Surgical Compared with Conservative Treatment for Acute Nondisplaced or Minimally Displaced Scaphoid Fractures A Systematic Review and Meta-Analysis of Randomized Controlled Trials

Buijze, G.A.; Doornberg, J.N.; Ham, J.S.; Ring, D.; Bhandari, M.; Poolman, R.W.

Link to publication

Surgical Compared with Conservative Treatment for Acute Nondisplaced or Minimally Displaced Scaphoid Fractures

A Systematic Review and Meta-Analysis of Randomized Controlled Trials

By Geert A. Buijze, MD, Job N. Doornberg, MD, PhD, John S. Ham, MD, PhD, David Ring, MD, PhD,
Mohit Bhandari, MD, and Rudolf W. Poolman, MD, PhD

*Investigation performed at the Department of Orthopaedic Surgery, Academic Medical Center, Amsterdam,
and at the Department of Joint Research, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands*

Background: There is a current trend in orthopaedic practice to treat nondisplaced or minimally displaced fractures with early open reduction and internal fixation instead of cast immobilization. This trend is not evidence-based. In this systematic review and meta-analysis, we pool data from trials comparing surgical and conservative treatment for acute nondisplaced and minimally displaced scaphoid fractures, thus aiming to summarize the best available evidence.

Methods: A systematic literature search of the medical literature from 1966 to 2009 was performed. We selected eight randomized controlled trials comparing surgical with conservative treatment for acute nondisplaced or minimally displaced scaphoid fractures in adults. Data from included studies were pooled with use of fixed-effects and random-effects models with standard mean differences and risk ratios for continuous and dichotomous variables, respectively. Heterogeneity across studies was assessed with calculation of the I^2 statistic.

Results: Four hundred and nineteen patients from eight trials were included. Two hundred and seven patients were treated surgically, and 212 were treated conservatively. Most trials lacked scientific rigor. Our primary outcome parameter, standardized functional outcome, which was assessed for 247 patients enrolled in four trials, significantly favored surgical treatment ($p < 0.01$). With regard to our secondary parameters, we found heterogeneous results that favored surgical treatment in terms of satisfaction (assessed in one study), grip strength (six studies), time to union (three studies), and time off work (five studies). In contrast, we found no significant differences between surgical and conservative treatment with regard to pain (two studies), range of motion (six studies), the rates of nonunion (six studies) and malunion (seven studies), and total treatment costs (two studies). The rate of complications was higher in the surgical treatment group (23.7%) than in the conservative group (9.1%), although this difference was not significant ($p = 0.13$). There was a nearly significantly higher rate of scaphotrapezial osteoarthritis in the surgical treatment group ($p = 0.05$).

Conclusions: Based on primary studies with limited methodological quality, this study suggests that surgical treatment is favorable for acute nondisplaced and minimally displaced scaphoid fractures with regard to functional outcome and time off work; however, surgical treatment engenders more complications. Thus, the long-term risks and short-term benefits of surgery should be carefully weighed in clinical decision-making.

Traditionally, nondisplaced and minimally displaced fractures have been considered by most surgeons to be stable. Several studies have demonstrated predictable rates of healing in association with conservative treatment for these types of fractures, ranging from 90% to 100%[1-7]. In contrast, displaced scaphoid fractures—defined by most authors as those with >1 mm of displacement—have been recognized as unstable fractures that are associated with a

significant risk of nonunion if not treated surgically[8]. Currently, there is a trend in orthopaedic practice toward early open reduction and internal fixation for the treatment of fractures that traditionally have been treated conservatively[9-11]. Possible explanations for this trend are increased patient expectations and more complete follow-up, including the emphasis on patient-based outcome measures of function and health status[12].

Recent reports have advocated surgical treatment for nondisplaced and minimally displaced scaphoid fractures[13-15]. Two previous meta-analyses on the treatment of acute scaphoid fractures indicated that current evidence failed to prove the superiority of operative as compared with conservative treatment[16,17]. However, since the publication of those reviews, additional data have been published, and the discussion about the indications for operative treatment is ongoing[13,18]. Even though the present meta-analysis has substantial overlap with previous ones, we believe that it adds methodological quality and incorporates recently published data to provide the best currently available evidence to support clinical decision-making[19].

The aim of the present review was to evaluate the evidence from randomized controlled clinical trials comparing surgical and conservative treatment for acute nondisplaced and minimally displaced fractures of the scaphoid. Furthermore, we aimed to evaluate methodological limitations in current studies evaluating these therapeutic options for scaphoid fractures. The primary outcome measures of the two treatment groups were patient-rated and physician-rated wrist function.

Materials and Methods

The present study is reported following the QUOROM (Quality of Reporting of Meta-analyses) guidelines[20].

Types of Studies, Participants, and Interventions

In order to be included in the present systematic review, a study had to be a randomized or quasi-randomized controlled trial that compared surgical and conservative treatment of acute nondisplaced or minimally displaced scaphoid fractures in adult patients. A quasi-randomized trial is one in which participants are allocated to treatment in a manner that is not strictly random (e.g., date of birth, hospital record number).

The present review excluded (1) trials that included children or patients with congenital deformities or degenerative conditions and (2) trials that focused on the treatment of delayed union (four weeks to six months after a fracture) or nonunion (more than six months after a fracture).

Surgical intervention included internal fixation by means of either an open or percutaneous approach. Conservative modalities included all types of cast immobilization regardless of the length of the cast, immobilization of the thumb, or position of the hand in the cast.

Types of Outcome Measures

The primary outcome measure of this review was functional outcome based on validated hand and wrist function scores, including the Disabilities of the Arm, Shoulder and Hand (DASH) score[21], the Patient Evaluation Measure (PEM)[22], the Patient-Rated Wrist Evaluation (PRWE)[23], and the modified Green and O'Brien score[24]. The secondary outcome measures of this review were patient satisfaction, pain, physician-rated functional outcome (such as range of wrist motion and grip strength), time to union, time off work, return to previous activity, costs, and negative outcomes (including the rates of infection, malunion, and nonunion). Nonunion was defined as failure of the fracture to unite at more than six months after injury, with radiographic evidence of a fracture line.

Search Methods for Identification of Studies

Our methods for the identification of studies were similar to the Cochrane Bone, Joint and Muscle Trauma Group methods used in reviews. A systematic search, from inception of the database in 1966 to February 27, 2009, was performed with the following terms: scaph* and fractur* and random*. Databases included the Cochrane Bone, Joint and Muscle Trauma Group Specialised Register, the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, CINAHL, and reference lists of articles. Researchers in the field were contacted to inquire about any additional unpublished trials or trials in progress.

