Clin Orthop Relat Res (2010) 468:723–734
DOI 10.1007/s11999-009-1081-6

CLINICAL RESEARCH

# Diagnosing Suspected Scaphoid Fractures

## A Systematic Review and Meta-analysis

**Zhong-Gang Yin MD, Jian-Bing Zhang MD, Shi-Lian Kan MD, Xiao-Gang Wang MD**

Received: 4 February 2009 / Accepted: 26 August 2009 / Published online: 15 September 2009


**Abstract** Imaging protocols for suspected scaphoid fractures among investigators and hospitals are markedly inconsistent. We performed a systematic review and meta-analysis to assess and compare the diagnostic performance of bone scintigraphy, MRI, and CT for diagnosing suspected scaphoid fractures. Twenty-six studies were included. Sensitivity, specificity, and diagnostic odds ratio were pooled separately and summary receiver operating characteristic curves were fitted for each modality. Meta-regression analyses were performed to compare these modalities. We obtained likelihood ratios derived from the pooled sensitivity and specificity and, using Bayes' theorem, calculated the posttest probability by application of the tests. The pooled sensitivity, specificity, natural logarithm of the diagnostic odds ratio, and the positive and negative likelihood ratios were, respectively, 97%, 89%, 4.78, 8.82, and 0.03 for bone scintigraphy; 96%, 99%, 6.60, 96, and 0.04 for MRI; and 93%, 99%, 6.11, 93, and 0.07 for CT. Bone scintigraphy and MRI have equally high sensitivity and high diagnostic value for excluding scaphoid fracture; however, MRI is more specific and better for confirming scaphoid fracture. We believe additional studies are needed to assess diagnostic performance of CT, especially paired design studies or randomized controlled trials to compare CT with MRI or bone scintigraphy.

**Level of Evidence:** Level III, diagnostic study. See the Guidelines for Authors for a complete description of levels of evidence.

## Introduction

Scaphoid fractures are the most common type of carpal fractures and occur frequently in young men [40]. Because their potential complications, including nonunion, avascular necrosis, and osteoarthritis, are made more likely by a delay in diagnosis and treatment, early diagnosis and treatment for these fractures are critical to improving outcomes [3]. The diagnosis of a scaphoid fracture usually can be established on the basis of clinical examination and radiographs, which typically include four views: posteroanterior, lateral, semipronated oblique, and posteroanterior with ulnar deviation [3]. However, in the acute phase after injury, some fractures are radiographically occult. To avoid undertreatment of these occult fractures, patients with suspected scaphoid fractures (high clinical probability of a scaphoid fracture but negative or equivocal radiographs) usually are treated with cast immobilization followed by repeat clinical examination and radiographs [46]. As the prevalence of true fractures among patients with suspected scaphoid fractures might be only 5% to 10% [1], the majority of these patients are overtreated, which results in lost work days and productivity and increased healthcare costs [11].

To avoid undertreatment and overtreatment, accurate and early diagnosis is required to confirm and exclude scaphoid fracture as the diagnosis. Investigators have recommended various imaging modalities to achieve earlier definitive diagnosis for suspected scaphoid fractures, including bone scintigraphy [13, 49], MRI [18, 42], CT

Each author certifies that he or she has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

Z.-G. Yin (✉), J.-B. Zhang, S.-L. Kan, X.-G. Wang
Department of Hand and Microsurgery, Tianjin Hospital, 300211 Tianjin, China
e-mail: zgyin@hotmail.com

[41], and high-frequency sonography [20]. An international survey of hospital practices revealed marked inconsistency in acute scaphoid fracture imaging protocols, which the authors believed were likely to be multifactorial but also probably reflected a deficiency in scientific evidence regarding the best practice for imaging scaphoid fractures [24]. Clinical decisions regarding the use and interpretation of a diagnostic test require assessment of diagnostic performance. Ideally, the assessment is based on the aggregate of pertinent knowledge available rather than on single studies alone or on personal experience [16].

The purposes of this meta-analysis were to (1) obtain and compare summary estimates of sensitivity, specificity, and diagnostic odds ratio of bone scintigraphy, MRI, and CT for diagnosing suspected scaphoid fractures; (2) use summary receiver operating characteristic (SROC) curves to estimate and compare the overall diagnostic performance of the three modalities; and (3) obtain the likelihood ratios derived from pooled sensitivity and specificity and, using Bayes' theorem, calculate the posttest probability of scaphoid fractures by application of these imaging modalities.

## Materials and Methods

We searched PubMed (January 1966 to October 2008) with the following search strategy: (scaphoid bone[MeSH] OR carpal bone[MeSH] AND fracture[MeSH]) AND (predictive value[WORD] OR test[WORD] OR accuracy[WORD] OR sensitiv*[WORD] OR specificity[WORD] OR sensitivity and specificity[MeSH] OR diagnos*[Title/Abstract] OR diagnosis[MeSH] OR diagnosis[SH] OR false negative[WORD] OR false positive[WORD] OR detection [WORD]); we searched EMBASE (OVID, January 1980 to October 2008) with the keyword scaphoid fracture; and we hand-searched the references of relevant studies without date limitation. There was no language restriction on the search. To be included, the study had to meet the following criteria. (1) The study was a clinical investigation that assessed the diagnostic performance of bone scintigraphy, MRI, CT, or ultrasound for suspected scaphoid fractures. (2) The study used followup images (radiographs, CT, MRI, or bone scintigraphy) or clinical followup and/or combined images as the reference test. (3) The study provided sufficient information to reconstruct a $2 \times 2$ contingency table of the performance of the index test. (4) The study was published as a full report in English. We tried to determine if different studies from the same institution used the same patients because one author published several reports. When data were presented in more than one article, the article with the most details or the most recent article was chosen. Two reviewers (ZGY, JBZ) independently selected the studies, and disagreements were discussed to reach a consensus.

The computer search yielded 2440 citations: 1637 from PubMed and 803 from EMBASE, of which 26 studies ultimately were included in this review [2, 4, 6, 7, 9, 10, 14, 19, 21, 23, 25, 27, 30, 32, 34, 37, 39, 43, 44, 47, 48, 50–52, 54, 55] (Fig. 1).

Two reviewers (ZGY, JBZ) independently extracted the following data from each study: year of publication, patient demographics, sample size, imaging technique, reference test, $2 \times 2$ table of the index test, and prevalence of scaphoid fracture; and used QUADAS criteria, which consisted of 14 questions answered "yes," "no," or "unclear," to assess the methodologic quality of the studies (Table 1) [53]. The item "was the period between reference standard and index test short enough" was omitted because image and clinical followup were considered the reference standards. Disagreements were resolved by discussion.

Two imaging modalities were reported in seven studies, so 15 studies with 1102 participants reported on the diagnostic accuracy of bone scintigraphy, 10 studies with 513 participants reported on MRI, six studies with 211 participants reported on CT, and two studies with 72 participants reported on ultrasound (Tables 2, 3). The results for

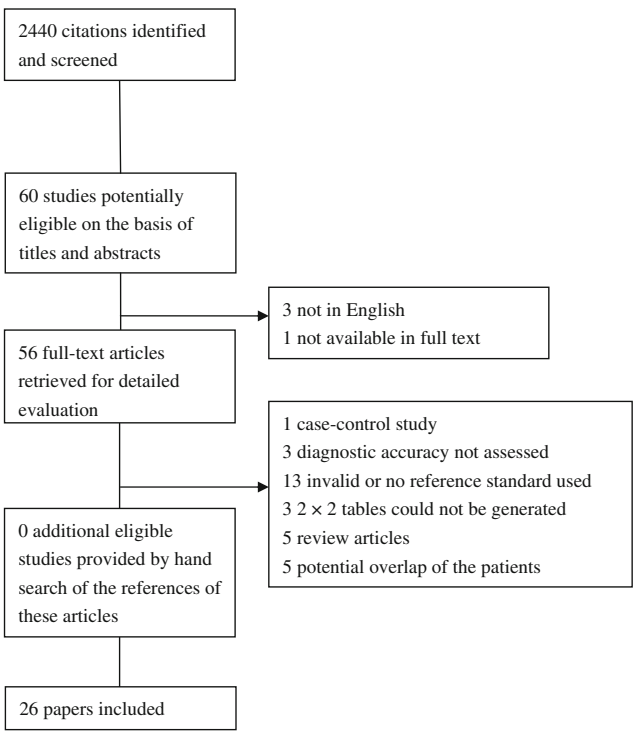


**Fig. 1** A flowchart shows the results of the literature search and selection for this systematic review. The computer search yielded 2440 citations; 26 studies ultimately were included.

