


6-week radiographs unsuitable for diagnosis of suspected scaphoid fractures

Wouter H. Mallee¹  · Jos J. Mellema² · Thierry G. Guitton³ · J. Carel Goslings⁴ · David Ring⁵ · Job N. Doornberg¹ · Science of Variation Group

Received: 13 December 2015 / Published online: 30 March 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Introduction Six week follow-up radiographs are a common reference standard for the diagnosis of suspected scaphoid fractures. The main purpose of this study was to evaluate the interobserver reliability and diagnostic performance characteristics of 6-weeks radiographs for the detection of scaphoid fractures. In addition, two online techniques for evaluating radiographs were compared.

Materials and methods A total of 81 orthopedic surgeons affiliated with the Science of Variation Group assessed initial and 6-week scaphoid-specific radiographs of a consecutive series of 34 patients with suspected scaphoid fractures. They were randomized in two groups for evaluation, one used a standard website showing JPEG files and one a more sophisticated image viewer (DICOM). The goal

was to identify the presence or absence of a (consolidated) scaphoid fracture. Interobserver reliability was calculated using the multirater kappa measure. Diagnostic performance characteristics were calculated according to standard formulas with CT and MRI upon presentation in the emergency department as reference standards.

Results The interobserver agreement of 6-week radiographs for the diagnosis of scaphoid fractures was slight for both JPEG and DICOM ($k = 0.15$ and $k = 0.14$, respectively). The sensitivity (range 42–79 %) and negative predictive value (range 79–94 %) were significantly higher using a DICOM viewer compared to JPEG images. There were no differences in specificity (range 53–59 %), accuracy (range 53–58 %), and positive predictive value (range 14–26 %) between the groups.

Conclusions Due to low agreement between observers for the recognition of scaphoid fractures and poor diagnostic performance, 6-week radiographs are not adequate for evaluating suspected scaphoid fractures. The online evaluation of radiographs using a DICOM viewer seem to improve diagnostic performance characteristics compared to static JPEG images and future reliability and diagnostic studies should account for variation due to the method of delivering medical images.

Level of evidence Diagnostic level II.

Keywords Diagnostics · Fracture · Radiographs · Reference standard · Scaphoid · Occult

W. H. Mallee and J. J. Mellema shared first authorship based on equal contribution.

Members of the Science of Variation Group are given in the Acknowledgments.

✉ Wouter H. Mallee
w.h.mallee@amc.uva.nl

¹ Department of Orthopedic Surgery, Academic Medical Center Amsterdam, AMC, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

² Hand and Upper Extremity Service, Massachusetts General Hospital, Boston, MA, USA

³ Department of Plastic Surgery, University Medical Center Groningen, Groningen, The Netherlands

⁴ Department of Trauma Surgery, Academic Medical Center Amsterdam, Amsterdam, The Netherlands

⁵ Department of Surgery and Perioperative Care, Dell Medical School, Austin, USA

Introduction

In management of suspected scaphoid fractures, overtreatment (i.e. immobilization and restrictions of activities) must be balanced against the risks of nonunion associated with

undertreatment [1]. Overtreatment can be limited by establishing early definitive diagnosis using bone scintigraphy [2–4], magnetic resonance imaging (MRI) [5–8] and computed tomography (CT) [5, 8, 9]. However, there is no consensus on scaphoid imaging protocols due to limited evidence regarding diagnostic performance of these advanced imaging techniques [10].

The absence of a consensus reference standard for the diagnosis of scaphoid fractures makes the interpretation of diagnostic performance characteristics and improvement of diagnostic imaging tests difficult [11]. Latent class analysis can be used to estimate diagnostic test accuracy without using a reference standard [1, 12], but this approach has considerable limitations and must be viewed with skepticism [13]. The most commonly used reference standard in studies that evaluated diagnostic tests for scaphoid fractures are scaphoid-specific radiographs made 6 weeks after initial injury [5, 8, 9, 11, 14–18], while some authors question the use of follow-up radiographs as reference standard [19–21].

The Science of Variation Group, a collaborative effort to improve the study of variation in interpretation and classification of injuries, performed numerous studies by evaluating images using JPEG format [22–24]. Since this could limit diagnostic performance due to lack of several functions (window level, zoom, lower quality image), a new online tool was created using an embedded DICOM viewer. This tool mimics clinical practice, however, larger data files and use of multiple functions increases duration of assessment. It is unknown if this tool could be of true value.