Selection of Studies

Two trained orthopaedic research physicians (G.A.B. and J.N.D.) assessed abstracts of all studies identified during the initial search. Full copies of the report of potentially relevant studies were independently assessed by those two physicians with use of the above criteria. Disagreements were resolved by means of discussion, with arbitration by an experienced orthopaedic surgeon with training in clinical epidemiology (R.W.P.) when differences of opinion remained.

Assessment of Methodological Quality

The quality of the selected studies was independently assessed, without masking the source of authorship of the trial reports, by two physicians (G.A.B. and J.N.D.) with use of a quality-assessment tool derived from the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group[25]. To evaluate the methodological quality of the selected studies, we scored the methods of allocation and blinding and the loss to follow-up. The GRADE criteria were used to evaluate the quality of evidence according to outcome. Disagreement was resolved by means of discussion and, if necessary, by means of scrutiny by an independent orthopaedic surgeon with a doctorate in study methodology (R.W.P.).

Data Collection

The two physicians independently extracted data from all eligible studies with use of a piloted data-extraction form. Disagreement was resolved as described above. All authors of the selected trials were contacted to obtain unpublished original raw data in order to complete the data-extraction form. For all continuous outcomes, means and standard deviations were extracted for analysis. If means and confidence intervals were reported instead, standard deviations were calculated from

these values. If the required data could not be derived from the published data, as was the case for five of the eight selected studies, we referred to unpublished data obtained from the authors. Incomplete data (e.g., means without standard deviations) could not be used and were excluded from this meta-analysis.

Data Pooling Across Studies

Outcome measures in the eligible studies were reported with use of several different measurement tools. To increase the generalizability of our results where possible, comparable continuous outcome data were pooled across studies with use of the method of standard mean differences with random effects. This model calculates the error term for both within-study and interstudy variability in the meta-analysis. Data from studies using the same measurement tools were plainly pooled. If different measurement tools were used, data were standardized for comparison and were pooled with use of the following methods.

For our primary outcome measure, we pooled comparable scores from different patient-reported functional outcome instruments when these instruments scored disability on a 100-point scale, with 100 representing the least or most disability. If more than one outcome instrument was used, the average outcome was used for analysis. Using the same method, we pooled comparable scores of pain as reported on a 100-point scale, with 0 and 100 representing the least and most pain, respectively.

Range of motion was reported as flexion and extension and/or radial and ulnar deviation as a percentage of the value for the uninjured wrist. When range of motion was reported in more than one plane, outcomes were standardized by calculating the means of different planes, with equal weighting of the different planes. We then calculated the percentage lost in comparison with the uninjured side by subtracting values from 100% (e.g., 13% lost instead of 87% remaining).

Grip strength measurements were reported as a percentage of the value for the uninjured hand in most studies and were reported in kilograms in some studies. When the percentage compared with the uninjured side was provided, we calculated the percentage lost in comparison with the uninjured side by subtracting values from 100% as described above. When only the measurements in kilograms were provided, data were standardized according to the average values from the Guides to the Evaluation of Permanent Impairment[26]. The derived values of average grip strength were 47 kg for men and 24 kg for women. We then calculated the percentage lost in comparison with the average grip strength values, adjusted for sex.

In addition, three experienced orthopaedic surgeons (J.S.H., D.R., and M.B.) graded complications in terms of severity in order to pool the data across studies for comparison on the basis of patient-important end points[27]. Three subgroups of complications were created: low, moderate, and severe. Low-grade complications were defined as transient or short-term consequences and included superficial infection. Moderate-grade complications were defined as long-term consequences that did not require additional surgery and included osteoarthritis. Severe-grade complications were defined as (1) long-term consequences that required additional surgery, including screw malposition, hardware removal, malunion, nonunion, and osteonecrosis, and (2) complex regional pain syndrome.

Data Analysis

For both treatment groups in each study, standard mean differences and 95% confidence intervals were calculated for continuous outcomes and risk ratios (i.e., relative risk)[28] and 95% confidence intervals were calculated for dichotomous outcomes. Treatment effect was defined as significant if $p < 0.05$. Heterogeneity between studies was tested with use of both the chi-square test (with significance defined as $p < 0.1$) and the I^2 test (with substantial heterogeneity defined as values of >50%)[29].

Source of Funding

Results

Literature Search

The search resulted in 120 potentially eligible studies, eight of which met our inclusion criteria (Fig. 1)[13,14,18,30-34]. All included studies were randomized controlled trials. In total, 419 patients from eight trials were included. Two hundred and seven patients were treated surgically, and 212 were treated conservatively.

Description of Included Studies

The study characteristics are summarized in a table in the Appendix.

Adolfsson et al. reported on fifty-three patients (fourteen women and thirty-nine men) with a mean age of thirty-one years (range, fifteen to seventy-five years) who had a recent (less than fourteen-day-old) nondisplaced fracture of the waist of the scaphoid (a Herbert[35] type-B1 or B2 fracture)[30]. The authors randomly allocated twenty-five patients to operative treatment (percutaneous screw fixation) and twenty-eight patients to conservative treatment (cast immobilization). Patients were followed for a minimum of sixteen weeks; 26% were lost to follow-up.

Arora et al. randomly allocated forty-seven patients with nondisplaced waist (Herbert type-B2) fractures of the scaphoid to percutaneous scaphoid fixation (twenty-three patients) and cast treatment (twenty-four patients)[31]. The follow-up group consisted of forty-four patients (twelve women and thirty-two men) with a mean age of thirty-three years (range, twenty to fifty-six years). Patients were followed for a minimum of twenty-four weeks; 6% were lost to follow-up.

Bond et al. reported on twenty-five patients (three women and twenty-two men) with a mean age of twenty-four years (range, eighteen to thirty-four years) who had sustained an acute nondisplaced fracture of the scaphoid waist (a Herbert type-A2 or B2 fracture)[14]. Eleven patients were randomized to percutaneous screw fixation, and fourteen were randomized to cast immobilization. The mean duration of follow-up for each group was twenty-five months; no patient was lost to follow-up.
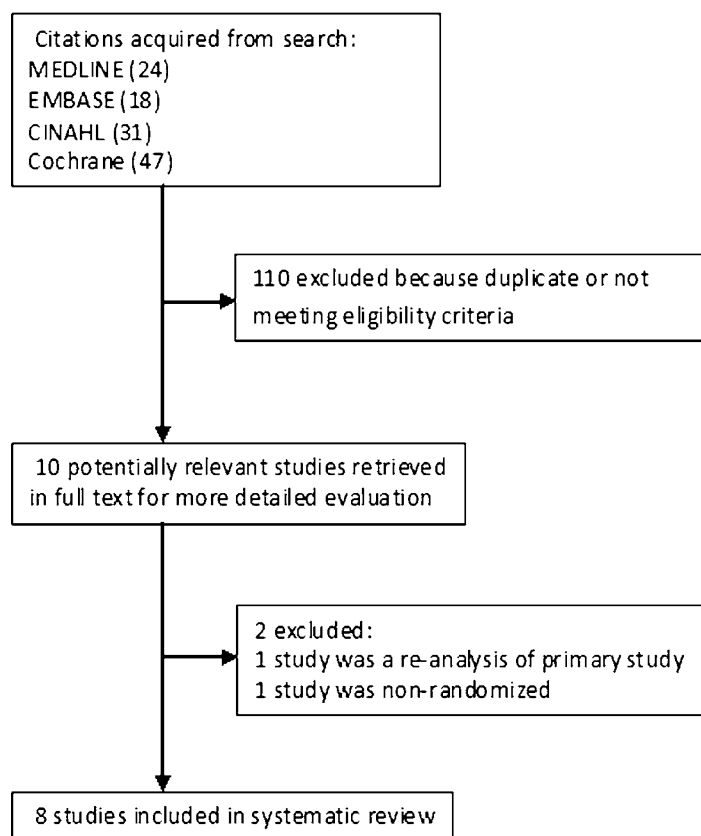
Fig. 1

