless screw fixation for a scaphoid nonunion (Acutrak, Acumed™, Hillsboro, USA). Demographics are presented in Table 1. The CT scans used for the union assessment were taken at a median of 11.6 weeks (range 6.0–25.3 weeks) after surgery. In this sample, 51 out of 60 patients achieved union (26 scaphoid fractures, 25 scaphoid nonunion). Forty-five patients had X-ray/CT scans available showing full union at a mean 9.7 months (range 3–25 months) after operative intervention. The remaining six were classified as united at their last follow-up based on CT scans showing partial union, absence of pain, and good wrist function.

The interrater agreement for the evaluation of no/partial/full union was overall substantial (Table 2). Interrater agreement for no union was almost perfect ($\kappa = 0.81$), moderate for partial union ($\kappa = 0.57$), and substantial for union ($\kappa = 0.61$). Fracture cases had overall substantial agreement, whereas the agreement was moderate in nonunion cases (Table 2). Interrater agreement for the

Table 2

Overall Interrater Agreement and Pairwise Agreement Between Raters for the CT Scan Assessment of Union After Operative Intervention for Scaphoid Fracture and Nonunion

	Interrater Agreement No/Partial/Full Union			Interrater Agreement >/<50% Union	
	Kappa	CI		Kappa	CI
All Cases	0.62	0.50–0.74	All Cases	0.50	0.35–0.64
Rater 1+2	0.78		Rater 1+2	0.70	
Rater 1+3	0.53		Rater 1+3	0.45	
Rater 2+3	0.57		Rater 2+3	0.39	
Fracture n = 30	0.69	0.52–0.86	Fracture n = 30	0.48	0.27–0.69
Rater 1+2	0.77		Rater 1+2	0.72	
Rater 1+3	0.60		Rater 1+3	0.33	
Rater 2+3	0.71		Rater 2+3	0.42	
Nonunion n = 30	0.50	0.34–0.65	Nonunion n = 30	0.50	0.29–0.71
Rater 1+2	0.78		Rater 1+2	0.68	
Rater 1+3	0.42		Rater 1+3	0.57	
Rater 2+3	0.34		Rater 2+3	0.34	

Table 3

Intrater Agreement Overall (n = 60), Fracture (n = 30), Nonunion (n = 30), Patients with CT Scan <12 weeks (n = 33), and Patients with CT Scan ≥ 12 Weeks after Operative Intervention (n = 27)

Intrater Agreement	Kappa	CI
Overall	0.71	0.63–0.77
Fracture	0.68	0.58–0.88
Nonunion	0.72	0.62–0.82
CT scan < 12 weeks	0.58	0.46–0.69
CT scans ≥ 12 weeks	0.82	0.73–0.90

evaluation of >/<50% union was overall moderate, both in the evaluation for scaphoid fracture and scaphoid nonunion (Table 2). Agreement was consistently substantial between rater 1 and 2 ($\kappa = 0.68$ –0.78), and agreement with rater 3 varied from fair to substantial ($\kappa = 0.33$ –0.71) (Table 2).

Intrater agreement was overall substantial both in scaphoid fractures and nonunion (Table 3).

A post hoc subgroup analysis compared cases with CT scans taken within 12 weeks after the operative intervention (n = 33) to those scanned at 12 weeks or later (n=27). In this analysis, evaluation of >/< 50% union showed moderate kappa values in both fracture and nonunion groups, with higher kappa values observed for patients ≥12 weeks (Figure 1). For the assessment of no/partial/full union, interrater agreement was fair for cases <12 weeks after surgery and substantial for cases ≥12 weeks after surgery (Figure 1). Intrater agreement was moderate for cases <12 weeks versus almost perfect in patients ≥12 weeks after operative intervention (Table 3).

Discussion

This study demonstrates that CT scans show substantial agreement in assessing whether there is no/partial/full union within the clinically relevant time window for expected union, which can be used reliably in patients following operative intervention for scaphoid fracture and scaphoid nonunion.

Matzon et al⁹ reported overall moderate agreement ($\kappa = 0.59$) in 32 patients who underwent operative treatment (16 fractures/16 nonunion). Agreement was not improved when both X-ray and CT scans were available, although surgeons believed more certain in their evaluations. To our knowledge, this study is the only report on the reliability of assessing union after operative treatment. Hanneman et al⁵ reported a moderate overall interrater agreement ($\kappa = 0.57$) in categorizing no/partial/full union in 44 patients with conservatively treated scaphoid fractures, slightly lower than our findings ($\kappa = 0.62$). They observed substantial agreement for cases

of no union and union, whereas agreement for partial union was moderate. Hence, the agreement in each type of consolidation might be underpowered because of the sample size. In our study agreement for no union was almost perfect, and similar with moderate agreement for partial union and substantial agreement for full union.

