CrossMark

HANDSURGERY

The required minimum length of video sequences for obtaining a reliable interobserver diagnosis in wrist arthroscopies

Steffen Löw[1] · Holger Erne[2] · Alexander Schütz[3] · Christoph Eingartner[1] · Christian K. Spies[4]

Abstract

Purpose   To examine the relationship between video length for wrist arthroscopy and interobserver reliability.

Materials and methods   100 consecutive wrist arthroscopies were documented by long and short videos of the radiocarpal and the midcarpal joints. The long videos were about twice as long as the short videos. They were presented randomly to two independent and blinded examiners. Their diagnoses were compared to the diagnoses made by the surgeon who performed the arthroscopies. Kappa coefficients were calculated.

Results   Kappa statistics were inconsistent and did not show that the long video provided an obvious advantage over the short video. The Kappa coefficients of the two examiners for the assessment of the cartilage status were 0.524 and 0.700 for the long videos and 0.465 and 0.639 for the short videos, respectively. The examiners diagnosed twice as many false-positive cartilage lesions on short videos than on long videos. The assessment of ligament lesions was more accurate on long than on short videos.

Conclusions   The results confirmed the hypothesis that the reproducibility of diagnoses based on video documents was influenced by the length of the video sequences. Therefore, it may be advisable for video documentation to be done diligently. The video sequence of the radiocarpal joint should last about 60 s, and that of a midcarpal joint should last about 45 s. Videos of difficult joints should last appropriately longer.

Level of evidence   Diagnostic II.

Keywords   Interobserver reliability · Scapholunate ligament · Triangular fibrocartilage complex · Video documentation · Wrist arthroscopy

✉ Steffen Löw
   steffen_loew@hotmail.com

   Holger Erne
   holger.erne@mri.tum.de

   Alexander Schütz
   alexander.schuetz2@googlemail.com

   Christoph Eingartner
   christoph.eingartner@ckbm.de

   Christian K. Spies
   christian.spies@vulpiusklinik.de

1   Section of Hand Surgery, Division of Trauma and Orthopedic Surgery, Caritas Krankenhaus, Uhlandstraße 7, 97980 Bad Mergentheim, Germany

2   Department for Plastic and Hand Surgery, Klinikum rechts der Isar TUM, Ismaninger Straße 22, 81675 Munich, Germany

3   Sporthopaedicum, Hildegard-von Bingen-Str. 1, 93053 Regensburg, Germany

4   Department of Hand Surgery, Vulpius Klinik, Vulpiusstrabe 29, 74906 Bad Rappenau, Germany

Introduction

Arthroscopy is an important tool in the diagnosis and treatment of wrist pathologies [1–4]. Imaging is mandatory for documentation and reproducibility of the diagnoses. If the diagnosis is doubtful despite photo documentation [5, 6] performed by one surgeon, in certain cases it could be better to repeat a wrist arthroscopy when the patient will be treated again by another surgeon. Attempts have been made to improve the reproducibility by adding videos to photo documents. However, in a study of Löw et al. [7] videos were criticized for being too short to detect the correct diagnosis.

🔗 Springer

The purpose of this study was to analyze the relationship between video length and interobserver reliability for detection of intra-articular structures and/or lesions assessed by two independent examiners.

The hypothesis of this study was that the interobserver reliability would be higher based on long video documents than based on short video documents. We further expected a lower rate of false-positive cartilage lesions after reviews of longer videos compared to reviews of shorter videos.