Study flow diagram of the systematic review.

Dias et al. reported on eighty-eight patients with a mean age of 29.5 years (range, sixteen to sixty-one years) who had a clear bicortical nondisplaced or minimally displaced fracture of the waist of the scaphoid (a Herbert type-A2, B2, or B5 fracture)[32]. The study group included nine female and seventy-nine male patients. Forty-four patients were allocated to each treatment group, consisting of internal fixation with a Herbert screw or cast immobilization. Minimal initial fracture displacement was present in association with three fractures in the conservative treatment group and eight fractures in the surgical treatment group. Patients were followed for a minimum of fifty-two weeks; 8% were lost to follow-up.

McQueen et al. randomly allocated sixty consecutive patients with a nondisplaced or minimally displaced fracture of the waist of the scaphoid (a Herbert type-B1 or B2 fracture) to percutaneous screw fixation or immobilization in a cast[13]. Each group consisted of thirty patients. Minimal initial fracture displacement was present in association with two fractures in the conservative group and five fractures in the surgical group. There were ten female and fifty male patients with a mean age of 29.4 years (range, seventeen to sixty-five years). Patients were followed for a minimum of fifty-two weeks; 8% were lost to follow-up.

Saedén et al. randomized sixty-one patients with sixty-two acute fractures of the scaphoid (AO type-C2 and C3 fractures) to surgical treatment with use of a Herbert screw

(thirty-two fractures) and conservative treatment with use of a cast (thirty fractures)[33]. The mean age at the time of injury was thirty-three years (29 ± 13 years for the operative treatment group and 37 ± 20 for the conservative treatment group). Thirteen fractures were in females, and forty-nine were in males. Patients were followed for a minimum of twelve years; 18% were lost to follow-up.

In 2007, Vinnars et al.[34] studied the data for a subgroup of fifty-two patients from the same randomized trial reported by Vinnars et al. in 2008[18]. They included acute nondisplaced fractures of the scaphoid (Herbert type-A2, B1, B2, and B3 fractures). There were sixteen female patients and thirty-six male patients. The median age was thirty-two years for the casting group and twenty-nine years for the surgical treatment group. Both groups included twenty-six patients. Only additional data from this report (i.e., data not reported by Vinnars et al. in 2008) were used for analysis. The mean duration of follow-up was 10.2 years, and no patient was lost to follow-up.

In 2008, Vinnars et al. reported on eighty-five patients (nineteen women and sixty-six men) with a mean age of thirty-two years who had sustained an acute nondisplaced scaphoid fracture (a Herbert type-A2, B1, B2, or B3 fracture)[18]. Forty-three patients were allocated to receive surgical treatment, and forty-two were allocated to receive conservative treatment. Patients were followed for a mean of 10.2 years; 11% were lost to follow-up.

TABLE I GRADE Quality Assessment of Trials for Impact of Surgical Treatment Compared With Conservative Treatment (Control)							➤
		Quality Assessment					
Outcome	No. of Studies	Design	Limitations	Inconsistency*	Indirectness†	Imprecision‡	Other Considerations
Functional outcome (range of scores: 0 to 100; better indicated by less)	4	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Satisfaction (range of scores: 0 to 5; better indicated by more)	1	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious‡‡	None
Pain (range of scores: 0 to 100; better indicated by less)	2	Randomized trial	Serious**	Serious§§	No serious indirectness	Serious††	None
Range of wrist motion (range of scores: 0 to 100; better indicated by less)	6	Randomized trial	Serious**	Serious§§	No serious indirectness	Serious††	None
Grip strength (range of scores: 0 to 100; better indicated by less)	6	Randomized trial	Serious**	Serious§§	No serious indirectness	Serious††	None
Time to union (better indicated by less)	3	Randomized trial	Serious**	Serious§§	No serious indirectness	Serious‡‡	None
Infection	7	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Nonunion	6	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Malunion	7	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Time off work (better indicated by less)	5	Randomized trial	Serious**	Serious§§	No serious indirectness	Serious††	None
Return to previous activity	2	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Total costs (better indicated by less)	2	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None
Complications	7	Randomized trial	Serious**	No serious inconsistency	No serious indirectness	Serious††	None

*Inconsistency refers to an unexplained heterogeneity of results. †Indirectness refers to an indirect comparison of groups (e.g., differences between populations). ‡Imprecision refers to studies with relatively few patients and few events and thus with wide confidence intervals around the estimate of the effect. §The values are given as the risk ratio, with the 95% confidence interval in parentheses. #The values are given as the standard mean difference, with the 95% confidence interval in parentheses, or as the difference in the number of cases per 1000. **Unclear allocation concealment in three studies, patients not blinded in any study, outcome assessors blinded in only one study, >25% loss to follow-up in one study. ††Confidence interval includes possible benefits from both types of treatment. ‡‡Sample size not optimal for adequate precision. §§Unexplained heterogeneity.

Outcome Measure Reporting
The reported outcomes of the included studies are summarized in a table in the Appendix.

Four of the eight studies evaluated the primary outcome measure, functional outcome based on validated function scores[13,18,31,32]. Two studies evaluated the DASH score[18,31], including one study that evaluated the DASH and PRWE scores[18], one evaluated the PEM score[32], and one evaluated the modified Green and O'Brien score[13].