As the reliability and accuracy of 6-week radiographs for suspected scaphoid fractures remain subject of discussion and important for the interpretation of diagnostic accuracy of alternative imaging modalities, CT and MRI in particular, there is a need to assess its reliability as well as diagnostic performance characteristics. Therefore, the purpose of this study was to evaluate the interobserver reliability and diagnostic performance characteristics of 6-week radiographs for the recognition of scaphoid fractures in patients with suspected scaphoid fractures. In addition, this study compared the online evaluation of radiographs in JPEG and DICOM format.

Methods

Study design

Orthopaedic surgeons affiliated with the Science of Variation Group were asked to log on to <http://www.scienceofvariationgroup.org> or <http://www.traumaplatform.org> for an online evaluation of suspected scaphoid fractures. In

an invitation email observers were informed that participation would be credited on the study by acknowledgement or group authorship [25, 26] and links were provided that directed to the respective web-based study platforms. Our Institutional Review Board approved this study.

Subjects

The initial and 6-week radiographs were used from our previous study [5] of a consecutive series of 34 patients aged 18 years or greater with a suspected scaphoid fracture (tenderness of the scaphoid and normal radiographic findings after a fall on the outstretched hand). All patients presented within 24 h after injury and underwent CT and MRI within 10 days after wrist injury between April 2008 and October 2008 in a level I trauma center.

The number of subjects in reliability studies is determined based on an appropriate balance between the number of observers evaluating each subject and the number of subjects [27]. Our web-based study platforms (i.e. Science of Variation Group and Traumaplatform) aim to increase the number of observers in interobserver reliability studies for maximizing power and generalizability and to allow comparison between and within subgroups. For this reason, we prefer to select a limited number of subjects to limit burden on observers and increase participation rate (i.e. number of observers).

Observers

Orthopedic surgeons trained in hand surgery and listed in the Science of Variation Group as active members were randomized (1:1) by computer-generated random numbers (Microsoft Excel, Redmond, WA, USA) to assess the selected radiographs online in JPEG or DICOM format.

Online evaluation

Scaphoid-specific radiographs at baseline and 6 weeks after initial trauma were presented and consisted of four views: (1) a posteroanterior view with the wrist in ulnar deviation, (2) a lateral view with the wrist in 15° extension, (3) a lateral view with the wrist in 30° of pronation, and (4) a posteroanterior view with the X-ray beam directed from distal to proximal and with the wrist positioned in 40° of angulation. Observers were asked to answer 1 question for each of the 34 cases: Is there a (consolidated) scaphoid fracture?

Before starting the online evaluation and upon log on to the website, observers received a short description of the study procedure. Observers assigned to the JPEG group evaluated radiographs that were converted to images in JPEG format (<http://www.scienceofvariationgroup.org>).

org) and observers assigned to the DICOM group evaluated radiographs provided by an online DICOM viewer (<http://www.traumaplatform.org>). Both groups evaluated the same initial and 6-week radiographs, however, the JPEG group was not able to use the window level, scroll, and zoom options available in the online DICOM viewer software.

Statistical analysis

A post hoc power analysis was performed using the method as described by Guitton and Ring [23]. It was calculated that 81 observers provided 5.8 % power to detect a 0.003 difference in kappa value (i.e. interobserver reliability) between the JPEG and DICOM group using a two-sample z test ($\alpha = 0.05$). However, 81 observers provided 100 % power to detect a clinically relevant difference in kappa value, defined as a difference of one category as describe by Landis and Koch [28] ($\Delta_{\text{kappa}} = 0.20$), between the groups with $\alpha = 0.05$.

Interobserver reliability was calculated using the multirater kappa as described by Siegel and Castellan [29]. The kappa statistic is a frequently used measure of chance-corrected agreement between observers and interpreted according to the guidelines of Landis and Koch [28]: a value of 0.01 to 0.20 indicates slight agreement; 0.21 to 0.40, fair agreement; 0.41 to 0.60, moderate agreement; 0.61 to 0.80, substantial agreement; and 0.81 to 0.99, almost perfect agreement. A two-sample z test was used to compare kappa values and P values of <0.05 were considered significant. For a better understanding of the underlying data, the proportion of agreement was calculated for each case (in absolute percentages, %) and defined as the proportion of observers agreeing with the most provided answer.