**Table 1.** QUADAS criteria

| Methodologic criteria | Information required for "yes" |
|---|---|
| 1. Was the spectrum of patients representative of the patients who will receive the test in practice? | Patients with suspected scaphoid fracture were consecutively and/or prospectively recruited |
| 2. Were selection criteria clearly described? | Clear definition of suspected scaphoid fracture was provided |
| 3. Is the reference standard likely to classify the target condition correctly? | Followup image, or plus clinical examination and/or combined image, was used as reference standard, and the limit of followup was at least 6 weeks; the imaging modalities could be plain radiographs, CT, MRI, or bone scintigraphy |
| 4. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis? | All patients received a reference standard |
| 5. Did patients receive the same reference standard regardless of the index test result? | All patients received same reference standard |
| 6. Was the reference standard independent of the index test? | Index test results did not form part of reference standard |
| 7. Was the execution of the index test described in sufficient detail to permit replication of the test? | Clear description of test techniques and definitions of positive and negative test results were mentioned |
| 8. Was the execution of the reference standard described in sufficient detail to permit its replication? | |
| 9. Were the index test results interpreted without knowledge of the results of the reference standard? | Index test was interpreted without knowledge of reference standard results and vice versa, or test was clearly interpreted before results of other test were available |
| 10. Were the reference standard results interpreted without knowledge of the results of the index test? | |
| 11. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? | Data on patient age, gender, and presenting symptoms and physical signs were available |
| 12. Were uninterruptible/intermediate test results reported? | Results were available for all patients who entered the study |
| 13. Were withdrawals from the study explained? | Reasons why results were not available for all patients who entered the trial were reported or results were available for all patients |

The QUADAS criteria were taken from Table 2 in: Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25. Available at: http://www.biomedcentral.com/1471-2288/3/25.

ultrasound were not pooled because too few studies and patients were included (Table 4).

Most studies had some methodologic limitations (Table 5). Only two studies satisfied all criteria. Eighteen studies met at least 70% (nine items) of the criteria, and five studies met fewer than 50% (seven items) of the criteria. Greater than 70% of studies met the following items: adequate patient spectrum, avoidance of partial verification, independent reference test, adequate description of index test, blind assessment of index test, clinical data available, reporting of uninterruptible/intermediate test results, and explanation of withdrawals from the study; 54% (14 of 26) of studies met the item for valid reference test; and only 35% (nine of 26) and 27% (seven of 26) of studies met the items for describing details of the reference test and blind assessment of the reference test, respectively.

We calculated pooled estimates of sensitivity, specificity, and the natural logarithm of the diagnostic odds ratio (ln DOR) using the random-effects model. To pool the sensitivity and specificity, the variances of the raw proportions (r/n) were stabilized using a Freeman-Tukey-type arcsine square root transformation [35, 38]. To prevent division by 0 when pooling the ln DOR, conventional correction was applied by adding 0.5 to each cell when the $2 \times 2$ table for a study contained one or more 0 values. The diagnostic performance of each test also was assessed using SROC curves, which allow for a trend in DOR with threshold [36]. These curves can be defined by a regression model: $D = a + bS$, where D is ln DOR and S, a measure of the diagnostic threshold, is equivalent to the sum of logits of the true-positive rate and the false-positive rate. A limitation of this method is that it also requires adding 0.5 to each cell in the $2 \times 2$ table containing one or more 0 values. A b that is not different from 0 indicates the absence of a threshold effect. The SROC curve can be displayed graphically by plotting the predicted sensitivity across a range of values of 1 − specificity. Models were unweighted to provide parameter estimates similar to the random-effects model [28, 36]. We used the area under the curve and the Q* index, the intersection point of the SROC curve with a diagonal line from the left upper corner to the right lower corner of the ROC space, which corresponds to the highest common value of sensitivity and specificity for the test, as the global measures of test efficacy. A perfect

**Table 2.** Details of included studies

| Study | Year | Imaging | Male/female | Mean age (years) | Incidence of scaphoid fracture | Reference test |
|---|---|---|---|---|---|---|
| Rolfe et al. [43] | 1981 | BS | NR | 23 | 9% | Followup radiographs |
| Nielsen et al. [39] | 1983 | BS | 61/39* | 33 | 11% | 2 months' followup radiographs and clinical examination |
| Stordahl et al. [47] | 1984 | BS | 18/12† | 31 | 32% | 6 weeks' followup radiographs |
| Wilson et al. [54] | 1986 | BS | NR | NR | 5% | Day 10 radiographs |
| Brismar [10] | 1988 | BS | 117/70 | NR | 11% | Repeat radiographs |
| Tiel-van Buul et al. [50] | 1993 | BS | NR | NR | 17% | 6 weeks' followup radiographs |
| Waizenegger et al. [52] | 1994 | BS | NR | NR | 8% | Followup of 6–12 months |
| Murphy et al. [37] | 1995 | BS | 55/44* | 36 | 13% | Day 14 radiographs or bone scan |
| Thorpe et al. [48] | 1996 | BS MRI | 20/38* | 22 | 7% | 6 weeks' followup radiographs and clinical examination |
| Breitenseher et al. [7] | 1997 | MRI | 23/19 | 30.5 | 33% | 6 weeks' followup radiographs |
| Hunter et al. [27] | 1997 | MRI | 28/8 | 26 | 33% | 2 weeks' followup radiographs and clinical examination |
| Fowler et al. [19] | 1998 | BS MRI | 21/22 | 32 | 14% | Combined images or followup of at least 1 year |
| Kitsis et al. [30] | 1998 | BS MRI | 9/13 | 34 | 14% | 8 weeks' followup radiographs and clinical examination |
| Bretlau et al. [9] | 1999 | MRI | 27/25‡ | 44 | 16% | 8–14 weeks' followup radiographs |
| Gäbler et al. [21] | 2001 | MRI | 77/44 | 30 | 23% | Followup of 6 weeks |
| Hauger et al. [25] | 2002 | US | 35/19 | 26 | 9% | 10–14 days' images and clinical signs |
| Akdemir et al. [2] | 2004 | BS | 18/14 | 31 | 25% | Followup of 2 and 12 months with clinical signs and radiographs |
| Breederveld and Tuinebreijer [6] | 2004 | BS CT | NR | NR | 31% | Followup of 6 weeks or 14 months |
| Senall et al. [44] | 2004 | US | NR | 35 | 50% | ≥ 8 weeks' followup radiographs |
| Groves et al. [23] | 2005 | BS CT | 17/34 | 40.2 | 12% | 6 weeks' followup radiographs and MRI |
| Kumar et al. [32] | 2005 | MRI | 17/5 | 27 | 27% | Day 10 radiographs or MRI |
| Memarsadeghi et al. [34] | 2006 | MRI CT | 17/12 | 34 | 38% | 6 weeks' followup radiographs |
| Cruickshank et al. [14] | 2007 | CT | 26/21 | NR | 15% | Day 10–14 radiographs or MRI |
| You et al. [55] | 2007 | CT | NR | NR | 20% | Clinical and image followup |
| Beeres et al. [4] | 2008 | BS MRI | 50/50 | 42 | 20% | Combined images or clinical and radiographic examinations |
| Ty et al. [51] | 2008 | CT | 12/8 | 40 | 20% | ≥ 6 weeks' followup radiographs |

* One patient had bilateral injuries; †two patients were excluded because scaphoid fractures were already evident on review of the initial radiographs; ‡five of the 52 patients were excluded owing to technical problems with the MRI; two of the 47 enrolled patients had lost followup radiographs; NR = not reported; BS = bone scintigraphy; US = ultrasound.

test has an area under the curve of 1.0 and a Q* of 1.0, and a test with no diagnostic value has an area under the curve of 0.5 and a Q* of 0.5.