Materials and methods

The study was approved by the institutional ethics committee of the local university. In a prospective study, videos of 100 consecutive wrist arthroscopies (57 right, 43 left) were made by the first author. Indications for the arthroscopies were the evaluation of the scapholunate (SL) ligament in eight patients, the triangular fibrocartilage complex (TFCC) in 74 patients, and to assess other issues (such as cartilage status) in 18 patients. All operations were done in a standardized manner as described by Löw et al. [1] under axillary plexus anesthesia and by tourniquet control with 4 kg of axial traction. The arthroscopy medium was 0.9 % saline. The 3–4 and 4–5 radiocarpal portals and the radial and ulnar midcarpal portals were used. The optic was always inserted radially. Intra-articular structures were visualized using a 2.7 mm 30° optic (Stryker, Kalamazoo, MI, USA). Two videos of the wrist showing the radiocarpal (cartilage of the scaphoid, lunate, scaphoid fossa, lunate fossa; palmar radiocarpal ligaments; scapholunate and lunotriquetrum (LT) ligament; triangular fibrocartilage complex; synovia), and midcarpal (cartilage of the capitate, hamate, triquetrum; SL and LT joint gap) structures were produced. The first video was twice as long as the second video. Cartilage lesions were classified according to Outerbridge, with grades 2 and 3 combined into one grade [8]. Tears of the palmar radiocarpal, and the SL and LT ligaments were graded as partial or complete. Additionally, the SL ligament was graded according to its examination by the probe from the midcarpal joint. Grading was performed according to whether the probe could or could not be inserted into the SL interval, whether the probe could be turned inside the interval, and whether the optic could be inserted into the SL interval. Grading was based on Geissler's classification [9] excluding its radiocarpal assessment. The TFCC was assessed for a trampoline effect [10] and a lesion was classified according to Palmer [11]. If existing, a synovitis was rated separately as present or absent.

The first author randomly mixed the 200 pairs of long and short videos and presented them for assessment by two independent surgeons (hereafter "examiners") who were highly experienced in wrist arthroscopy. The examiners were neither involved in the treatment of the patients nor informed about the study's hypothesis. The videos were presented pseudonymized, to blind the examiners as well as the first author to the actual diagnosis (hereafter "surgeon"). The examiners were asked to assess all intra-articular structures/lesions mentioned above and give a diagnosis, which was compared to the actual diagnosis of the surgeon.

Statistical methods

Since the Kolmogorov–Smirnov test revealed significant deviation from normal distribution, the lengths of the video sequences were compared using the Wilcoxon test. Cohen's Kappa coefficients were calculated to measure the interobserver agreement. Concerning the cartilage status, Kappa coefficients were calculated for the presence or absence of a cartilage lesion anywhere in the joint and separately for each carpal bone. For the latter, Kappa coefficients were also calculated for the assessment according to Outerbridge. Additionally, the rates of false-positive estimations including the 95 % confidence intervals were calculated. The contingency tables were analyzed for specifics that influenced Kappa statistics. Relevant findings were outlined in a descriptive manner.

Results

The prevalence of intra-articular pathologies—cartilage and ligament lesions—according to the surgeon's diagnosis is outlined in Table 1.

Length of the video documents

The median length of long videos was 56.50 s for the radiocarpal and 41.50 s for the midcarpal joint. These lengths were approximately twice as long as the short video sequences, which lasted 26.50 and 23.00 s, respectively. The difference was statistically significant ($P < 0.001$) as shown in Fig. 1a, b.

Cartilage

Comparing long with short videos, according to Kappa statistics no difference was observed. Assessing the Outerbridge classification of each carpal bone, Kappa coefficients indicated inhomogeneous results ranging from poor to substantial agreement (Tables 2, 3). Both examiners rated twice as many false-positive cartilage lesions on

Table 1 Intra-articular findings according to the diagnoses of the surgeon

	Intact	Grade 1	Grade 2/3	Grade 4
Radiocarpal cartilage status (outerbridge)				
Scaphoid Fossa	86	7	4	3
Scaphoid	88	5	4	3
Lunate Fossa	82	10	5	3
Lunate	74	3	8	15
Midcarpal cartilage status (outerbridge)				
Scaphoid	97	1	2	0
Lunate	97	1	1	1
Capitate	85	8	5	2
Hamate	70	9	13	8
Triquetrum	96	2	1	1
Ligaments				
SL ligament	78× intact	12× partial tear	10× complete tear	
SL findings from midcarpal	34× probe not insertable	21× probe insertable	29× probe can be turned inside SL gap	9× optic insertable
LT ligament	84× intact	15× partial tear	1× complete tear	
Palmar radiocarpal ligaments	98× intact	1× partial tear	1× complete tear	
TFCC				
Trampoline sign [10]	48× positive		52× negative	
Palmer's classification [11]	12× intact			
	4× 1A; 9× 1B			
	26× 2A; 1× 2B; 34× 2C; 9× 2D; 1× 2E			
	Combined lesions: 1× 1A/2C, 1× 1B/2A; 2× 1B/2C			
Presence of synovitis				
26× radiocarpal	65× ulnocarpal		34× midcarpal	

short than on long videos (Table 4). The differences were not statistically significant.