Diagnostic performance characteristics (sensitivity, specificity, accuracy, positive predictive value, and negative predictive values) of 6-week radiographs for the recognition of (consolidated) scaphoid fractures were calculated according to standard formulas. The reference standard for the diagnosis of scaphoid fractures was CT and MRI. A panel of three observers, an attending musculoskeletal radiologist, an attending trauma surgeon who treats fractures, and an attending orthopaedic surgeon, evaluated the images for the presence of a scaphoid fracture until a consensus opinion was reached [5]. The 95 % confidence intervals (95 % CIs) were calculated using the formula for the standard error of proportion, based on normal approximation method for binomial proportions, and differences were considered significant when the 95 % CIs did not overlap [30].

Results

Observer characteristics

A total of 288 invitation emails were sent, of which 143 went to the JPEG group and 145 to the DICOM group. Fifty-seven respondents started with the evaluation in the JPEG group, of which 53 (93 %) completed the online evaluation, and 45 respondents started in the DICOM group, of which 28 (62 %) completed the online evaluation. After incomplete responses were excluded, 53 (65 %) observers were left in the JPEG group and 28 (35 %) in the DICOM group. Observers were predominately male (95 %), from the US (78 %), hand and wrist surgeons (96 %), and in independent practice for more than 5 years (68 %) (Table 1).

Reliability of 6-week radiographs for scaphoid fractures

The interobserver reliability of 6-week radiographs for the diagnosis of scaphoid fractures was the same for the JPEG and DICOM viewer group and slight in both groups ($k = 0.15$ and $k = 0.14$, respectively; $P = 0.75$). In addi-

Table 1 Observer characteristics

	JPEG ($n = 53$)		DICOM viewer ($n = 28$)	
	n	%	n	%
Sex				
Men	50	94	27	96
Women	3	5.7	1	3.6
Area				
United States	41	77	22	79
Europe	7	13	3	11
Other	5	9.4	3	11
Specialization				
Hand and wrist	53	100	25	89
Shoulder and elbow	–	–	2	7.1
Trauma	–	–	1	3.6
Years in independent practice				
0–5	18	34	8	29
6–10	9	17	4	14
11–20	16	30	10	36
21–30	10	19	6	21
Fractures per year				
0–10	12	23	5	18
11–20	33	62	7	25
More than 20	8	15	16	57

Table 2 Interobserver agreement for the recognition of (consolidated) scaphoid fractures based on 6-week radiographs (JPEG versus DICOM viewer)

	JPEG (<i>n</i> = 53)			DICOM viewer (<i>n</i> = 28)			<i>P</i> value
	Kappa	Agreement	95 % CI	Kappa	Agreement	95 % CI	
Overall	0.15	Slight	0.13 to 0.16	0.14	Slight	0.12 to 0.16	0.75
Area							
United States	0.14	Slight	0.12 to 0.16	0.16	Slight	0.14 to 0.18	0.24
Europe	0.18	Slight	0.09 to 0.26	0.28	Fair	0.06 to 0.50	0.40
Other	0.04	Slight	−0.06 to 0.15	0.18	Slight	−0.04 to 0.41	0.28
Years in independent practice							
0–5	0.12	Slight	0.09 to 0.15	0.06	Slight	0.00 to 0.13	0.094
More than 5 years	0.16	Slight	0.13 to 0.18	0.16	Slight	0.13 to 0.18	0.94
Fractures per year							
0–10	0.17	Slight	0.11 to 0.23	0.22	Fair	0.09 to 0.35	0.47
11–20	0.14	Slight	0.12 to 0.15	0.21	Fair	0.12 to 0.29	0.12
More than 20	0.12	Slight	0.03 to 0.22	0.13	Slight	0.10 to 0.16	0.94

tion, subgroup analysis showed that interobserver agreement ranged from slight to fair and no significant differences in kappa value between subgroups were detected (Table 2). The average proportion of agreement was 68 % in the JPEG group and 68 % in the DICOM group (Table 3).