A dummy variable indicating the type of imaging modality was included as a covariate in the meta-regression model and the SROC model to compare the sensitivity, specificity, and overall diagnostic accuracy of bone scintigraphy, MRI, and CT with each other. A p value less than 0.05 of the regression coefficient of this variable was considered to indicate a significant difference. In the SROC model, the antilogarithm transformation of the estimated parameter can be interpreted as a relative DOR of the covariate. It indicates the change in diagnostic performance of the test under study per unit increase in the covariate.

We calculated derived likelihood ratios from the pooled sensitivities and specificities [56] and then used Bayes' theorem to calculate the probability of scaphoid fracture, conditioned by the likelihood ratios as a function of the pretest probability [15].

We assessed heterogeneity separately for sensitivity, specificity, and DOR using Cochran's Q test. Heterogeneity was defined as $p < 0.1$. The $I^2$ statistic was used to measure the percentage of variability among summary indices

**Table 3.** Details of included studies

| Study | Imaging technique | Period between test and injury | TP | FP | FN | TN |
|---|---|---|---|---|---|---|
| Bone scintigraphy | | | | | | |
| Rolfe et al. [43] | $^{99m}$Tc-MDP | 2–60 days | 9 | 17 | 0 | 73 |
| Nielsen et al. [39] | 10–15 mCi $^{99m}$Tc-MDP | 10 days | 11 | 43 | 0 | 47 |
| Stordahl et al. [47] | 10–15 mCi $^{99m}$Tc-DP | 2 weeks | 9 | 0 | 0 | 19 |
| Wilson et al. [54] | 500 MBq $^{99m}$Tc-MDP | Within 10 days | 2 | 6 | 0 | 34 |
| Brismar [10] | NR | 2–3 weeks | 21 | 9 | 0 | 157 |
| Tiel-van Buul et al. [50] | 200 MBq $^{99m}$Tc-MDP | 3–34 days | 21 | 14 | 0 | 90 |
| Waizenegger et al. [52] | 600 MBq $^{99m}$Tc-MDP | Within 14 days | 7 | 12 | 0 | 65 |
| Murphy et al. [37] | Three-phase $^{99m}$Tc-MDP | 4 days | 13 | 7 | 0 | 80 |
| Thorpe et al. [48] | 350–750 MBq $^{99m}$Tc-MDP | 3–4 weeks | 4 | 3 | 0 | 52 |
| Fowler et al. [19] | $^{99m}$Tc-MDP | 10–35 days | 5 | 2 | 1 | 35 |
| Kitsis et al. [30] | 550 MBq $^{99m}$Tc-HDP | 2–4 weeks | 3 | 1 | 0 | 18 |
| Akdemir et al. [2] | Three-phase 740 MBq $^{99m}$Tc-MDP | 2 weeks | 8 | 0 | 0 | 24 |
| Breederveld and Tuinebreijer [6] | Three-phase $^{99m}$Tc-MDP | 1–7 days | 7 | 2 | 2 | 18 |
| Groves et al. [23] | 400 MBq $^{99m}$Tc-MDP | NR | 6 | 4 | 0 | 41 |
| Beeres et al. [4] | 500 MBq $^{99m}$Tc-HDP | 3–5 days | 20 | 8 | 0 | 72 |
| MRI | | | | | | |
| Thorpe et al. [48] | T1, T2, STIR | 3–4 weeks | 4 | 1 | 0 | 54 |
| Breitenseher et al. [7] | T1, T2, STIR, 1.0 T | 0–7 days | 14 | 0 | 0 | 28 |
| Hunter et al. [27] | T1, T2, STIR, 1.5 T | 0–7 days | 10 | 1 | 0 | 19 |
| Fowler et al. [19] | T1, T2, STIR, 1.0 T | 10–35 days | 6 | 0 | 0 | 37 |
| Kitsis et al. [30] | T1, T2, 0.5 T | 2–4 weeks | 3 | 0 | 0 | 19 |
| Bretlau et al. [9] | T1, STIR | 2–10 days | 7 | 0 | 0 | 38 |
| Gäbler et al. [21] | T1, T2, STIR, 1.0 T | 0–7 days | 28 | 0 | 0 | 93 |
| Kumar et al. [32] | T1, STIR, 1.5 T | Within 1 day | 6 | 0 | 0 | 16 |
| Memarsadeghi et al. [34] | T1, T2, STIR, 1.0 T | 1–6 days | 11 | 0 | 0 | 18 |
| Beeres et al. [4] | T1, T2, 1.5 T | Within 24 hours | 16 | 0 | 4 | 80 |
| CT | | | | | | |
| Breederveld and Tuinebreijer [6] | Spiral, 1-mm slices | 0–4 days | 9 | 0 | 0 | 20 |
| Groves et al. [23] | 16 detector, 0.5-mm slices | NR | 6 | 0 | 0 | 45 |
| Memarsadeghi et al. [34] | Multidetector, 0.5-mm slices | 1–6 days | 8 | 0 | 3 | 18 |
| Cruickshank et al. [14] | 1-mm slices | 0–3 days | 7 | 0 | 0 | 40 |
| You et al. [55] | Multidetector, 1.0-mm slices | NR | 7 | 0 | 0 | 28 |
| Ty et al. [51] | 1.2-mm slices | NR | 4 | 0 | 0 | 16 |
| Ultrasound | | | | | | |
| Hauger et al. [25] | 12-MHz transducer | 0–7 days | 5 | 1 | 0 | 48 |
| Senall et al. [44] | 10-MHz transducer | 0–16 days | 7 | 1 | 2 | 8 |

MDP = methylene diphosphonate; DP = diphosphonate; HDP = hydroxymethylene diphosphonate; STIR = short tau inversion recovery; T1 = T1-weighted; T2 = T2-weighted; NR = not reported; TP = number of true positives; FP = number of false positives; FN = number of false negatives; TN = number of true negatives.

**Table 4.** Diagnostic accuracy of ultrasound for diagnosing suspected scaphoid fracture

| Study | Sensitivity (95% CI) | Specificity (95% CI) | ln DOR (95% CI) |
|---|---|---|---|
| Hauger et al. [25] | 1.00 (0.48–1.00) | 0.98 (0.89–1.00) | 5.87 (2.55–9.19) |
| Senall et al. [44] | 0.78 (0.40–0.97) | 0.89 (0.52–1.00) | 3.33 (0.73–5.94) |

CI = confidence interval; ln DOR = natural logarithm of the diagnostic odds ratio.