When evaluating no union versus union in conservatively treated scaphoid fractures, Hanneman et al⁵ reported substantial agreement ($\kappa = 0.66$), consistent with the findings by Buijze et al¹² ($\kappa = 0.69$) using sagittal reconstructions only and among 59 observers. In these studies, union was subjectively defined by the raters, whereas our study used a criterion of >/<50% union, resulting in moderate agreement ($\kappa = 0.50$). Despite using CT scans with imaging metal artifact reduction techniques, metal artifacts from the headless compression screw can complicate the assessment of union in scaphoid fractures and especially in cases of scaphoid nonunion with graft material. Therefore, a direct comparison of the findings is not fully appropriate.

In evaluating no/partial/full union, we found better agreement in fracture cases (substantial) compared to nonunion cases (moderate). In the subgroup analysis of scaphoid fractures and nonunion by Matzon et al,⁹ agreement was similar to our findings. However, the sample size of their study was too small in calculating kappa values for each group independently, and imaging was performed at 1.3–19 months after operative intervention, compared to 2–6 months in our study.

Considering the timing of the CT scan, the agreement approaches a κ value of 1.0 as the CT scan is performed later, as the union is expected to be more visible. Vice versa, nonunion is easier defined when a gap is visible in the longer term after the operative treatment. In subgroup analysis we found improved inter- and intrater agreement in CT scans taken ≥12 weeks after operative intervention versus those taken <12 weeks. In contrast, Hanneman et al⁵ reported moderate agreement in CT scans at week 6 and 12 and substantial agreement in patients 24 weeks after conservatively treated scaphoid fractures. However, the sample size for each time event in this study was too small in calculating kappa values. Farracho et al¹⁴ suggested cone beam CT scan is accurate and reliable 6 weeks after treatment (12/52 patients had operative intervention for a scaphoid fracture) and reported substantial agreement in >/<50% union ($\kappa = 0.64$). For comparison, our results in scaphoid fractures and nonunion cases <12 weeks after operative intervention showed the agreement was moderate ($\kappa = 0.44$). Generally, a CT scan taken less than 12 weeks after operative intervention is valuable for detecting patients at risk of nonunion, such as those with displacement, implant movement, and deformity.¹⁵ Our results suggest that evaluating union and consolidation

INTERRATER AGREEMENT TIME DEPENDENT

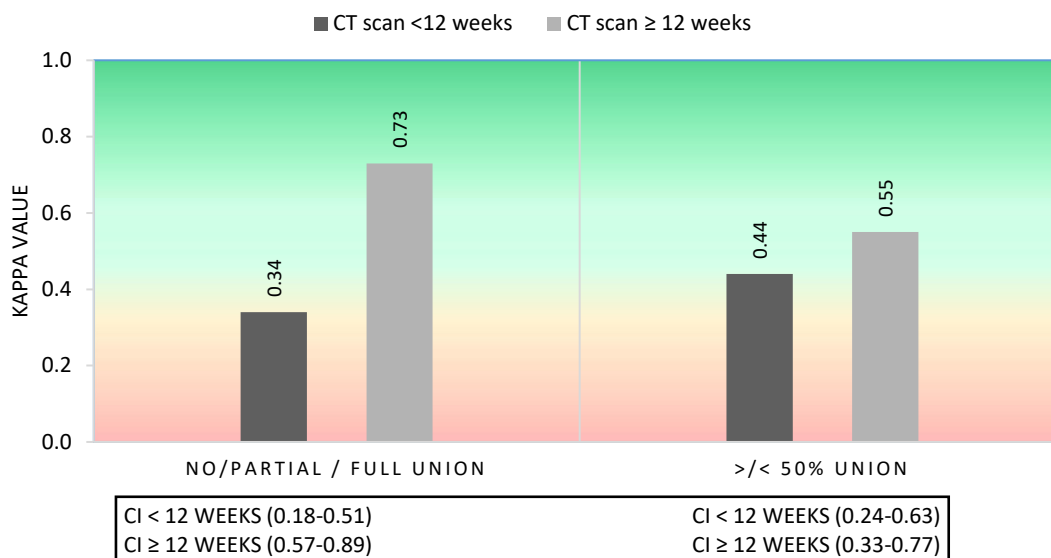


Figure 1. Interrater agreement for each category, separated for patients with CT scans <12 weeks (n = 33) and ≥12 weeks (n = 27) after operative intervention.

type yields superior inter- and intrarater agreement, when the CT scans are 12 weeks or more after operative intervention.

In this study, the pairwise interrater agreement was consistently substantial between rater one and two ($\kappa = 0.68$ – 0.78). However, the agreement with rater three was distinctly lower, especially when evaluating >/<50% union, because it was only fair to moderate. This discrepancy might be attributed to bias from the senior hand surgeon's (rater three) clinical experience in cases of partial union, who usually progress to full union, as he tends to assess more patients with >50% union.⁴ However, Hannemann et al⁵ also reported fair-moderate agreement in 25% to 74% consolidation ($\kappa = 0.22$ – 0.42), and Matzon et al⁹ reported fair agreement ($\kappa = 0.31$) in 25% to 50% union.