Diagnostic performance characteristics of 6-week radiographs for scaphoid fractures

The sensitivity of 6-week radiographs for the diagnosis of scaphoid fractures ranged from 42 to 79 % and was significantly higher in the DICOM group compared to the JPEG group with MRI, CT, and MRI with CT combined as reference standard. Specificity ranged from 53 to 59 %, accuracy ranged from 53 to 58 %, and positive predictive value ranged from 14 to 26 % and were not significantly different between the DICOM and JPEG group with MRI, CT and MRI with CT combined as reference standard. The negative predictive value ranged from 79 to 94 % and was significantly higher using the DICOM viewer compared to JPEG images with MRI, CT, and MRI with CT combined as reference standard (Table 4).

Discussion

Scaphoid-specific radiographs at 6 weeks follow-up are most commonly used as reference standard for scaphoid fractures despite its alternatives, such as latent class analysis and MRI, but its use remains subject of discussion [1, 5, 7–9, 12, 14–18]. This study was designed to evaluate the interobserver reliability and diagnostic performance characteristics of 6-week radiographs for the recognition of

scaphoid fractures in patients with suspected scaphoid fractures and to compare the online evaluation of radiographs using images in JPEG and DICOM format. We found that the interobserver reliability for 6-week radiographs was slight in both the JPEG and DICOM group. The diagnostic performance characteristics of 6-week radiographs were poor as well, but significantly better when radiographs were evaluated using a DICOM viewer compared to JPEG images.

The strengths of our study include the large number of observers, which allowed a more complex study design with randomization and subgroup analysis, the use of prospectively collected data from our previous study [5] that evaluated a consecutive series of 34 patients with a suspected scaphoid fracture that returned for follow-up after 6 weeks and underwent CT and MRI scans, and the use of DICOM viewers for the online evaluation of radiographs that resembles evaluation in clinical practice. The limitations include the heterogeneous group of surgeons that evaluated the radiographs, which were from multiple countries and different levels of experience and therefore more likely to disagree compared to observers from a single institute with the same level of experience. A possible limitation was the use of a reference standard for the diagnosis of scaphoid fractures that was based on CT and MRI findings and the consensus agreement of three senior authors.

In this study, the interobserver reliability for the recognition of scaphoid fractures based on 6-week radiographs was low in the JPEG and DICOM group and comparable with agreement reported in previous studies [19–21]. Tiel-van Buul et al. [19] selected follow-up radiographs (2 and 6 weeks after injury) of a consecutive series of 60 patients with suspected scaphoid fractures that

Table 3 Proportion of agreement for the recognition of (consolidated) scaphoid fractures based on 6-week radiographs (JPEG and DICOM viewer)

Case no.	JPEG (<i>n</i> = 53)		DICOM Viewer (<i>n</i> = 28)	
	Most provided answer	PA*	Most provided answer	PA*
1	Present	79	Absent	57
2	Absent	64	Present	86
3	Absent	66	Absent	57
4	Present	57	Absent	75
5	Absent	62	Absent	61
6	Absent	70	Absent	75
7	Absent	57	Present	82
8	Present	51	Absent	75
9	Present	85	Present	57
10	Present	68	Absent	57
11	Present	74	Absent	54
12	Present	72	Present	93
13	Absent	60	Absent	64
14	Absent	83	Absent	71
15	Absent	85	Absent	79
16	Present	62	Present	68
17	Absent	70	Present	86
18	Absent	77	Present	79
19	Absent	55	Present/absent	50
20	Present	53	Absent	61
21	Absent	77	Absent	64
22	Absent	87	Present	61
23	Absent	66	Present	57
24	Present	55	Present	61
25	Absent	74	Present	75
26	Present	70	Absent	64
27	Absent	87	Absent	57
28	Present	62	Absent	61
29	Absent	66	Absent	79
30	Absent	57	Present	75
31	Present	60	Absent	79
32	Absent	87	Absent	71
33	Absent	62	Absent	61
34	Absent	60	Present	61

* Proportion of agreement: the proportion of observers agreeing with the most provided answer

were rated by 4 observers and found slight to fair inter-observer agreement (range $k = 0.20$ to $k = 0.39$). A similar study by Tiel-van Buul et al. [20] reported slight to moderate agreement (range $k = 0.19$ to $k = 0.50$) among 3 observers that evaluated 6-week radiographs of a consecutive series of 78 patients with clinically suspected scaphoid fractures. Low et al. [21] found fair agreement (range $k = 0.30$ to $k = 0.40$) for scaphoid-specific follow-up radiographs between 4 observers that rated 50 patients with a suspected scaphoid fracture.