**Table 5.** Methodologic quality assessment according to the QUADAS criteria in Table 1

| Study | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rolfe et al. [43] | − | + | − | ? | − | + | + | − | − | − | − | + | + |
| Nielsen et al. [39] | + | − | + | + | + | + | + | − | + | ? | + | + | + |
| Stordahl et al. [47] | − | − | + | + | + | + | + | + | − | − | + | + | + |
| Wilson et al. [54] | + | + | − | + | + | + | + | + | ? | ? | + | + | + |
| Brismar [10] | ? | − | − | + | + | + | + | − | ? | ? | − | + | + |
| Tiel-van Buul et al. [50] | + | + | + | + | + | + | + | − | + | + | + | + | + |
| Waizenegger et al. [52] | + | − | + | − | − | + | + | − | + | − | ? | + | + |
| Murphy et al. [37] | + | + | − | + | + | + | − | − | + | + | + | + | + |
| Thorpe et al. [48] | + | − | + | + | + | + | − | − | + | ? | ? | − | + |
| Breitenseher et al. [7] | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Hunter et al. [27] | + | + | − | − | − | + | + | − | + | − | + | − | − |
| Fowler et al. [19] | + | − | − | + | − | − | + | + | + | − | + | + | + |
| Kitsis et al. [30] | + | + | + | + | + | + | − | − | + | ? | + | + | + |
| Bretlau et al. [9] | + | + | + | + | + | + | + | + | + | + | + | − | + |
| Gäbler et al. [21] | + | + | + | + | + | + | − | − | + | + | + | + | + |
| Hauger et al. [25] | + | + | − | + | + | + | + | − | + | ? | + | + | + |
| Akdemir et al. [2] | + | − | + | + | + | + | + | − | + | ? | + | + | + |
| Breederveld and Tuinebreijer [6] | + | + | + | + | − | + | − | − | + | − | + | + | + |
| Senall et al. [44] | + | + | + | + | + | + | − | + | + | ? | + | + | + |
| Groves et al. [23] | + | − | − | + | − | − | + | − | − | − | + | + | + |
| Kumar et al. [32] | + | + | − | − | − | + | + | + | + | − | + | + | + |
| Memarsadeghi et al. [34] | + | + | + | + | + | + | + | + | + | + | + | + | + |
| Cruickshank et al. [14] | + | + | − | + | + | + | − | − | + | + | + | − | + |
| You et al. [55] | − | − | − | + | − | − | + | − | + | − | ? | + | + |
| Beeres et al. [4] | + | + | − | + | − | − | + | + | + | − | + | + | + |
| Ty et al. [51] | + | + | + | + | + | + | + | − | + | ? | + | − | − |

All criteria were scored yes (+), no (−), or unclear (?).

caused by heterogeneity rather than chance [26]. An $I^2$ value of 0 indicates no heterogeneity and greater than 50% suggests substantial heterogeneity. To explore sources of heterogeneity, we performed univariable meta-regression analysis if the included studies were no fewer than 10. The following covariates were tested: sample size ($\leq$ 50 patients versus $>$ 50 patients), prevalence of scaphoid fracture ($\leq$ 15% versus $>$ 15%), period between injury and index tests ($\leq$ 10 versus $>$ 10 days), and study quality (high versus low). We considered a study high quality when it scored positive on at least four of the following items: prospective design with consecutive recruitment, appropriate reference standard, avoidance of partial verification, avoidance of differential verification, and interpretation of reference test without knowledge of index test results. Any covariate showing an association with sensitivity, specificity, or DOR ($p < 0.1$) was selected, and subgroups of studies identified by such covariates underwent separate meta-analysis. We also performed sensitivity analysis by excluding the outlier studies when there was substantial heterogeneity. Outlier studies were detected by means of Galbraith plots [22].

The analyses were performed using Stata®/SE9.0 (StataCorp LP, College Station, TX) and Meta-DiSc, Version 1.4 (Hospital Universitario Ramo′n y Cajal, Madrid, Spain).

## Results

Bone scintigraphy, MRI, and CT all had high pooled sensitivities, specificities, and ln DORs (Figs. 2, 3, 4; Table 6). We found no differences in sensitivity among the three tests. The specificity of bone scintigraphy was worse than that of MRI ($p < 0.001$) and CT ($p = 0.001$). No difference ($p = 0.94$) in specificity was found between MRI and CT. The DOR of MRI was greater ($p = 0.009$) than that of bone scintigraphy. We found no differences in DOR between MRI and CT ($p = 0.63$) and between CT and bone scintigraphy ($p = 0.12$). There was no heterogeneity

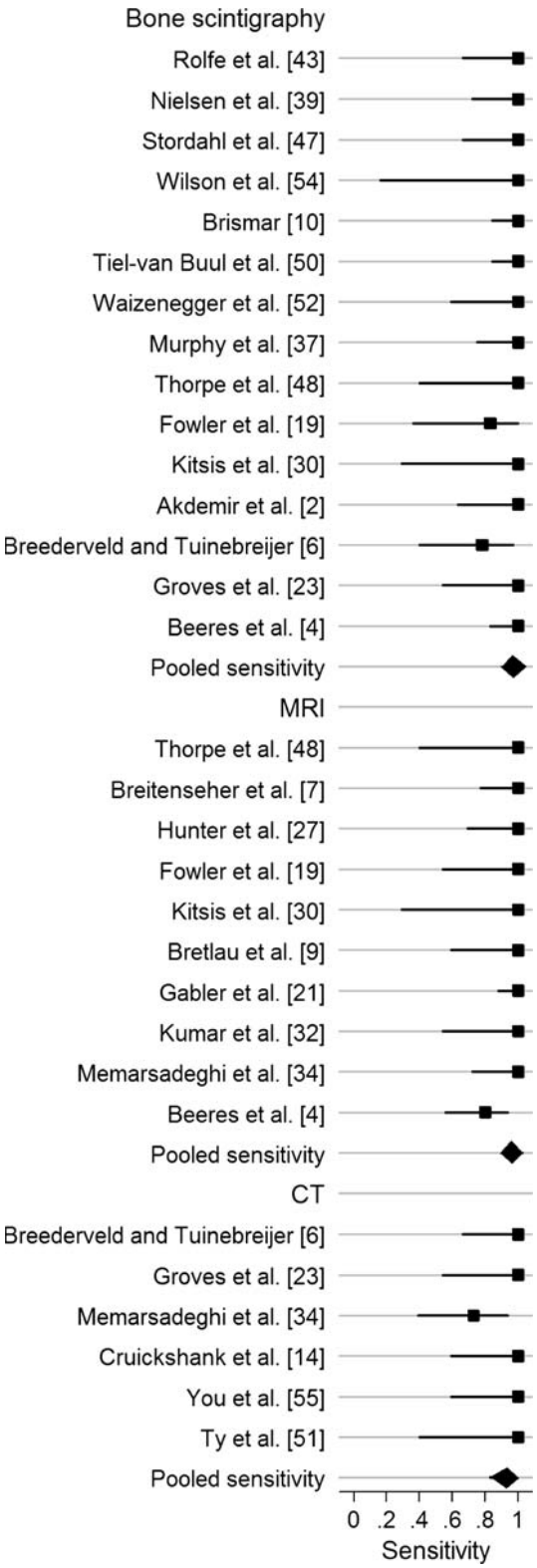


**Fig. 2** The pooled sensitivity for bone scintigraphy was 97% (95% CI, 93%–99%) for bone scintigraphy, 96% (95% CI, 91%–99%) for MRI, and 93% (95% CI, 83%–98%) for CT.