Table 5 shows Kappa coefficients calculated for assessment of SL, LT, palmar radiocarpal ligaments, TFCC, and synovitis. The following paragraphs outline the specifics of the contingency tables that help explain Kappa statistics results.

SL ligament

Kappa coefficients revealed moderate to almost perfect agreement with no obvious advantage of long video sequences. Examiner 1 identified six of the ten complete SL ruptures by assessing short videos, but he diagnosed nine of the 10 complete ruptures correctly assessing long videos. For examiner 2, there was no difference. In 94 and 103 out of all 200 midcarpal videos, respectively, the two examiners observed that the probe could be inserted or even turned inside the SL joint from midcarpal. Examiner 1 rated this as "partial tear" in 32 cases and as "complete tear" in one case, whereas examiner 2 rated this as "partial tear" in 22 cases and as "complete tear" in three cases.

Both examiners rated the insertion of the optic into the SL joint as a complete SL lesion.

LT ligament

The assessment of the LT joint was more accurate using long videos compared to short videos (Table 5). Both examiners identified the single complete LT ligament tear only on the long video. Examiner 1 identified seven, whereas examiner 2 identified two of 15 partial LT ligament lesions on long videos. Only one partial lesion was detected by examiner 1 on a short video, while examiner 2 saw none.

Palmar radiocarpal ligaments

Examiner 1 rated 35 of the short and 12 of the long video sequences as inadequate for the assessment of the palmar radiocarpal ligaments. In these cases, they criticized the moment when the ligaments were displayed for being too short to adequately assess their integrity. Examiner 2 rated only one short and none of the long video sequences as
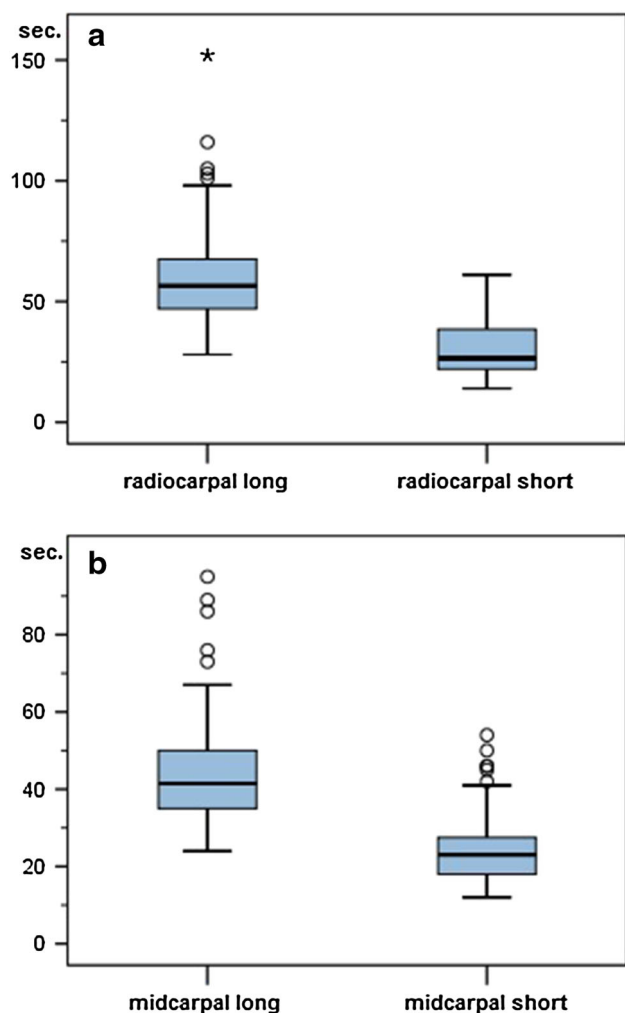
Fig. 1 Video lengths of the radiocarpal (**a**) and midcarpal (**b**) joints in wrist arthroscopy. The median length of the videos corresponds to the documentation of "simple" wrists. For such wrists, the long video sequences lasted 56.50 s for the radiocarpal and 41.50 s for the midcarpal joint. "Difficult" wrists necessitated correspondingly longer video documents. The short video sequences in this study lasted 26.5 s for the radiocarpal and 23.0 s for the midcarpal compartment

inadequate. In these cases, the examiners felt that the moment when the ligaments were displayed was too short to adequately assess the ligaments.