We found that the diagnostic performance characteristics of 6-week radiographs for scaphoid fractures were poor

with MRI, CT, and MRI with CT combined as reference standard using radiographs in JPEG and DICOM format. Six-week radiographs seem better at excluding scaphoid fractures (negative predictive value ranged from 79 to 94 %) than recognizing a scaphoid fracture (positive predictive value ranged from 14 to 26 %). Moreover, our data suggest that almost 50 % of the ratings were inaccurate (accuracy ranged from 53 to 58 %). Low et al. [21]. reported low negative predictive value (range 30 to 40 %) and high positive predictive value (range 75 to 88 %) of follow-up radiographs in patients with suspected scaphoid fractures with MRI as reference standard, which were not

Table 4 Diagnostic performance of 6-week radiographs for the recognition of (consolidated) scaphoid fractures (JPEG versus DICOM viewer)

	JPEG (<i>n</i> = 53)		DICOM viewer (<i>n</i> = 28)	
	%	95 % CI	%	95 % CI
Reference standard: MRI				
Sensitivity	42	37–47	64	57–71
Specificity	56	54–59	53	50–57
Accuracy	53	51–56	56	52–59
Positive predictive value	20	17–23	26	22–30
Negative predictive value	79	76–81	85	82–88
Reference standard: CT				
Sensitivity	56	50–62	79	72–85
Specificity	59	56–61	55	51–58
Accuracy	58	56–61	58	55–61
Positive predictive value	19	16–22	23	19–27
Negative predictive value	89	87–90	94	91–96
Reference standard: MRI + CT				
Sensitivity	52	45–59	75	67–83
Specificity	58	55–60	53	50–56
Accuracy	57	55–59	56	52–59
Positive predictive value	14	12–17	18	14–21
Negative predictive value	90	88–92	94	92–96

consistent with our findings. These differences can be explained as the prevalence influences the negative predictive value and positive predictive value [9, 31]. Our study evaluated the radiographs of a consecutive series of patients (prevalence 18 %) and Low et al. selected patients retrospectively if they had both follow-up radiographs and MRI after injury (prevalence 75 %).

Our results show that the method of presenting radiographs may affect their evaluation by surgeon observers. We found that the interobserver reliability was the same in the JPEG and DICOM group, but the diagnostic performance was better when radiographs were evaluated using a DICOM viewer compared to static JPEG images. The ability to window level, scroll, and zoom using a DICOM viewer improved the diagnosis of scaphoid fractures, in terms of sensitivity and negative predictive value, significantly. Since the format of medical images could be a source of variation between surgeons, it should be accounted for in future reliability and diagnostic studies.

Given the low agreement and poor diagnostic accuracy of 6-week radiographs for the recognition of scaphoid fractures in this study, surgeons and patients must accept that they are dealing with probabilities rather than certainties in the management of scaphoid fractures. For example, we cannot reduce the probability of missing a fracture to 0 % with a negative predictive value of less than 100 %. Using 6-week radiographs as reference standard for studying suspected scaphoid fractures is not advised for future studies. To date, observer experience, training,

image presentation, training, and simplification of classifications are shown to have a limited effect on the reliability and accuracy of diagnosis and classification of fractures. At this time it remains unclear what interventions will improve reliability and accuracy, but our collaborative plans to continue studying variation between surgeons to attempt to reduce it.

Acknowledgments Science of Variation Group: A. Ilyas, R. P. Calfee, C. H. Fernandes, C. O. Miranda, C. Taleb, C. L. Moreno-Serrano, M. A. Pirela-Cruz, D. A. Osei, D. M. Kalainov, D. Polatsch, E. T. Walbeehm, E. D. Schumer, F. Suarez, H. W. Grunwald, G. W. Balfour, M. Nancollas, H. L. Kimball, J. A. Greenberg, J. McAuliffe, J. Choueka, J. Wint, K. J. Prommersberger, K. M. Rumball, G. A. Kraan, L. Lattanza, L. Paz, L. Catalano, M. J. Richard, M. Soong, M. Darowish, R. M. Costanzo, M. M. Wood, M. Baskies, M. J. Palmer, M. W. Grafe, O. M. Semenkin, J. A. Ortiz, J. E. Forigua, P. Blazar, W. A. Batson, R. De Bedout, R. R. L. Gray, R. Papandrea, R. L. Hutchison, R. R. Slater, R. Klinefelter, S. Spruijt, S. A. Kennedy, S. Beldner, S. Dodds, S. Kaplan, T. Aparid, T. G. Stackhouse, P. A. Martineau, L. F. Naquira, G. M. Pess, T. Siff, C. M. Jones, J. Adams, J. M. Erickson, S. Kakar, B. P. D. Wills, F. T. D. Kaplan, P. W. Owens, M. W. Kessler, E. Hofmeister, K. D. Bickel, J. T. Capo, C. J. Walsh, W. C. Hammert, Camilo Jose Romero Barreto, T. Baxamusa, R. M. Fricker, N. Hoekzema, C. A. Goldfarb, R. S. Gilbert, J. Taras, M. Behrman, M. Calcagni, F. L. Walter, M. P. van den Bekerom.