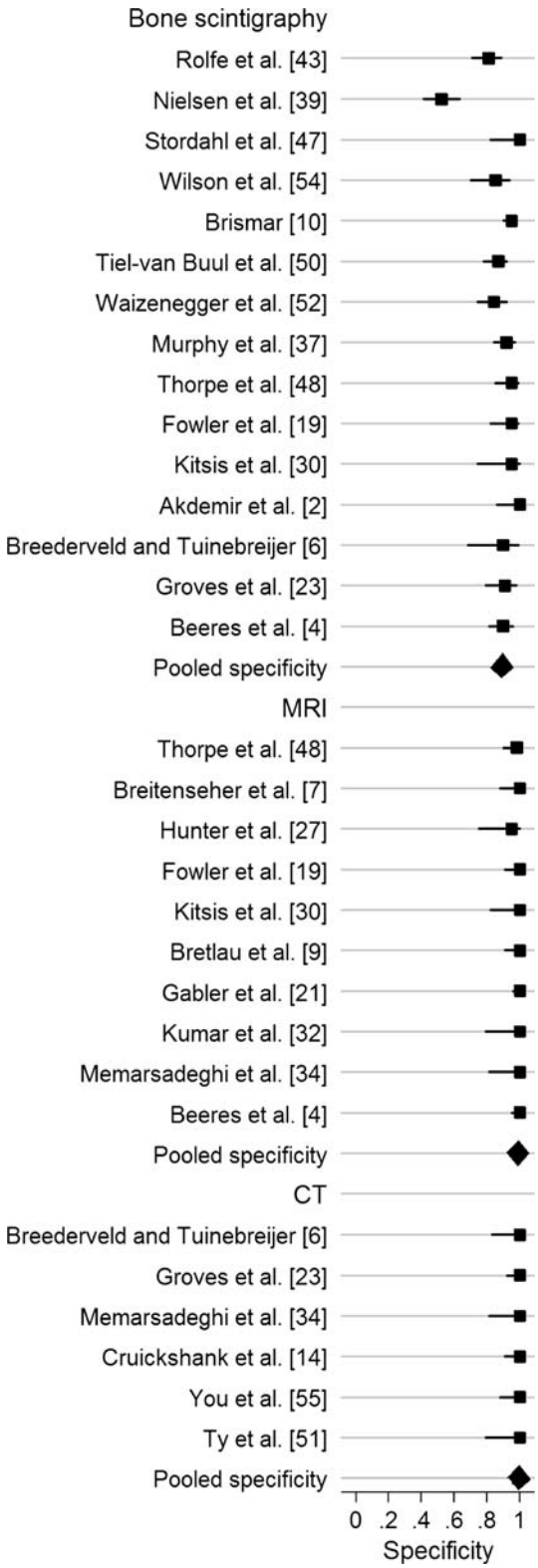


**Fig. 3** The pooled specificity for bone scintigraphy was 89% (95% CI, 83%–94%) for bone scintigraphy, 99% (95% CI, 96%–100%) for MRI, and 99% (95% CI, 96%–100%) for CT.

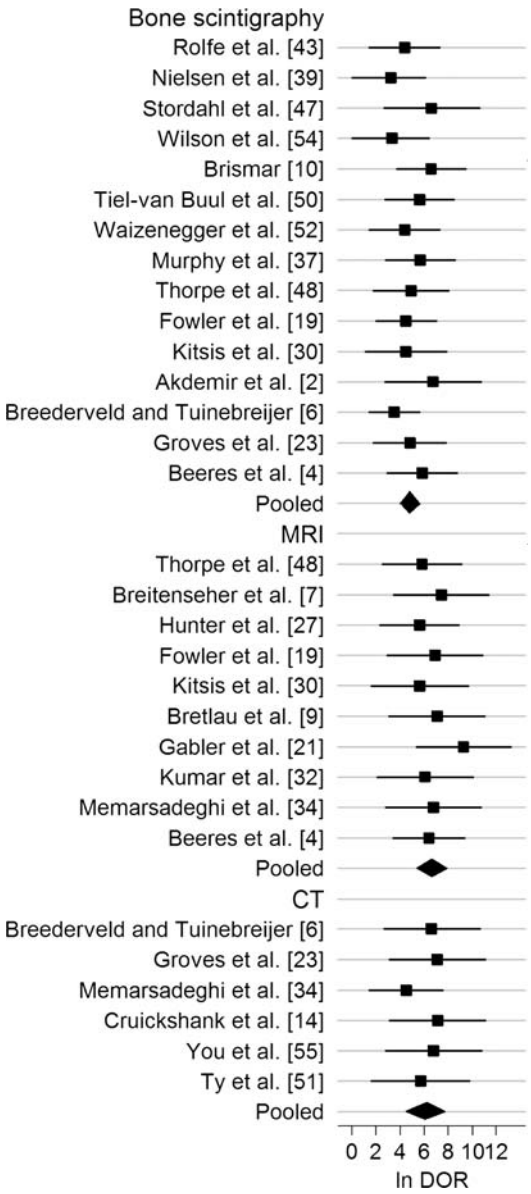


**Fig. 4** The pooled ln DOR was 4.78 (95% CI, 4.02–5.54) for bone scintigraphy, 6.60 (95% CI, 5.43–7.76) for MRI, and 6.11 (95% CI, 4.56–7.66) for CT.

in sensitivity and DOR for all modalities and in specificity for MRI and CT. Heterogeneity was significant for specificity with bone scintigraphy (Table 7). Univariate meta-regression analysis showed the period between injury and index test was associated ($p = 0.08$) with variation in specificity of bone scintigraphy. Longer period was associated with higher specificity and lower heterogeneity: 91% (95% confidence interval [CI], 87%–95%; heterogeneity, $I^2 = 64\%$) for more than 10 days and 82% (95% CI, 65%–95%; heterogeneity, $I^2 = 92\%$) for 10 days or less. We found none of the other variables, such as sample size, prevalence, or study quality, were a source of variability. Four outlier studies [2, 10, 39, 47] were identified on the Galbraith plot (Fig. 5). After exclusion of the four studies, the pooled sensitivity, specificity, and ln DOR were, respectively, 96% (95% CI, 91%–99%; heterogeneity, $I^2 = 0$), 88% (95% CI, 85%–91%; heterogeneity, $I^2 = 12\%$), and 4.58 (95% CI, 3.73–5.44; heterogeneity, $I^2 = 0$), which were similar to the pooled estimates with inclusion of the outliers. In examining the four outliers, we could not identify the potential reasons for the outlier results.

The SROC model showed all modalities had an area under the curve close to 1, a Q* index close to 1, and the parameter b not different from 0, indicating high diagnostic performance and no evidence of a threshold effect for any modality (Table 8). Comparative SROC analysis showed MRI had higher diagnostic performance than bone scintigraphy (relative DOR = 4.85; 95% CI, 1.42–16.52; $p = 0.01$) and similar diagnostic performance to CT (relative DOR = 1.41; 95% CI, 0.04–4.96; $p = 0.56$); a difference between CT and bone scintigraphy was not found (relative DOR = 3.61; 95% CI, 0.76–17.07; $p = 0.10$) (Fig. 6).

The positive and negative likelihood ratios derived from the pooled sensitivities and specificities were 8.82 and 0.03 for bone scintigraphy, 96 and 0.04 for MRI, and 93 and 0.07 for CT. For each modality, we calculated the posttest probability according to the pretest probability and the derived likelihood ratios (Fig. 7). With a pretest probability of 18%, which was the mean prevalence of true fracture derived from 1165 patients in 21 included studies in which patients with suspected scaphoid fracture were consecutively and/or prospectively recruited, negative results from bone scintigraphy, MRI, and CT would reduce the posttest probability to 0.7%, 0.9%, and 1.5%, respectively. Positive findings from MRI or CT would increase the posttest probability to 95%; however, positive findings from bone scintigraphy would increase the posttest probability to only 66%.

**Table 6.** Pooled estimates of sensitivity, specificity, and ln DOR

| Imaging modalities | Number of studies | Number of patients | Sensitivity (95% CI) | Specificity (95% CI) | ln DOR (95% CI) |
|---|---|---|---|---|---|
| Bone scintigraphy | 15 | 1102 | 97% (93%–99%) | 89% (83%–94%) | 4.78 (4.02–5.54) |
| MRI | 10 | 513 | 96% (91%–99%) | 99% (96%–100%) | 6.60 (5.43–7.76) |
| CT | 6 | 211 | 93% (83%–98%) | 99% (96%–100%) | 6.11 (4.56–7.66) |

CI = confidence interval; ln DOR = natural logarithm of the diagnostic odds ratio.