One study evaluated patient satisfaction[14], and two studies evaluated pain[31,32]. Seven studies evaluated range of motion and grip strength[13,14,18,30-33]. Three studies evaluated time to union[13,14,31], seven evaluated infection[13,14,18,30-33], six evaluated the rate of nonunion[14,18,30-33], and seven evaluated the rate of

TABLE I (continued)

Summary of Findings					
No. of Patients		Effect			
Surgical	Conservative	Relative§	Absolute#	Quality	Importance
123	124	–	–0.62 (–0.89 to –0.36)	Low	Critical
11	14	–	2.06 (1.06 to 3.07)	Low	Critical
60	65	–	0.49 (–0.76 to 1.74)	Very low	Critical
159	161	–	–0.16 (–0.7 to 0.37)	Very low	Critical
159	161	–	–2.59 (–4.24 to –0.94)	Very low	Critical
61	63	–	–4.2 (–7.69 to –0.7)	Very low	Important
2/186 (1.1%)	0/188 (0%)	3.25 (0.35 to 30.34)	0 more per 1000	Low	Critical
3/166 (1.8%)	5/169 (3.0%)	0.72 (0.18 to 2.86)	5 fewer per 1000	Low	Critical
0/186 (0%)	5/188 (2.7%)	0.17 (0.02 to 1.39)	0 fewer per 1000	Low	Critical
110	108	–	–1.69 (–2.7 to –0.68)	Very low	Important
65/69 (94.2%)	61/63 (96.8%)	0.98 (0.91 to 1.05)	19 fewer per 1000	Low	Important
47	49	–	–0.1 (–0.5 to 0.3)	Low	Important
46/194 (23.7%)	18/197 (9.1%)	1.96 (0.82 to 4.67)	0 more per 1000	Low	Critical

malunion[13,14,18,30-33]. Six studies evaluated time off work[13,14,31-34], three evaluated return to previous activity[13,18,30], and two evaluated costs[31,34]. All studies evaluated complications[13,14,18,30-34].

Methodological Quality
The methodological quality of the eligible trials was limited, as shown in a table in the Appendix. Even though all studies were randomized controlled trials, concealment of allocation was unclear in three studies[30,31,33], no study was blinded, and outcome assessors were blinded in only one study[13]. The rate of loss to follow-up, reported in all studies, ranged from 0% to 26%. The authors of four of the eight studies clearly stated that their analysis was based on intention-to-treat principles[13,18,32,34]. The authors of three studies reported on cross-over to surgical treatment[18,32,33]; however, Saedén et al.[33] did not clearly state whether they performed an intention-to-treat analysis. The quality assessments of all outcome measurements as well as the effect of surgical as compared with conservative treatment are described in Table I.

Functional Outcome
Four[13,18,31,32] of the eight selected studies evaluated the primary outcome measure of this meta-analysis: functional outcome based on validated function scores. Their data could be successfully pooled, and the groups were homogeneous. Analysis revealed a significant difference in treatment effect, in favor of surgical treatment (standard mean difference = –0.62, 95% confidence interval = –0.89 to –0.36, p < 0.01, I^2 = 5%) (Fig. 2).

Study or Subgroup	Operative			Conservative			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Arora	4	4.2	21	13	14.1	23	17.6%	-0.83 [-1.45, -0.21]	
Dias	3.9	1	39	5.2	1.8	42	31.3%	-0.88 [-1.33, -0.42]	
McQueen	6.3	7.9	23	12.3	19.2	24	20.0%	-0.40 [-0.98, 0.18]	
Vinnars 2008	4.5	1	40	5	1.5	35	31.1%	-0.39 [-0.85, 0.06]	
Total (95% CI)			**123**			**124**	**100.0%**	**-0.62 [-0.89, -0.36]**	

Heterogeneity: Tau² = 0.00; Chi² = 3.16, df = 3 (P = 0.37); I² = 5%
Test for overall effect: Z = 4.62 (P < 0.00001)

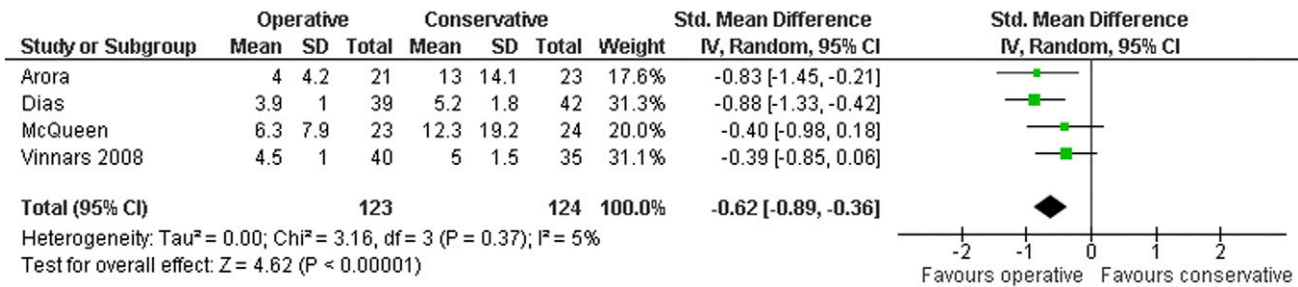Favours operative   Favours conservative

Fig. 2

Table and forest plot illustrating functional outcome based on validated outcome scores. SD = standard deviation, Std. = standard, CI = confidence interval, df = degrees of freedom. IV = inverse variance.

Secondary Outcomes

Patient satisfaction was only reported in one study[14]. Within that study, there was a significant difference in treatment effect, in favor of surgical treatment (standard mean difference = 2.06, 95% confidence interval = 1.06 to 3.07, p < 0.01).

Measurements of pain could be pooled from two studies[31,32]. Analysis of these data showed significant heterogeneity and no significant differences in the pooled treatment effect (standard mean difference = 0.49, 95% confidence interval = –0.76 to 1.74, p = 0.44, I² = 91%).

Seven studies evaluated range of motion and grip strength[13,14,18,30-33], but the data from one of those studies[30] could not be included because no standard deviations were provided. Thus, data for range of motion were pooled across six studies[13,14,18,31-33]. Analysis of those data revealed no significant difference in pooled treatment effect and considerable heterogeneity (standard mean difference = –0.16, 95% confidence interval = –0.70 to 0.37, p = 0.55, I² = 81%). Data for grip strength were pooled across the same six studies[13,14,18,31-33]. Analysis of those data revealed substantial heterogeneity between treatment groups and a significant difference in pooled treatment effect in favor of surgical treatment (standard mean difference = –2.59, 95% confidence interval = –4.24 to –0.94, p < 0.01, I² = 97%).

There was a significantly shorter time to union in the surgical treatment groups as reported in three studies[13,14,31] (standard mean difference = –4.20, 95% confidence interval = –7.69 to –0.70, p = 0.02, I² = 96%); however, heterogeneity was

considerable. The rate of infection was not significantly different across seven studies[13,14,18,30-33] (risk ratio = 3.25, 95% confidence interval = 0.35 to 30.34, p = 0.3, I² = 0%). There were no significant differences in the rates of nonunion across six studies[14,18,30-33] (risk ratio = 0.72, 95% confidence interval = 0.18 to 2.86, p = 0.64, I² = 0%) or malunion across seven studies[13,14,18,30-33] (risk ratio = 0.17, 95% confidence interval = 0.02 to 1.39, p = 0.1, I² = 0%).