Compliance with ethical standards

Conflict of interest Each author certifies that he has no commercial associations (e.g., consultancies, stock ownership, equity interest, patent/licensing arrangement) that might pose a conflict of interest in connection with the submitted article.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Duckworth AD, Ring D, McQueen MM (2011) Assessment of the suspected fracture of the scaphoid. *J Bone Joint Surg Br* 93(6):713–719. doi:[10.1302/0301-620x.93b6.26506](https://doi.org/10.1302/0301-620x.93b6.26506)
- Rhemrev SJ, de Zwart AD, Kingma LM, Meylaerts SA, Arndt JW, Schipper IB, Beeres FJ (2010) Early computed tomography compared with bone scintigraphy in suspected scaphoid fractures. *Clin Nucl Med* 35(12):931–934. doi:[10.1097/RLU.0b013e3181f9de26](https://doi.org/10.1097/RLU.0b013e3181f9de26)
- Akdemir UO, Atasever T, Sipahioglu S, Turkolmez S, Kazimoglu C, Sener E (2004) Value of bone scintigraphy in patients with carpal trauma. *Ann Nucl Med* 18(6):495–499
- Beeres FJ, Rhemrev SJ, den Hollander P, Kingma LM, Meylaerts SA, le Cessie S, Bartlema KA, Hamming JF, Hogervorst M (2008) Early magnetic resonance imaging compared with bone scintigraphy in suspected scaphoid fractures. *J Bone Joint Surg Br* 90(9):1205–1209. doi:[10.1302/0301-620x.90b9.20341](https://doi.org/10.1302/0301-620x.90b9.20341)
- Mallee W, Doornberg JN, Ring D, van Dijk CN, Maas M, Goslings JC (2011) Comparison of CT and MRI for diagnosis of suspected scaphoid fractures. *J Bone Joint Surg Am* 93(1):20–28. doi:[10.2106/JBJS.I.01523](https://doi.org/10.2106/JBJS.I.01523)
- De Zwart AD, Beeres FJ, Ring D, Kingma LM, Coerkamp EG, Meylaerts SA, Rhemrev SJ (2012) MRI as a reference standard for suspected scaphoid fractures. *Br J Radiol* 85(1016):1098–1101. doi:[10.1259/bjr/73145885](https://doi.org/10.1259/bjr/73145885)
- Tibrewal S, Jayakumar P, Vaidya S, Ang SC (2012) Role of MRI in the diagnosis and management of patients with clinical scaphoid fracture. *Int Orthop* 36(1):107–110. doi:[10.1007/s00264-011-1350-3](https://doi.org/10.1007/s00264-011-1350-3)
- Memarsadeghi M, Breitenseher MJ, Schaefer-Prokop C, Weber M, Aldrian S, Gabler C, Prokop M (2006) Occult scaphoid fractures: comparison of multidetector CT and MR imaging—initial experience. *Radiology* 240(1):169–176. doi:[10.1148/radiol.2401050412](https://doi.org/10.1148/radiol.2401050412)
- Adey L, Souer JS, Lozano-Calderon S, Palmer W, Lee SG, Ring D (2007) Computed tomography of suspected scaphoid fractures. *J Hand Surg Am* 32(1):61–66. doi:[10.1016/j.jhsa.2006.10.009](https://doi.org/10.1016/j.jhsa.2006.10.009)
- Groves AM, Kayani I, Syed R, Hutton BF, Bearcroft PP, Dixon AK, Ell PJ (2006) An international survey of hospital practice in the imaging of acute scaphoid trauma. *AJR Am J Roentgenol* 187(6):1453–1456. doi:[10.2214/ajr.05.0686](https://doi.org/10.2214/ajr.05.0686)
- Ring D, Lozano-Calderon S (2008) Imaging for suspected scaphoid fracture. *J Hand Surg Am* 33(6):954–957. doi:[10.1016/j.jhsa.2008.04.016](https://doi.org/10.1016/j.jhsa.2008.04.016)
- Buijze GA, Mallee WH, Beeres FJP, Hanson TE, Johnson WO, Ring D (2011) Diagnostic performance tests for suspected scaphoid fractures differ with conventional and latent class analysis. *Clin Orthop Relat Res* 469(12):3400–3407. doi:[10.1007/s11999-011-2074-9](https://doi.org/10.1007/s11999-011-2074-9)
- Pepe MS, Janes H (2007) Insights into latent class analysis of diagnostic test performance. *Biostatistics* 8(2):474–484. doi:[10.1093/biostatistics/kx1038](https://doi.org/10.1093/biostatistics/kx1038)
- Gabler C, Kukla C, Breitenseher MJ, Trattinig S, Vecsei V (2001) Diagnosis of occult scaphoid fractures and other wrist injuries. Are repeated clinical examinations and plain radiographs still state of the art? *Langenbeck's Arch Surg/Dtsch Ges Chir* 386(2):150–154