**Table 7.** Test for heterogeneity

| Modality | Sensitivity | | | Specificity | | | Diagnostic odds ratio | | |
|---|---|---|---|---|---|---|---|---|---|
| | Q | p | $I^2$ | Q | p | $I^2$ | Q | p | $I^2$ |
| Bone scintigraphy | 9.01 | 0.83 | 0 | 91.83 | < 0.0001 | 85% | 8.13 | 0.88 | 0 |
| MRI | 8.54 | 0.48 | 0 | 4.87 | 0.85 | 0 | 2.89 | 0.97 | 0 |
| CT | 5.48 | 0.36 | 9% | 0.22 | 1.00 | 0 | 1.73 | 0.89 | 0 |

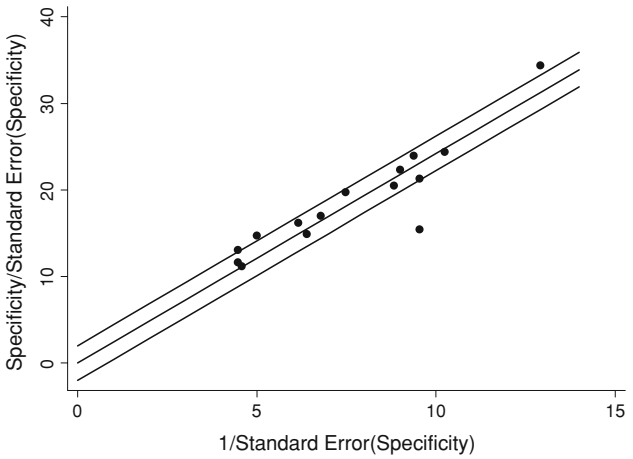


**Fig. 5** A Galbraith plot for specificity of studies describing bone scintigraphy identified four outlier studies.

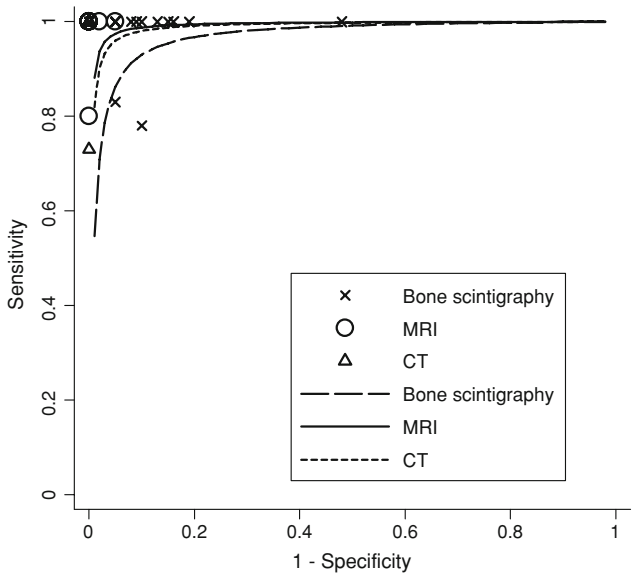


**Fig. 6** The SROC curve for MRI was closest to the left upper corner, followed by CT and bone scintigraphy, indicating MRI had the highest overall diagnostic performance, then CT and bone scintigraphy.

## Discussion

Early diagnosis and treatment for scaphoid fractures are critical to improving outcomes [3]; however, early diagnosis sometimes can be difficult to establish because some fractures are radiographically occult in the acute phase after injury [46]. There is marked inconsistency in imaging protocols for suspected scaphoid fractures among investigators and hospitals [13, 18, 24, 41, 42, 49]. We performed this systematic review and meta-analysis to produce and compare summary estimates of sensitivity, specificity, and DOR of bone scintigraphy, MRI, and CT for diagnosing suspected scaphoid fractures; estimate and compare the overall diagnostic performance of the three modalities using the SROC curve approach; and calculate the probability of scaphoid fractures associated with positive or negative imaging results using Bayes' theorem.

This study has several limitations. First, we included only studies published as a full report in English. We identified three potentially eligible nonEnglish studies [8, 31, 33], two of which [8, 31] involved potential overlap of patient population with an included study [7]. Second, we did not test for publication bias, which hampers most meta-analyses of diagnostic tests [45]. Routine testing for publication bias may not be an especially useful paradigm in the context of systematic reviews of diagnostic tests [5]. If diagnostic test data are subject to the same sort of bias as clinical trials, then one might expect publication bias would lead to an overestimate of diagnostic performance in

**Table 8.** Results of the summary receiver operating characteristic curve model

| Modality | Parameter | | | Area under the curve (standard error) | Q* point (standard error) |
|---|---|---|---|---|---|
| | a | b | p | | |
| Bone scintigraphy | 5.00 | −0.11 | 0.66 | 0.9676 (0.0095) | 0.916 (0.0149) |
| MRI | 5.65 | −0.03 | 0.95 | 0.9923 (0.0038) | 0.9644 (0.0102) |
| CT | 7.69 | 0.86 | 0.27 | 0.9886 (0.0073) | 0.955 (0.017) |

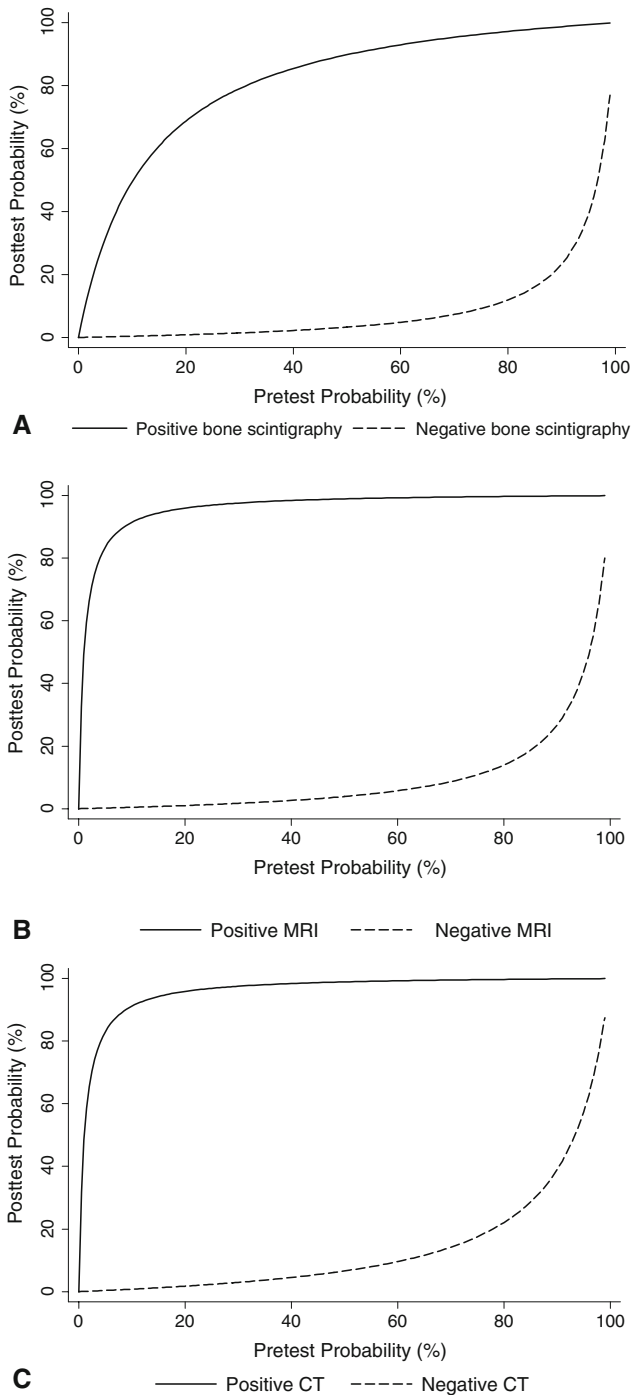


**Fig. 7A–C** (**A**) In patients with low pretest probability, positive results with bone scintigraphy cannot accurately confirm fracture, but negative results can accurately exclude fracture. (**B**) In patients with low pretest probability, positive results with MRI can reliably confirm fracture; negative results can accurately exclude fracture. (**C**) In patients with low pretest probability, positive results with CT can reliably confirm fracture; negative results can accurately exclude fracture.

meta-analysis [17]. Third, most included studies had a variable level of methodologic limitations. One major limitation is the lack of an ideal reference standard for diagnosing a true scaphoid fracture. The most commonly used standard is the absence of radiographic evidence of a scaphoid fracture on scaphoid-specific radiographs obtained a minimum of 6 weeks after injury, but we believe this is somewhat unsatisfying [21]. The other two major limitations were interpreting the reference test with knowledge of the test result and reporting no details of the reference test. However, it is not entirely clear how individual aspects of quality may affect estimates of diagnostic accuracy and to what magnitude [12]. Finally, the comparisons of the three imaging techniques presented here were indirect, which are prone to confounding.