With regard to time off work, the data from one study[32] could not be included because they were incomplete, even after contacting the authors. The analysis of time off work from five studies[13,14,31,33,34] revealed a significant difference in favor of surgical treatment (standard mean difference = –1.69, 95% confidence interval = –2.70 to –0.68, p < 0.01, I² = 89%) (Fig. 3). With regard to return to previous activity, one study had incomplete data[30] and therefore only two studies[13,18] could be analyzed. Analysis of those data revealed the groups to be perfectly homogeneous. There was no significant difference in treatment effect (risk ratio = 0.98, 95% confidence interval = 0.91 to 1.05, p = 0.54, I² = 0%). The data from two studies on the total costs of treatment[31,34] were pooled, and the analysis showed no significant difference between treatment groups (standard mean difference = –0.10, 95% confidence interval = –0.50 to 0.30, p = 0.2, I² = 0%).

Data on complications were reported in all studies[13,14,18,30-34]. The data were graded in terms of severity in order to be pooled across studies and were analyzed in three subgroups. Low and

Study or Subgroup	Operative			Conservative			Weight	Std. Mean Difference IV, Random, 95% CI	Std. Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
Arora	1.1	1.2	21	7.9	7.1	23	22.8%	-1.28 [-1.94, -0.63]	
Bond	8	0.7	11	15	0.7	14	7.7%	-9.67 [-12.70, -6.65]	
McQueen	3.8	3.3	28	11.4	8.6	24	23.2%	-1.18 [-1.78, -0.59]	
Saeden	6	3	24	15	10	21	22.9%	-1.23 [-1.88, -0.59]	
Vinnars 2007	67	61	26	101	101	26	23.4%	-0.40 [-0.95, 0.15]	
Total (95% CI)			**110**			**108**	**100.0%**	**-1.69 [-2.70, -0.68]**	

Heterogeneity: Tau² = 1.06; Chi² = 37.53, df = 4 (P < 0.00001); I² = 89%
Test for overall effect: Z = 3.28 (P = 0.001)
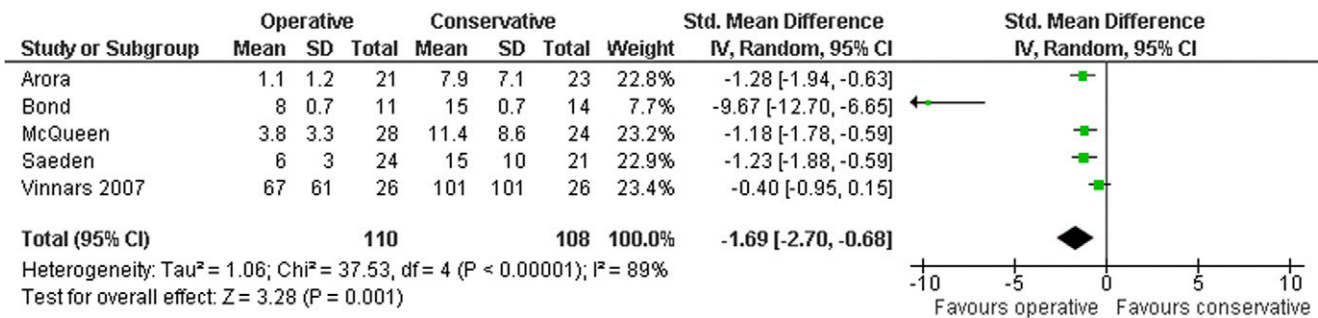
Favours operative   Favours conservative

Fig. 3

Table and forest plot illustrating time off work. SD = standard deviation, Std.= standard, CI = confidence interval, df = degrees of freedom. IV = inverse variance.
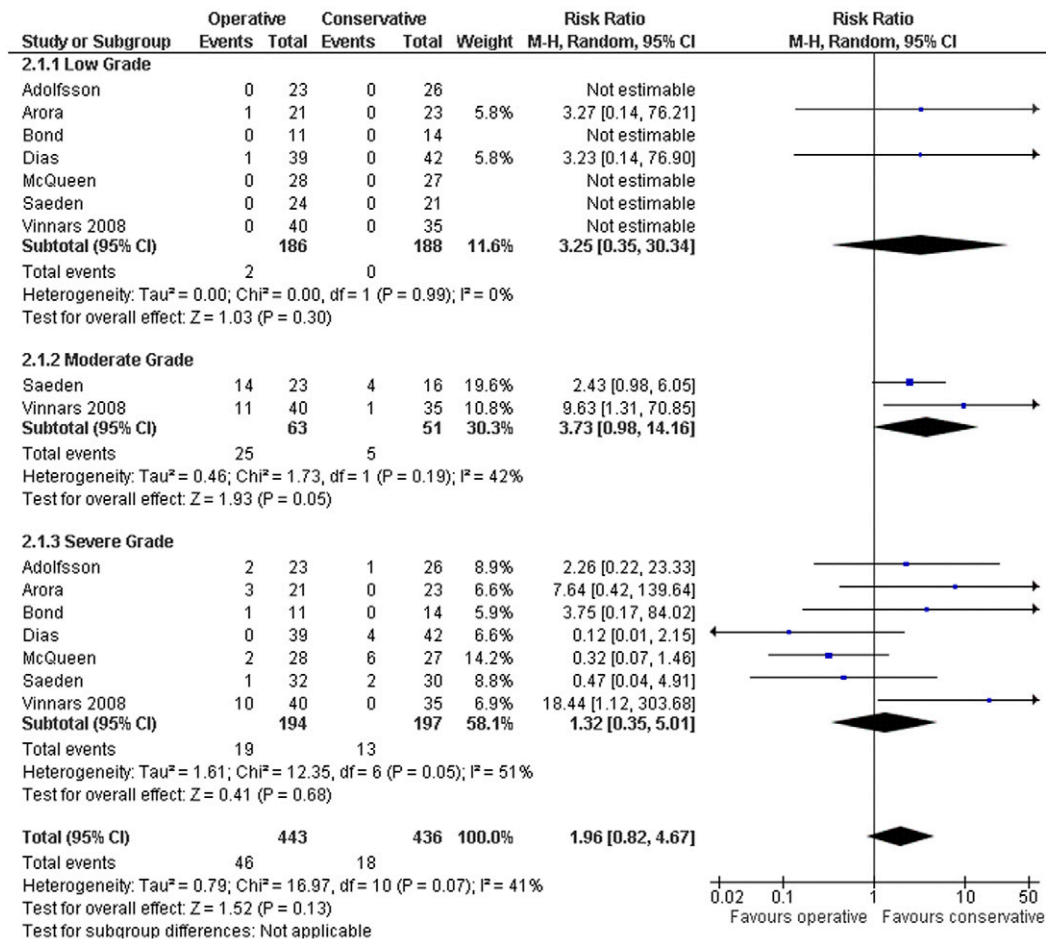
Fig. 4