TFCC

Interpretation of the trampoline sign by examiner 2 revealed a higher interobserver reliability on long videos compared to short videos. According to Palmer's classification, no correlation between the length of the video and interobserver reliability was found. Examiner 1 detected one of 12 ulnar TFCC lesions on long videos. No lesions were identified on short videos. Examiner 2 detected TFCC lesions in three long and three short videos.

Synovitis

Neither Kappa coefficients nor analysis of the contingency tables outlined the specific influence of the video length on the inter-rater agreement for assessment of synovitis.

Discussion

According to Kappa statistics, long videos have no advantage over short videos for the assessment of the most important intra-articular structures and/or lesions—cartilage, SL ligament, and TFCC—in wrist arthroscopy. Nevertheless, both examiners rated twice as many presumptive cartilage lesions in the short videos as they observed in the long videos. This illustrates the need for a slow-motion video to adequately display articular surfaces. Among the studies that examine interobserver reliability for the assessment of cartilage, the authors of only one study mentioned the length of the video sequences [12]. Their videos for knee arthroscopies lasted one min each for the patellofemoral, medial, and lateral compartments. Despite these relatively long sequences, Kappa coefficients between 0.43 and 0.49 revealed only moderate agreement. Cameron et al. [13] videotaped the arthroscopies of six cadaveric knees. They calculated the percentage of agreement between the grades determined during arthroscopy and at subsequent arthrotomy. They found the Outerbridge classification moderately accurate for grading chondral lesions arthroscopically. Marx et al. [14] examined the interobserver reliability of the grading of cartilage lesions among surgeons at different institutions. They concluded that the arthroscopic grading of articular cartilage lesions would be reproducible, although their calculated Kappa coefficients ranged between 0.34 and 0.87. None of these studies considered the video length to be relevant in the assessment of cartilage lesions. In a prospective study, Javed et al. [15] found an interobserver variation of 18 % for the evaluation of articular surface defects. They stated that the examiners' level of experience would influence the accuracy of cartilage assessment. In this study, the surgeon and the two reviewers were highly and equally experienced in wrist arthroscopy.

As it is, we have to be aware that the assessment of articular surfaces by viewing video documents differs among surgeons [12–14]. Spahn et al. [16] have shown that poor interobserver reliability can be expected if different surgeons examine the same joint arthroscopically. Consequently, it seems necessary to improve the quality of video documentations in wrist arthroscopy. Determining a minimum length for such videos as in this study is an important part of quality improvement.

In this study, the video lengths differed within a large range (Fig. 1a, b). This seems to be due to technical issues

Table 2 Kappa coefficients for cartilage assessment of examiner 1

	Cartilage assessment according to outerbridge		Cartilage lesion present/absent	
	Short videos	Long videos	Short videos	Long videos
Wrist in total	–	–	0.465	0.524
Radiocarpal				
Scaphoid Fossa	0.317	0.268	0.393	0.393
Scaphoid	0.217	0.386	0.249	0.596
Lunate Fossa	0.112	0.267	0.247	0.387
Lunate	0.357	0.412	0.397	0.481
Midcarpal				
Scaphoid	0.010	0.000	−0.015	0.000
Lunate	0.010	0.000	−0.015	0.000
Capitate	0.349	0.433	0.530	0.682
Hamate	0.154	0.208	0.196	0.347
Triquetrum	0.012	0.010	−0.016	−0.016