- Munk B, Frokjaer J, Larsen CF, Johannsen HG, Rasmussen LL, Edal A, Rasmussen LD (1995) Diagnosis of scaphoid fractures. A prospective multicenter study of 1052 patients with 160 fractures. *Acta Orthop Scand* 66(4):359–360
- Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, Imhof H, Trattinig S (1997) Radiographically occult scaphoid fractures: value of MR imaging in detection. *Radiology* 203(1):245–250. doi:[10.1148/radiology.203.1.9122402](https://doi.org/10.1148/radiology.203.1.9122402)
- Groves AM, Cheow HK, Balan KK, Courtney HM, Bearcroft PW, Dixon AK (2005) Case report: false negative 16 detector multislice CT for scaphoid fracture. *Br J Radiol* 78(925):57–59. doi:[10.1259/bjr/55015850](https://doi.org/10.1259/bjr/55015850)
- Mallee WH, Doornberg JN, Ring D, Maas M, Muhl M, van Dijk CN, Goslings JC (2014) Computed tomography for suspected scaphoid fractures: comparison of reformations in the plane of the wrist versus the long axis of the scaphoid. *Hand (N Y)* 9(1):117–121. doi:[10.1007/s11552-013-9556-z](https://doi.org/10.1007/s11552-013-9556-z)
- Tiel-van Buul MM, van Beek EJ, Broekhuizen AH, Nooitgedacht EA, Davids PH, Bakker AJ (1992) Diagnosing scaphoid fractures: radiographs cannot be used as a gold standard! *Injury* 23(2):77–79
- Tiel-van Buul MM, van Beek EJ, Borm JJ, Gubler FM, Broekhuizen AH, van Royen EA (1993) The value of radiographs and bone scintigraphy in suspected scaphoid fracture. A statistical analysis. *J Hand Surg (Edinburgh, Scotland)* 18(3):403–406
- Low G, Raby N (2005) Can follow-up radiography for acute scaphoid fracture still be considered a valid investigation? *Clin Radiol* 60(10):1106–1110. doi:[10.1016/j.crad.2005.07.001](https://doi.org/10.1016/j.crad.2005.07.001)
- Doornberg J, Lindenhovius A, Kloen P, van Dijk CN, Zurakowski D, Ring D (2006) Two and three-dimensional computed tomography for the classification and management of distal humeral fractures. Evaluation of reliability and diagnostic accuracy. *J Bone Joint Surg Am* 88(8):1795–1801
- Guitton TG, Ring D (2011) Interobserver reliability of radial head fracture classification: two-dimensional compared with three-dimensional CT. *J Bone Joint Surg Am* 93(21):2015–2021. doi:[10.2106/jbjs.j.00711](https://doi.org/10.2106/jbjs.j.00711)
- Neuhaus V, Bot AG, Guitton TG, Ring DC, Abdel-Ghany MI, Abrams J, Abzug JM, Adolfsson LE, Balfour GW, Bamberger HB, Barquet A, Baskies M, Batson WA, Baxamusa T, Bayne GJ, Begue T, Behrman M, Beingsner D, Biert J, Bishop J, Alves MB, Boyer M, Brilej D, Brink PR, Brunton LM, Buckley R, Cagnone JC, Calfee RP, Campinhos LA, Cassidy C, Catalano L 3rd, Chivers K, Choudhari P, Cimerman M, Conflitti JM, Costanzo RM, Crist BD, Cross BJ, Dantuluri P, Darowish M, de Bedout R, DeCoster T, Dennison DG, DeNoble PH, DeSilva G, Dienstknecht T, Duncan SF, Duralde XA, Durchholz H, Egol K, Ekholm C, Elias N, Erickson JM, Esparza JD, Fernandes CH, Fischer TJ, Fischmeister M, Forigua Jaime E, Getz CL, Gilbert RS, Giordano V, Glaser DL, Gosens T, Grafe MW, Filho JE, Gray RR, Gulotta LV, Gummerson NW, Hammerberg EM, Harvey E, Haverlag R, Henry PD, Hobby JL, Hofmeister EP, Hughes T, Itamura J, Jebson P, Jenkinson R, Jeray K, Jones CM, Jones J, Jubel A, Kaar SG, Kabir K, Kaplan FT, Kennedy SA, Kessler MW, Kimball HL, Kloen P, Klostermann C, Kohut G, Kraan GA, Kristan A, Loebenberg MI, Malone KJ, Marsh L, Martineau PA, McAuliffe J, McGraw I, Mehta S, Merchant M, Metzger C, Meylaerts SA, Miller AN, Wolf JM, Murachovsky J, Murthi A, Nancollas M, Nolan BM, Omara T, Omara T, Ortiz JA, Overbeck JP, Castillo AP, Pesantez R, Polatsch D, Porcellini G, Prayson M, Quell M, Ragsdell MM, Reid JG, Reuver JM, Richard MJ, Richardson M, Rizzo M, Rowinski S, Rubio J, Guerrero CG, Satora W, Schandelmaier P, Scheer JH, Schmidt A,