Our meta-analysis showed MRI and CT had high sensitivities and specificities; bone scintigraphy was comparable to them in sensitivity but had lower specificity. It has been recognized increased $^{99m}$Tc-methylene diphosphonate (MDP) uptake is sensitive to metabolic change; however, it is not specific regarding the underlying causes, such as fracture, bone bruise, soft tissue injury, and inflammation at adjacent joints [23]. Our analysis also showed the heterogeneity in specificity for bone scintigraphy among studies was substantial and the period between injury and the test was associated with the variation: longer period was associated with higher specificity and reduced heterogeneity. The reason for this finding is unclear. It might be, during the delay in testing, some injuries increasing $^{99m}$Tc-MDP uptake, except for fracture, might have been alleviated. Although the sensitivity for CT has been questioned [3], our analysis did not show a difference in sensitivity between CT and MRI or bone scintigraphy. However, there was imprecision around the estimate of sensitivity for CT (wide CI), indicating the need for more data. Two short-cut systematic reviews in this area have been published [18, 42]. One reported the average sensitivity and specificity were, respectively, 96% and 89% for bone scintigraphy; 98% and 99% for MRI; and 94% and 96% for CT, which were similar to our results [42]. The review included more studies than for the same modality in our review; however, it did not report the search strategy and inclusion criteria or list the publications. The other review was restricted to comparison between MRI and bone scintigraphy and only included four comparative diagnostic cohort studies [18], of which one was not included in our review because of invalid reference test. Both reviews preferred MRI; however, neither performed statistical pooling, comparison, and assessment of heterogeneity.

The SROC analysis showed the three modalities all have high overall diagnostic performance. MRI is superior to bone scintigraphy, but CT is not superior to MRI or bone scintigraphy. The major advantage of SROC analysis is that it can assess the threshold effect and compare the performance of tests at different tradeoff points [36].

However, the following issues should be considered when interpreting our findings. The continuity correction with adding 0.5 to each cell in the 2 × 2 table containing one or more 0 values was performed in 13 of 15 data sets on bone scintigraphy and all data sets on MRI and CT, so downward bias would be introduced to the estimated SROC curves. These results were derived from small samples, especially results for CT, which were derived from 211 patients in six studies. It has been recognized in the diagnostic study literature that small studies tend to overestimate the effect size [45]. Small samples also mean low power in detecting the threshold effect and the difference between imaging modalities. Although we found no threshold effect for each modality, we still included the variable S in the model when the modalities were compared with each other, which yielded conservative results.

The positive likelihood ratios of MRI and CT are greater than 90, but that of bone scintigraphy was less than 10. All three modalities have negative likelihood ratios less than 0.1. The likelihood ratios indicate the extent of change in the odds of disease after a test result. As a general rule, positive likelihood ratios greater than 10 and negative likelihood ratios less than 0.1 are considered to provide strong evidence to rule in or rule out diagnoses, respectively, in most circumstances [15]. Our findings allow for calculation of the posttest probability of scaphoid fracture, provided the pretest probability has been estimated before the test. With a pretest probability of 18%, negative results from bone scintigraphy, MRI, and CT are enough to exclude a scaphoid fracture. Positive findings from MRI or CT provide strong evidence for the presence of a scaphoid fracture; however, positive findings from bone scintigraphy are insufficient to confirm fracture. The prevalence of true scaphoid fracture we presented was greater than that used by some investigators to calculate the posttest probability [1, 42], but similar to the result of a recent epidemiologic study [29], in which the prevalence of true scaphoid fracture in patients with suspected scaphoid fracture was 16%. The prevalence of abnormality in a study sample rarely can be generalized beyond the study except when the study is based on a suitable random sample [15]. Although our data were derived from 1165 patients in 21 studies in which patients with suspected scaphoid fractures were recruited consecutively and/or prospectively, it should be cautiously generalized.

Based on the current evidence, MRI is highly accurate for confirming and excluding the diagnosis of scaphoid fractures and might be used as the first choice in a patient with suspected scaphoid fracture. Bone scintigraphy is inappropriate for confirming scaphoid fractures. More studies are needed to assess the diagnostic performance of CT, especially paired design studies or randomized controlled trials to compare CT with MRI or bone scintigraphy.

**Acknowledgments** We thank Ma Ning for literature collection.

## References

1. Adey L, Souer JS, Lozano-Calderon S, Palmer W, Lee SG, Ring D. Computed tomography of suspected scaphoid fractures. *J Hand Surg Am.* 2007;32:61–66.
2. Akdemir UO, Atasever T, Sipahioğlu S, Türkölmez S, Kazimoğlu C, Sener E. Value of bone scintigraphy in patients with carpal trauma. *Ann Nucl Med.* 2004;18:495–499.
3. Amrami KK. Radiology corner: diagnosing radiographically occult scaphoid fractures—what's the best second test? *J Am Soc Surg Hand.* 2005;5:134–138.
4. Beeres FJ, Rhemrev SJ, den Hollander P, Kingma LM, Meylaerts SA, le Cessie S, Bartlema KA, Hamming JF, Hogervorst M. Early magnetic resonance imaging compared with bone scintigraphy in suspected scaphoid fractures. *J Bone Joint Surg Br.* 2008;90:1205–1209.
5. Begg CB. Systematic reviews of diagnostic accuracy studies require study by study examination: first for heterogeneity, and then for sources of heterogeneity. *J Clin Epidemiol.* 2005;58:865–866.
6. Breederveld RS, Tuinebreijer WE. Investigation of computed tomographic scan concurrent criterion validity in doubtful scaphoid fracture of the wrist. *J Trauma.* 2004;57:851–854.
7. Breitenseher MJ, Metz VM, Gilula LA, Gaebler C, Kukla C, Fleischmann D, Imhof H, Trattnig S. Radiographically occult scaphoid fractures: value of MR imaging in detection. *Radiology.* 1997;203:245–250.
8. Breitenseher MJ, Trattnig S, Gäbler C, Happel B, Bankier A, Kukla C, Rand T, Imhof H. [MRI in radiologically occult scaphoid fractures: initial experiences with 1.0 Tesla (whole body-middle field equipment) versus 0.2 Tesla (dedicated low-field equipment)] [in German]. *Radiologe.* 1997;37:812–818.
9. Bretlau T, Christensen OM, Edström P, Thomsen HS, Lausten GS. Diagnosis of scaphoid fracture and dedicated extremity MRI. *Acta Orthop Scand.* 1999;70:504–508.
10. Brismar J. Skeletal scintigraphy of the wrist in suggested scaphoid fracture. *Acta Radiol.* 1988;29:101–107.
11. Brydie A, Raby N. Early MRI in the management of clinical scaphoid fracture. *Br J Radiol.* 2003;76:296–300.
12. Cnossen JS, Vollebregt KC, de Vrieze N, ter Riet G, Mol BW, Franx A, Khan KS, vam der Post JA. Accuracy of mean arterial pressure and blood pressure measurements in predicting pre-eclampsia: systematic review and meta-analysis. *BMJ.* 2008;336:1117–1120.
13. Cooney WP 3rd. Scaphoid fractures: current treatments and techniques. *Instr Course Lect.* 2003;52:197–208.
14. Cruickshank J, Meakin A, Breadmore R, Mitchell D, Pincus S, Hughes T, Bently B, Harris M, Vo A. Early computerized tomography accurately determines the presence or absence of scaphoid and other fractures. *Emerg Med Australas.* 2007;19:223–228.
15. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329:168–169.
16. De Vries SO, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. *Acad Radiol.* 1996;3:361–369.
17. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess.* 2005;9:1–113, iii.