Table and forest plot illustrating the rate of complications. CI = confidence interval, df = degrees of freedom. M-H = Mantel-Haenszel.

severe-grade complications were reported in seven studies[13,14,18,30-33], and moderate-grade complications were reported in two[18,33]. The low and moderate-grade subgroups and the combined group were homogeneous; the severe grade showed moderate heterogeneity. Analysis within the low-grade subgroup (risk ratio = 3.25, 95% confidence interval = 0.35 to 30.34, p = 0.30, I^2 = 0%) and the severe-grade subgroup (risk ratio = 1.32, 95% confidence interval = 0.35 to 5.01, p = 0.68, I^2 = 51%) showed no significant difference in pooled treatment effect, and analysis of the moderate-grade subgroup (risk ratio = 3.73, 95% confidence interval = 0.98 to 14.16, p = 0.05, I^2 = 42%) was nearly significant in favor of conservative treatment. Meta-analysis of complications for overall treatment effect did not significantly favor one treatment (risk ratio = 1.96, 95% confidence interval = 0.82 to 4.67, p = 0.13, I^2 = 41%) (Fig. 4).

Discussion
Key Findings
According to the best available evidence, this meta-analysis suggests that (1) surgical treatment of nondisplaced and minimally displaced fractures results in significantly better patient-reported functional outcome, greater patient satisfaction, better grip strength, a shorter time to union, and earlier return to work and that (2) there are no significant differences between surgical and conservative treatment with regard to pain; range of motion; the rates of nonunion, malunion, infection, or complications; or total treatment cost. However, the results of this meta-analysis should be interpreted with caution because of the following limitations.

The wide range of follow-up in the primary studies, ranging from sixteen weeks to twelve years, represents an important limitation and a general source of heterogeneity for this meta-analysis. Our primary outcome measurement, functional outcome, generally improves over time, as illustrated in two of the primary studies[13,32] by a functional improvement trend between eight and fifty-two weeks of follow-up[32]. Post-immobilization active functional use of the hand and wrist plays a key role in improving function. Therefore, the fact that conservatively managed patients have had a relatively shorter mobilization time at the time of the latest follow-up evaluation (because of prolonged immobilization times) engenders bias

with regard to functional outcome, in favor of the surgical treatment group.

Several secondary outcome measurements showed significant heterogeneity. The six studies pooled for grip strength[13,14,18,31-33] showed a large variability, and none of the studies could be identified as causing major heterogeneity. Several authors have reported considerable variability in the relative strength of the two hands in a healthy population, and grip strength has been shown to be typically greater on the dominant side than on the nondominant side[36-38]. Therefore, the rate of dominant-hand injury, ranging from 34% to 64% across the five studies, could account for some of the grip strength variability among studies. We consider the physiological variability in grip strength within the population, the wide range in the rate of dominant-hand injury, and the wide range of follow-up times (as previously described) to be the main factors in this large variation among studies generating substantial heterogeneity.

Time to union was pooled from three studies[13,14,31] and showed substantial heterogeneity. The most plausible reason for this heterogeneity is the fact that time to union represents an inaccurate and imprecise measurement, especially for scaphoid fractures[39]. Moreover, the diagnosis of scaphoid union was ascertained with use of radiographs, which have poor reliability[40]. Computed tomography seems to be the most reliable imaging technique for predicting scaphoid fracture union and should therefore be used as a reference standard for healing[41]. Only one of the included trials[31] involved the use of a computed tomography scan as a standard of care to determine scaphoid fracture-healing.

Measurements of time off work were pooled from five studies[13,14,31,33,34], and we identified the study by Bond et al.[14] as the most important source of heterogeneity. A plausible reason for this was their study population, which consisted solely of full-time military personnel. Time off work in that study was defined as time until the patients returned to a full-duty status, which could most likely not be performed while wearing an arm cast. In other studies, time off work was shorter than the time of cast immobilization in many cases. Thus, many patients (especially non-manual laborers) returned to work prior to cast removal.

With regard to complications, although none of the differences for the subgroups or the combined group were significant, there were still some clinically important differences that should be addressed. Regarding low-grade complications, there were only two cases of postoperative infection among all 207 patients who were allocated to surgical treatment in the seven primary trials. With regard to the risk of osteoarthritis after conservative treatment—traditionally used to support a recommendation for surgery—the meta-analysis of the best available evidence suggests the opposite to be true. According to data from two studies[18,33], osteoarthritis develops in 40% of patients after surgery as compared with 10% of patients after cast immobilization, a nearly significant difference (p = 0.05). Regarding severe-grade complications, additional surgery was required in 7.7% of patients in the surgical treatment group as compared with 6.0% of those in the con-

servative treatment group, and complex regional pain syndrome was reported in four patients in the surgical treatment group and in one patient in the conservative treatment group. In the total study group, with all severity grades combined, complications were substantially more frequent in the surgical treatment group (23.7%) than in the conservative treatment group (9.1%), with osteoarthritis accounting for the largest difference between groups.

With regard to the conclusions from this meta-analysis and the recommendations for clinical decision-making, the above limitations and the low to very low quality of outcome factors (according to the GRADE system) need to be taken into account. Thus, short-term benefits, including faster return to function, seem to be superior following surgical treatment, but these benefits are transient and there is an increased risk for osteoarthritis with surgical treatment, as was concluded in two of the primary studies[18,32].

Strengths and Limitations

The strengths of this meta-analysis include a comprehensive literature search for reports in any language, the inclusion of only randomized clinical trials, duplicate data extraction, and duplicate assessment of quality of evidence with use of the GRADE system. Furthermore, all authors were contacted to clarify areas of uncertainty and to provide unreported data, which consisted of unreported standard deviations in most cases. We successfully obtained required raw data from four of the seven groups of authors[13,18,31,33,34], making it possible to pool comparable outcome measurements from more studies and therefore to augment the quantity of evidence.

Nonetheless, this meta-analysis had some limitations, including the heterogeneity of secondary outcome measures and low-quality evidence. With regard to substantial heterogeneity in several secondary outcomes, we were able to identify important sources such as fracture type and displacement, study population, and large physiological variations. Even though heterogeneity was substantial in several secondary outcome measurements, we considered the pooling of comparable outcomes and the identification of sources of heterogeneity to be the best option for meta-analysis. However, conclusions from these substantially heterogeneous outcomes should be interpreted with caution.