- Schubkegel TA, Schulte LM, Schumer ED, Sears BW, Shafritz AB, Shortt NL, Siff T, Silva DM, Smith RM, Spruijt S, Stein JA, Pemovska ES, Streubel PN, Swigart C, Swiontkowski M, Thomas G, Tolo ET, Turina M, Tyllianakis M, van den Bekerom MP, van der Heide H, van de Sande MA, van Eerten PV, Verbeek DO, Hoffmann DV, Vochteloo AJ, Wagenmakers R, Wall CJ, Wallensten R, Wascher DC, Weiss L, Wiater JM, Wills BP, Wint J, Wright T, Young JP, Zalavras C, Zura RD, Zyto K (2014) Scapula fractures: interobserver reliability of classification and treatment. *J Orthop Trauma* 28(3):124–129. doi:[10.1097/BOT.0b013e31829673e2](https://doi.org/10.1097/BOT.0b013e31829673e2)
25. Mellema JJ, Doornberg JN, Guitton TG, Ring D (2014) Biomechanical studies: science (f)or common sense? *Injury* 45(12):2035–2039. doi:[10.1016/j.injury.2014.09.014](https://doi.org/10.1016/j.injury.2014.09.014)
26. Doornberg JN, Guitton TG, Ring D (2013) Diagnosis of elbow fracture patterns on radiographs: interobserver reliability and diagnostic accuracy. *Clin Orthop Relat Res* 471(4):1373–1378. doi:[10.1007/s11999-012-2742-4](https://doi.org/10.1007/s11999-012-2742-4)
27. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. *Stat Med* 17(1):101–110
28. Landis JR, Koch GG (1977) An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 33(2):363–374
29. Siegel S, Castellan JN (1988) *Nonparametric statistics for the behavioral sciences*. McGraw-Hill, New York
30. Harper R, Reeves B (1999) Reporting of precision of estimates for diagnostic accuracy: a review. *BMJ* 318(7194):1322–1323
31. Altman DG, Bland JM (1994) Diagnostic tests 2: predictive values. *BMJ* 309(6947):102