The study has both limitations and strengths. In two patients, only CT scans in the longitudinal plane of the radius was available. The plane of the scaphoid long axis better visualize the fracture/nonunion site.⁷ Second, including clinical information would better reflect clinical practice and assist the observer in determining the degree of union. However, we investigated the reliability of the CT scans alone, with observers blinded to all clinical information, thus testing the CT scan as an independent diagnostic tool. Third, not all scaphoid fractures and nonunion received CT scans following operative treatment. Presumably, these were patients without pain symptoms and with satisfying bone healing addressed with the X-ray alone. This could imply exclusion of simple cases from the sample population. We believe the degree of selection bias is limited, as CT scans have been standard practice in our department since 2017, and only eight out of 60 patients in our sample were surgically treated before that date. Nevertheless, demographic data seems to be comparable to the general population of scaphoid fractures and nonunion in terms of union rate, sex, and age, and the study's sample size was appropriate.¹⁶

In conclusion, we found that CT scans demonstrate substantial agreement, in the assessment of no/partial/full union, after operative intervention for scaphoid fractures and nonunion, in which the presence of metal artifacts and bone grafts potentially deteriorates the vision among observers. The agreement was better in scaphoid fractures. Agreement in the assessment of >/<50% union was moderate. Inter- and intrarater agreement improves when CT scans

were taken 12 weeks or more after operative intervention, blinded to clinical information. To ensure that the postoperative period with a cast and functional restrictions is limited to what is necessary, this study supports the reliability of using CT scans and contributes to well-considered decision making in the union assessment after scaphoid fracture and nonunion surgery.

Conflicts of Interest

The study is a part of a Ph. D. project funded by the Department of Orthopedic Surgery, Herlev/Gentofte University Hospital, Copenhagen, Denmark. The funder does not have any role in the study design, execution of data analysis or submission of results. The research foundation at Herlev and Gentofte University Hospital, Copenhagen, Denmark has granted a scholarship of 6-month salary to the PhD project.

Acknowledgments

The author would acknowledge Jonas Alexander Larsen, Radiographer, Department of Radiography, Herlev-Gentofte University Hospital for his commitment and contribution to the study with gathering of CT scans and creation of randomized and blinded lists.

References

- Hackney LA, Dodds SD. Assessment of scaphoid fracture healing. *Curr Rev Musculoskelet Med.* 2011;4(1):16–22.
- Dias JJ. Definition of union after acute fracture and surgery for fracture nonunion of the scaphoid. *J Hand Surg Br.* 2001;26(4):321–325.
- Dias JJ, Taylor M, Thompson J, Brenkel IJ, Gregg PJBrenkel IJ, Gregg PJ. Radiographic signs of union of scaphoid fractures. An analysis of inter-observer agreement and reproducibility. *J Bone Joint Surg Br.* 1988;70(2):299–301.
- Singh HP, Forward D, Davis TR, Dawson JS, Oni JA, Downing ND. Partial union of acute scaphoid fractures. *J Hand Surg Br.* 2005;30(5):440–445.
- Hannemann PF, Brouwers L, van der Zee D, et al. Multiplanar reconstruction computed tomography for diagnosis of scaphoid waist fracture union: a prospective cohort analysis of accuracy and precision. *Skeletal Radiol.* 2013;42(10):1377–1382.
- Daniels AM, Wyers CE, Janzing HMJ, et al. The interobserver reliability of the diagnosis and classification of scaphoid fractures using high-resolution peripheral quantitative CT. *Bone Joint J.* 2020;102-B(4):478–484.

7. Bain GI, Bennett JD, Richards RS, Slethaug GP, Roth JH. Longitudinal computed tomography of the scaphoid: a new technique. *Skeletal Radiol*. May 1995;24(4): 271–273.
8. Drijkoningen T, Ten Berg PWL, Guitton TG, et al. Reliability of diagnosis of partial union of scaphoid waist fractures on computed tomography. *J Hand Microsurg*. 2018;10(3):130–133.
9. Matzon JL, Lutsky KF, Tulipan JE, et al. Reliability of radiographs and computed tomography in diagnosing scaphoid union after internal fixation. *J Hand Surg Am*. 2021;46(7):539–543.
10. Guss MS, Mitgang JT, Sapienza A. Scaphoid healing required for unrestricted activity: a biomechanical cadaver model. *J Hand Surg Am*. 2018;43(2): 134–138.
11. Donner A, Rotondi MA. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *Int J Biostat*. 2010;6(1):Article 31.
12. Buijze GA, Wijffels MM, Guitton TG, et al. Interobserver reliability of computed tomography to diagnose scaphoid waist fracture union. *J Hand Surg Am*. 2012;37(2):250–254.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
14. Farracho LC, Moutinot B, Neroladaki A, et al. Determining diagnosis of scaphoid healing: Comparison of cone beam CT and X-ray after six weeks of immobilization. *Eur J Radiol Open*. 2020;7:100251.
15. Grewal R, Suh N, Macdermid JC. Use of computed tomography to predict union and time to union in acute scaphoid fractures treated nonoperatively. *J Hand Surg Am*. 2013;38(5):872–877.
16. Duncumb JW, Robinson PG, Williamson TR, et al. Bone grafting for scaphoid nonunion surgery: a systematic review and meta-analysis. *Bone Joint J*. 2022;104-B(5):549–558.