Table 3 Kappa coefficients for cartilage assessment of examiner 2

	Cartilage assessment according to outerbridge		Cartilage lesion present/absent	
	Short videos	Long videos	Short videos	Long videos
Wrist in total	–	–	0.639	0.700
Radiocarpal				
Scaphoid Fossa	0.470	0.439	0.615	0.658
Scaphoid	0.427	0.368	0.584	0.470
Lunate Fossa	0.515	0.205	0.758	0.308
Lunate	0.531	0.600	0.574	0.653
Midcarpal				
Scaphoid	0.033	0.099	−0.048	0.221
Lunate	0.359	0.201	0.481	0.421
Capitate	0.270	0.467	0.393	0.576
Hamate	0.657	0.560	0.699	0.714
Triquetrum	0.279	0.483	0.267	0.580

based on the patients' anatomical structures. Synovitis and fibrosis may cause difficulties continuously examining the joint. Moreover, it is necessary, that the videos not stop at a site where the view is limited. This has been criticized in an earlier study [7]. All the more, it is necessary to display unclear parts of a joint, as these parts usually depict relevant articular pathologies. Therefore, in this study, single nonstop videos of each joint compartment were provided.

Regarding the SL ligament, complete lesions are not necessarily easy to diagnose when viewing a video document. The moment when the optic is inserted in between the scaphoid and the lunate should last appropriately long to allow an independent reviewer to recognize the capitate glancing through the SL gap. In this study, this resulted in obscured SL ligament lesions when the reviewers assessed the short videos. Kappa statistics were not able to depict this issue. Nevertheless, it makes sense not to pass difficult parts of the joint too quickly during video recording.

Limitations

Some bias may be assumed by the fact that the surgeon who performed the arthroscopies could have influenced the quality of the short video sequences, as he knew the study's aim. Another limitation of this study is the small number of actual intra-articular lesions, which resulted in low Kappa coefficients.

Conclusions

Despite the lack of statistical significance, adequate video documentation of our findings in wrist arthroscopies seems to necessitate adequate length of the video sequences. To avoid false-positive cartilage lesion diagnoses and to facilitate the detection of relevant ligament lesions, the video documents should last appropriately long. Assuming

Table 4 Percentage rates of false-positive findings and 95 % confidence intervals for cartilage assessment

	Examiner 1		Examiner 2	
	Short videos	Long videos	Short videos	Long videos
Wrist in total	4.1	2.04	22.45	13.73
	0.50, 13.98	0.05, 10.85	11.77, 36.62	4.63, 24.77
Radiocarpal				
Scaphoid Fossa	3.5	3.5	4.7	1.2
	0.73–9.86	0.73–9.86	1.28, 11.48	0.03, 6.31
Scaphoid	4.5	1.1	10.2	8
	1.25, 11.23	0.03, 6.17	4.78, 18.53	3.26, 15.70
Lunate Fossa	0	0	3.7	3.7
	0.00, 4.40	0.00, 4.40	0.76, 10.32	0.76, 10.32
Lunate	0	0	13.5	12.2
	0.00, 4.86	0.00, 4.86	6.68, 23.45	5.71, 21.84
Midcarpal				
Scaphoid	1	0	10.3	4.1
	0.03, 5.61	0.00, 3.73	5.06, 18.14	1.13, 10.22
Lunate	1	0	3.1	4.1
	0.03, 5.61	0.00, 3.73	0.64, 8.77	1.13, 10.22
Capitate	4.7	1.2	20	9.4
	1.30, 11.61	0.03, 6.38	12.10, 30.08	4.15, 17.71
Hamate	4.3	4.3	11.4	8.6
	0.89, 12.02	0.89, 12.02	5.07, 21.28	3.21, 17.73
Triquetrum	1	1	7.3	3.1
	0.03, 5.67	0.03, 5.67	2.98, 14.45	0.65, 8.86

Table 5 Kappa coefficients for assessment of ligaments, TFCC, and synovitis

	Examiner 1		Examiner 2	
	Short videos	Long videos	Short videos	Long videos
Ligaments				
SL partial/complete tear	0.436	0.573	0.834	0.763
SL findings from midcarpal	0.574	0.692	0.562	0.638
LT partial/complete tear	−0.009	0.463	−0.019	0.286
Palmar radiocarpal partial/complete tear	−0.019	0.109	0.271	0.593
TFCC				
Trampoline sign [10]	0.342	0.317	0.229	0.457
Palmer's classification [11]	0.26	0.297	0.249	0.226
Synovitis present/absent				
Radiocarpal	0.479	0.423	0.522	0.326
Ulnocarpal	0.253	0.375	0.189	0.352
Midcarpal	0.412	0.456	0.31	0.216

TFCC triangular fibrocartilage complex, *SL* Scapholunate ligament, *LT* lunotriquetral ligament

that the median length of the videos in this study adequately displays the findings in a simple wrist, we recommend that a sequence of the radiocarpal joint should last about 60 s and that the sequence of a midcarpal joint should last about 45 s. Videos of difficult joints should last appropriately longer.