The meta-analysis showed low-quality to very-low-quality evidence (according to the GRADE system) for treatment effect with any of the selected outcome measurements and, therefore, we can only give "weak" recommendations with regard to the best treatment option[27]. Required data on our primary outcome measurement, patient-rated functional outcome, was only reported in four of the eight studies. Nevertheless, data were homogeneous and significantly favored surgical treatment. In the GRADE system of rating quality of evidence, the highest quality for outcome is only obtainable if there is no serious methodological limitation. Otherwise, for each limitation, quality of evidence will be downgraded stepwise to moderate, low, or very low.

For randomized controlled trials investigating the effect of surgical and conservative treatment of scaphoid fractures,

some limitations seem to be to a great extent inevitable. First, there is no possibility of blinding patients and physicians in any trial comparing surgical and conservative treatment. Second, as both types of treatment have both important advantages and disadvantages, and the benefits and harms of most outcomes are minimal, confidence intervals of nearly every outcome include possible benefits from both types of treatment. This generates a limitation in the form of imprecision of treatment effect.

Some studies in this meta-analysis met the highest possible methodological quality for a surgical trial. Nonetheless, if one or more studies reporting on the same outcome have limited methodological quality, the GRADE system still scores the overall quality of that outcome as low. High-quality evidence within the GRADE framework is possible but difficult to obtain as all safeguards need to be achieved. However, when determining the strength of recommendation with regard to treatment, quality of evidence is one of the four determining factors. One also needs to consider the balance between the desirable and undesirable consequences, the variability in values and preferences, and costs[27]. Finally, only outcomes that are considered to be critical are the primary factors influencing a given recommendation and should be used to determine the overall quality of evidence supporting the recommendation.

Previous Literature

Traditional arguments in favor of surgical treatment for acute scaphoid fractures include (1) incomplete healing or nonunion in association with conservative treatment[35,42], (2) reduced range of motion following cast immobilization[35,43], (3) muscle weakness and reduced grip strength[35,43] after prolonged immobilization, and (4) osteoarthritis at the time of long-term follow-up[43]. In addition, the socioeconomic argument of delay in return to work or sports is becoming more important in current algorithms to decide between conservative and surgical treatment of suspected scaphoid fractures as well as other fractures. Recently, Modi et al.[44] performed a systematic review of surgical and conservative treatment of scaphoid fractures. The authors concluded that both treatments result in good outcomes and that surgical treatment should be reserved for patients desiring faster return to work or athletics.

Bhandari and Hanson[17] and Yin et al.[16] previously performed meta-analyses on this subject. Both of those meta-analyses indicated that there was no evidence from randomized trials to determine whether surgical treatment was superior to conservative treatment. However, since the publication of those studies, several randomized controlled trials have been conducted and published[13,18,31,34]. Furthermore, our literature research and meta-analysis added methodological rigor as we used guidelines from the QUORUM and GRADE working groups. In contrast to previous meta-analyses, we used functional outcome as a primary end point.

Implications for Future Research

All randomized trials that were included in this meta-analysis involved mainly comparable treatment methods for both groups, although outcome measures and measurement instruments differed to a great extent. Only four[13,18,31,32] of the seven primary trials involved the use of validated instruments for the evaluation of functional outcome. In current algorithms for decision-making, there is a trend to favor patient-rated functional outcome and return to function over more traditional outcomes such as time to union and measurement of radiographic evidence of fracture union[40].

Therefore, future randomized trials should use functional outcomes that are evaluated with use of validated outcome instruments as the primary end point. Substantially heterogeneous outcomes in this systematic review, such as grip strength and range of motion of the injured wrist, should be reported in comparison with the value for uninjured wrist. A large trial of high methodological rigor (e.g., adequate concealment of randomization and blinding of outcome assessors) would be needed to allow conclusions to be drawn regarding the treatment effect on grip strength and range of motion. Finally, we encourage future research groups to report their outcome data with use of means and standard deviations to increase the generalizability of results.

Appendix

(eA) Tables showing the characteristics, outcomes, and methodological quality of the trials included in the meta-analysis are available with the electronic version of this article on our web site at jbjs.org (go to the article citation and click on "Supporting Data"). ■

Geert A. Buijze, MD
Job N. Doornberg, MD, PhD
Department of Orthopaedic Surgery,
Academic Medical Centre,
Meibergdreef 9, 1105 AZ, Amsterdam,
The Netherlands.
E-mail address for G.A. Buijze: g.a.buijze@amc.uva.nl

John S. Ham, MD, PhD
Rudolf W. Poolman, MD, PhD
Department of Orthopaedic Surgery
and Department of Joint Research,
Onze Lieve Vrouwe Gasthuis, Postbus 95500,
1090 HM, Amsterdam, The Netherlands

David Ring, MD, PhD
Orthopaedic Hand and Upper Extremity Service,
Massachusetts General Hospital,
Harvard Medical School, Yawkey Center,
Suite 2100, 55 Fruit Street, Boston, MA 02114

Mohit Bhandari, MD
Hamilton Health Sciences-General Hospital,
1200 Main Street West, Hamilton ON L8N 3Z5,
Canada

References

1. Bhat M, McCarthy M, Davis TR, Oni JA, Dawson S. MRI and plain radiography in the assessment of displaced fractures of the waist of the carpal scaphoid. J Bone Joint Surg Br. 2004;86:705-13.

2. Dawson JS, Martel AL, Davis TR. Scaphoid blood flow and acute fracture healing. A dynamic MRI study with enhancement with gadolinium. J Bone Joint Surg Br. 2001;83:809-14.

3. Gellman H, Caputo RJ, Carter V, Aboulafia A, McKay M. Comparison of short and long thumb-spica casts for non-displaced fractures of the carpal scaphoid. J Bone Joint Surg Am. 1989;71:354-7.

4. Hambidge JE, Desai W, Schranz PJ, Compson JP, Davis TR, Barton NJ. Acute fractures of the scaphoid. Treatment by cast immobilisation with the wrist in flexion or extension? J Bone Joint Surg Br. 1999;81:91-2.

5. Kozin SH. Incidence, mechanism, and natural history of scaphoid fractures. Hand Clin. 2001;17:515-24.

6. Ring D, Jupiter JB, Herndon JH. Acute fractures of the scaphoid. J Am Acad Orthop Surg. 2000;8:225-31.

7. Thorleifsson R, Karlsson J, Sigurjonsson K. Fractures of the scaphoid bone. A follow-up study. Arch Orthop Trauma Surg. 1984;103:96-9.

8. Haisman JM, Rohde RS, Weiland AJ; American Academy of Orthopaedic Surgeons. Acute fractures of the scaphoid. J Bone Joint Surg Am. 2006;88:2750-8.