18. Foex B, Speake P, Body R. Best evidence topic report: magnetic resonance imaging or bone scintigraphy in the diagnosis of plain x ray occult scaphoid fractures. *Emerg Med J.* 2005;22:434–435.
19. Fowler C, Sullivan B, Williams LA, McCarthy G, Savage R, Palmer A. A comparison of bone scintigraphy and MRI in the early diagnosis of the occult scaphoid waist fracture. *Skeletal Radiol.* 1998;27:683–687.
20. Fusetti C, Poletti PA, Pradel PH, Garavaglia G, Platon A, Della Santa DR, Bianchi S. Diagnosis of occult scaphoid fracture with high-spatial-resolution sonography: a prospective blind study. *J Trauma.* 2005;59:677–681.
21. Gäbler C, Kukla C, Breitenseher MJ, Trattnig S, Vécsei V. Diagnosis of occult scaphoid fractures and other wrist injuries: are repeated clinical examinations and plain radiographs still state of the art? *Langenbecks Arch Surg.* 2001;386:150–154.
22. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med.* 1988;7:889–894.
23. Groves AM, Cheow HK, Balan KK, Bearcroft PW, Dixon AK. 16 detector multislice CT versus skeletal scintigraphy in the diagnosis of wrist fractures: value of quantification of 99Tcm-MDP uptake. *Br J Radiol.* 2005;78:791–795.
24. Groves AM, Kayani I, Syed R, Hutton BF, Bearcroft PP, Dixon AK, Ell PJ. An international survey of hospital practice in the imaging of acute scaphoid trauma. *AJR Am J Roentgenol.* 2006;187:1453–1456.
25. Hauger O, Bonnefoy O, Moinard M, Bersani D, Diard F. Occult fractures of the waist of the scaphoid: early diagnosis by high-spatial-resolution sonography. *AJR Am J Roentgenol.* 2002;178:1239–1245.
26. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557–560.
27. Hunter JC, Escobedo EM, Wilson AJ, Hanel DP, Zink-Brody GC, Mann FA. MR imaging of clinically suspected scaphoid fractures. *AJR Am J Roentgenol.* 1997;168:1287–1293.
28. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol.* 1995;48:119–130; discussion 131–132.
29. Jenkins PJ, Slade K, Huntley JS, Robinson CM. A comparative analysis of the accuracy, diagnostic uncertainty and cost of imaging modalities in suspected scaphoid fractures. *Injury.* 2008;39:768–774.
30. Kitsis C, Taylor M, Chandey J, Smith R, Latham J, Turner S, Wade P. Imaging the problem scaphoid. *Injury.* 1998;29:515–520.
31. Kukla C, Gaebler C, Breitenseher MJ, Trattnig S, Vécsei V. [Prospective comparison of MRI vs. direct magnification radiography in occult fractures of the scaphoid bone] [in German]. *Unfallchirurg.* 1998;101:32–36.
32. Kumar S, O'Connor A, Despois M, Galloway H. Use of early magnetic resonance imaging in the diagnosis of occult scaphoid fractures: the CAST Study (Canberra Area Scaphoid Trial). *N Z Med J.* 2005;118:U1296.
33. Lindequist S, Houshian S, Rosbach SB, Riegels-Nielsen P. [MRI scanning in suspected fractures of the scaphoid bone] [in Danish]. *Ugeskr Laeger.* 1998;160:7438–7441.
34. Memarsadeghi M, Breitenseher MJ, Schaefer-Prokop C, Weber M, Aldrian S, Gäbler C, Prokop M. Occult scaphoid fractures: comparison of multidetector CT and MR imaging—initial experience. *Radiology.* 2006;240:169–176.
35. Mills EJ, Nachega JB, Buchan I, Orbinski J, Attaran A, Singh S, Rachlis B, Wu P, Cooper C, Thabane L, Wilson K, Guyatt GH, Bangsberg DR. Adherence to antiretroviral therapy in Sub-Saharan Africa and North America: a meta-analysis. *JAMA.* 2006;296:679–690.
36. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med.* 1993;12:1293–1316.
37. Murphy DG, Eisenhauer MA, Powe J, Pavlofsky W. Can a day 4 bone scan accurately determine the presence or absence of scaphoid fracture? *Ann Emerg Med.* 1995;26:434–438.
38. Naguib M, Kopman AF, Ensor JE. Neuromuscular monitoring and postoperative residual curarisation: a meta-analysis. *Br J Anaesth.* 2007;98:302–316.
39. Nielsen PT, Hedeboe J, Thommesen P. Bone scintigraphy in the evaluation of fracture of the carpal scaphoid bone. *Acta Orthop Scand.* 1983;54:303–306.
40. Phillips TG, Reibach AM, Slomiany WP. Diagnosis and management of scaphoid fractures. *Am Fam Physician.* 2004;70:879–884.
41. Plancher KD. Methods of imaging the scaphoid. *Hand Clin.* 2001;17:703–721.
42. Ring D, Lozano-Calderón S. Imaging for suspected scaphoid fracture. *J Hand Surg Am.* 2008;33:954–957.
43. Rolfe EB, Garvie NW, Khan MA, Ackery DM. Isotope bone imaging in suspected scaphoid trauma. *Br J Radiol.* 1981;54:762–767.
44. Senall JA, Failla JM, Bouffard JA, van Holsbeeck M. Ultrasound for the early diagnosis of clinically suspected scaphoid fracture. *J Hand Surg Am.* 2004;29:400–405.
45. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol.* 2002;31:88–95.
46. Steinmann SP, Adams JE. Scaphoid fractures and nonunions: diagnosis and treatment. *J Orthop Sci.* 2006;11:424–431.
47. Stordahl A, Schjøth A, Woxholt G, Fjermeros H. Bone scanning of fractures of the scaphoid. *J Hand Surg Br.* 1984;9:189–190.
48. Thorpe AP, Murray AD, Smith FW, Ferguson J. Clinically suspected scaphoid fracture: a comparison of magnetic resonance imaging and bone scintigraphy. *Br J Radiol.* 1996;69:109–113.
49. Tiel-van Buul MM, Roolker W, Broekhuizen AH, Van Beek EJ. The diagnostic management of suspected scaphoid fracture. *Injury.* 1997;28:1–8.
50. Tiel-van Buul MM, van Beek EJ, Broekhuizen AH, Bakker AJ, Bos KE, van Royen EA. Radiography and scintigraphy of suspected scaphoid fracture: a long-term study in 160 patients. *J Bone Joint Surg Br.* 1993;75:61–65.
51. Ty JM, Lozano-Calderon S, Ring D. Computed tomography for triage of suspected scaphoid fractures. *Hand (NY).* 2008;3:155–158.
52. Waizenegger M, Wastie ML, Barton NJ, Davis TR. Scintigraphy in the evaluation of the "clinical" scaphoid fracture. *J Hand Surg Br.* 1994;19:750–753.
53. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol.* 2003;3:25. Available at: http://www.biomedcentral.com/1471-2288/3/25.
54. Wilson AW, Kurer MH, Peggington JL, Grant DS, Kirk CC. Bone scintigraphy in the management of X-ray-negative potential scaphoid fractures. *Arch Emerg Med.* 1986;3:235–242.
55. You JS, Chung SP, Chung HS, Park IC, Lee HS, Kim SH. The usefulness of CT for patients with carpal bone fractures in the emergency department. *Emerg Med J.* 2007;24:248–250.
56. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. *Stat Med.* 2008;27:687–697.