References

1. Löw S, Herold A, Eingartner C (2014) Standard wrist arthroscopy: technique and documentation. Oper Orthop Traumatol 26:539–546. doi:10.1007/s00064-014-0311-6
2. Unglaub F, Manz S, Bruckner T, Leclère FM, Hahn P, Wolf MB (2013) Dorsal capsular imbrication for dorsal instability of the

distal radioulnar joint. Oper Orthop Traumatol 25:609–614. doi:10.1007/s00064-012-0223-2

3. Kasapinova K, Kamiloski V (2015) Influence of associated lesions of the intrinsic ligaments on distal radius fractures outcome. Arch Orthop Trauma Surg 135:831–838. doi:10.1007/s00402-015-2203-0

4. Kirchberger MC, Unglaub F, Mühldorfer-Fodor M, Pillukat T, Hahn P, Müller LP, Spies CK (2015) Update TFCC: histology and pathology, classification, examination and diagnostics. Arch Orthop Trauma Surg 135:427–437. doi:10.1007/s00402-015-2153-6

5. Löw S, Prommersberger KJ, Pillukat T, van Schoonhoven J (2010) Intra- and interobserver reliability of digitally photodocumented findings in wrist arthroscopy. Handchir Mikrochir Plast Chir 42:287–292. doi:10.1055/s-0030-1252065

6. Löw S, Herold D, Mühldorfer-Fodor M, Pillukat T (2012) The effect of labeling photo documents in wrist arthroscopies on intra- and interobserver reliability. Arch Orthop Trauma Surg 132:1813–1818. doi:10.1007/s00402-012-1612-6

7. Löw S, Pillukat T, Prommersberger KJ, van Schoonhoven J (2013) The effect of additional video documentation to photo documentation in wrist arthroscopies on intra- and interobserver reliability. Arch Orthop Trauma Surg 133:433–438. doi:10.1007/s00402-012-1670-9

8. Outerbridge RE (1961) The etiology of chondromalacia patellae. J Bone Joint Surg Br 43:752–757

9. Geissler WB, Freeland AE, Savoie FH, McIntyre LW, Whipple TL (1996) Intracarpal soft-tissue lesions associated with an intraarticular fracture of the distal end of the radius. J Bone Joint Surg Am 78:357–365

10. Hermansdorfer JD, Kleinman WB (1991) Management of chronic peripheral tears of the triangular fibrocartilage complex. J Hand Surg Am 16:340–346

11. Palmer AK (1989) Triangular fibrocartilage complex lesions: a classification. J Hand Surg Am 14:594–606

12. Brismar BH, Wredmark T, Movin T, Leandersson J, Svensson O (2002) Observer reliability in the arthroscopic classification of osteoarthritis of the knee. J Bone Joint Surg Br 84:42–47

13. Cameron ML, Briggs KK, Steadman JR (2003) Reproducibility and reliability of the Outerbridge classification for grading chondral lesions of the knee arthroscopically. Am J Sports Med 31:83–86

14. Marx RG, Connor J, Lyman S, Amendola A, Andrish JT, Kaeding C, McCarty EC, Parker RD, Wright RW, Spindler KP, Network Multicenter Orthopaedic Outcomes (2005) Multirater agreement of arthroscopic grading of knee articular cartilage. Am J Sports Med 33:1654–1657

15. Javed A, Siddique M, Vaghela M, Hui AC (2002) Interobserver variations in intra-articular evaluation during arthroscopy of the knee. J Bone Joint Surg Br 84:48–49

16. Spahn G, Klinger HM, Baums M, Pinkepank U, Hofmann GO (2011) Reliability in arthroscopic grading of cartilage lesions: results of a prospective blinded study for evaluation of interobserver reliability. Arch Orthop Trauma Surg 131:377–381. doi:10.1007/s00402-011-1259-8