9. Canadian Orthopaedic Trauma Society. Nonoperative treatment compared with plate fixation of displaced midshaft clavicular fractures. A multicenter, randomized clinical trial. J Bone Joint Surg Am. 2007;89:1-10.

10. Handoll HH, Huntley JS, Madhok R. External fixation versus conservative treatment for distal radial fractures in adults. Cochrane Database Syst Rev. 2007;3:CD006194.

11. Jawa A, McCarty P, Doornberg J, Harris M, Ring D. Extra-articular distal-third diaphyseal fractures of the humerus. A comparison of functional bracing and plate fixation. J Bone Joint Surg Am. 2006;88:2343-7.

12. Doornberg JN, Ring D, Fabian LM, Malhotra L, Zurakowski D, Jupiter JB. Pain dominates measurements of elbow function and health status. J Bone Joint Surg Am. 2005;87:1725-31.

13. McQueen MM, Gelbke MK, Wakefield A, Will EM, Gaebler C. Percutaneous screw fixation versus conservative treatment for fractures of the waist of the scaphoid: a prospective randomised study. J Bone Joint Surg Br. 2008;90:66-71.

14. Bond CD, Shin AY, McBride MT, Dao KD. Percutaneous screw fixation or cast immobilization for nondisplaced scaphoid fractures. J Bone Joint Surg Am. 2001;83:483-8.

15. Bedi A, Jebson PJ, Hayden RJ, Jacobson JA, Martus JE. Internal fixation of acute, nondisplaced scaphoid waist fractures via a limited dorsal approach: an assessment of radiographic and functional outcomes. J Hand Surg Am. 2007;32:326-33.

16. Yin ZG, Zhang JB, Kan SL, Wang P. Treatment of acute scaphoid fractures: systematic review and meta-analysis. Clin Orthop Relat Res. 2007;460:142-51.

17. Bhandari M, Hanson BP. Acute nondisplaced fractures of the scaphoid. J Orthop Trauma. 2004;18:253-5.

18. Vinnars B, Pietreanu M, Bodestedt A, Ekenstam F, Gerdin B. Nonoperative compared with operative treatment of acute scaphoid fractures. A randomized clinical trial. J Bone Joint Surg Am. 2008;90:1176-85.

19. Poolman RW, Abouali JA, Conter HJ, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: why are they different? J Bone Joint Surg Am. 2007;89:1542-52.

20. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet. 1999;354:1896-900.

21. Atroshi I, Gummesson C, Andersson B, Dahlgren E, Johansson A. The Disabilities of the Arm, Shoulder and Hand (DASH) outcome questionnaire: reliability and validity of the Swedish version evaluated in 176 patients. Acta Orthop Scand. 2000;71:613-8.

22. Dias JJ, Bhowal B, Wildin CJ, Thompson JR. Assessing the outcome of disorders of the hand. Is the patient evaluation measure reliable, valid, responsive and without bias? J Bone Joint Surg Br. 2001;83:235-40.

23. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. J Orthop Trauma. 1998;12:577-86.

24. Cooney WP, Linscheid RL, Dobyns JH. Triangular fibrocartilage tears. J Hand Surg Am. 1994;19:143-54.

25. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schünemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S; GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328:1490.

26. Andersson GBJ, Cocchiarella L, editors. Guides to the evaluation of permanent impairment. 5th ed. Chicago: American Medical Association Press; 2000.

27. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ; GRADE Working Group. Going from evidence to recommendations. BMJ. 2008;336:1049-51.

28. Bergman S, Jacobsson LT, Herrström P, Petersson IF. Health status as measured by SF-36 reflects changes and predicts outcome in chronic musculoskeletal pain: a 3-year follow up study in the general population. Pain. 2004;108:115-23.

29. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539-58.

30. Adolfsson L, Lindau T, Arner M. Acutrak screw fixation versus cast immobilisation for undisplaced scaphoid waist fractures. J Hand Surg Br. 2001;26:192-5.

31. Arora R, Gschwentner M, Krappinger D, Lutz M, Blauth M, Gabl M. Fixation of nondisplaced scaphoid fractures: making treatment cost effective. Prospective controlled trial. Arch Orthop Trauma Surg. 2007;127:39-46.

32. Dias JJ, Wildin CJ, Bhowal B, Thompson JR. Should acute scaphoid fractures be fixed? A randomized controlled trial. J Bone Joint Surg Am. 2005;87:2160-8.

33. Saedén B, Törnkvist H, Ponzer S, Höglund M. Fracture of the carpal scaphoid. A prospective, randomised 12-year follow-up comparing operative and conservative treatment. J Bone Joint Surg Br. 2001;83:230-4.

34. Vinnars B, Ekenstam FA, Gerdin B. Comparison of direct and indirect costs of internal fixation and cast treatment in acute scaphoid fractures: a randomized trial involving 52 patients. Acta Orthop. 2007;78:672-9.

35. Herbert TJ, Fisher WE. Management of the fractured scaphoid using a new bone screw. J Bone Joint Surg Br. 1984;66:114-23.

36. Armstrong CA, Oldham JA. A comparison of dominant and non-dominant hand strengths. J Hand Surg Br. 1999;24:421-5.

37. Bohannon RW. Grip strength: a summary of studies comparing dominant and nondominant limb measurements. Percept Mot Skills. 2003;96(3 Pt 1):728-30.

38. Incel NA, Ceceli E, Durukan PB, Erdem HR, Yorgancioglu ZR. Grip strength: effect of hand dominance. Singapore Med J. 2002;43:234-7.

39. Ring D. Mini-incision fixation of nondisplaced scaphoid fracture nonunions. J Hand Surg Am. 2008;33:1121.

40. Dias JJ, Taylor M, Thompson J, Brenkel IJ, Gregg PJ. Radiographic signs of union of scaphoid fractures. An analysis of inter-observer agreement and reproducibility. J Bone Joint Surg Br. 1988;70:299-301.

41. Singh HP, Forward D, Davis TR, Dawson JS, Oni JA, Downing ND. Partial union of acute scaphoid fractures. J Hand Surg Br. 2005;30:440-5.

42. Dias JJ, Brenkel IJ, Finlay DB. Patterns of union in fractures of the waist of the scaphoid. J Bone Joint Surg Br. 1989;71:307-10.

43. Lindström G, Nyström A. Incidence of post-traumatic arthrosis after primary healing of scaphoid fractures: a clinical and radiological study. J Hand Surg Br. 1990;15:11-3.

44. Modi CS, Nancoo T, Powers D, Ho K, Boer R, Turner SM. Operative versus nonoperative treatment of acute undisplaced and minimally displaced scaphoid waist fractures—a systematic review. Injury. 2009;40:268-73.