

Variability in the Assessment of Fracture-Healing in Orthopaedic Trauma Studies

By Luis A. Corrales, MD, Saam Morshed, MD, MPH,
Mohit Bhandari, MD, MSc, FRCSC, and Theodore Miclau III, MD

Investigation performed at the Department of Orthopaedic Surgery, University of California, San Francisco, San Francisco, California; San Francisco General Hospital, San Francisco, California; and the Department of Orthopaedic Surgery, McMaster University, Hamilton, Ontario, Canada

Background: There is a lack of consensus among orthopaedic surgeons in the assessment of fracture-healing. We conducted a systematic review of recent clinical studies of long-bone fracture care that were published in three major orthopaedic journals to identify current definitions of fracture-healing.

Methods: MEDLINE and the computerized databases for *The Journal of Bone and Joint Surgery* (American Volume), *The Journal of Bone and Joint Surgery* (British Volume), and the *Journal of Orthopaedic Trauma* were searched from January 1996 through December 2006 with use of title, abstract, keyword, and medical subject headings. Therapeutic clinical studies of long-bone fractures of the appendicular skeleton in adults in which fracture-healing was assessed were selected. Two reviewers independently identified articles and extracted data. Any disagreement was resolved by consensus. We qualitatively and quantitatively summarized the definition of fracture union and the reliability of the assessment of radiographic fracture-healing.

Results: One hundred and twenty-three studies proved to be eligible. Union was defined on the basis of a combination of clinical and radiographic criteria in 62% of the studies, on the basis of radiographic criteria only in 37%, and on the basis of clinical criteria only in 1%. Twelve different criteria were used to define fracture union clinically, and the most common criterion was the absence of pain or tenderness at the fracture site during weight-bearing. In studies involving the use of plain radiographs, eleven different criteria were used to define fracture union, and the most common criterion was bridging of the fracture site. A quantitative measure of the reliability of the radiographic assessment of fracture union was reported in two studies.

Conclusions: We found a lack of consensus with regard to the definition of fracture-healing in the current orthopaedic literature. Without valid and reliable clinical or radiographic measures of union, the interpretation of fracture care studies remains difficult.

Each year in the United States, there are an estimated six million fractures, and as many as 10% will go on to nonunion¹. Trauma is a major cause of disability in the United States, with fractures being a leading contributor to hospital costs and lost productivity in the workplace². Despite advances in treatment, the determination of when a fracture has healed remains subjective, constituting a dilemma for the orthopaedic surgeon.

The precise assessment of fracture union not only is clinically important but also is fundamental to clinical fracture-

repair research. The effectiveness of both standard and novel fracture-healing therapies is fundamentally based on the demonstration of improved fracture-healing. A lack of valid or reliable measures to assess fracture-healing limits the integrity of such studies. In addition, the use of inconsistent criteria to assess fracture-healing across studies further limits the generalizability of their results.

Fracture-healing assessments have included serial physical and radiographic examinations³, responses to weight-bearing and to palpation⁴, manual assessments^{5,6}, plain radiographic

Disclosure: In support of their research for or preparation of this work, one or more of the authors received, in any one year, outside funding or grants in excess of \$10,000 from a Doris Duke Clinical Research Fellowship, an Orthopaedic Research and Education Foundation (OREF) Clinical Research Training Fellowship, and the National Institutes of Health. Neither they nor a member of their immediate families received payments or other benefits or a commitment or agreement to provide such benefits from a commercial entity. No commercial entity paid or directed, or agreed to pay or direct, any benefits to any research fund, foundation, division, center, clinical practice, or other charitable or nonprofit organization with which the authors, or a member of their immediate families, are affiliated or associated.

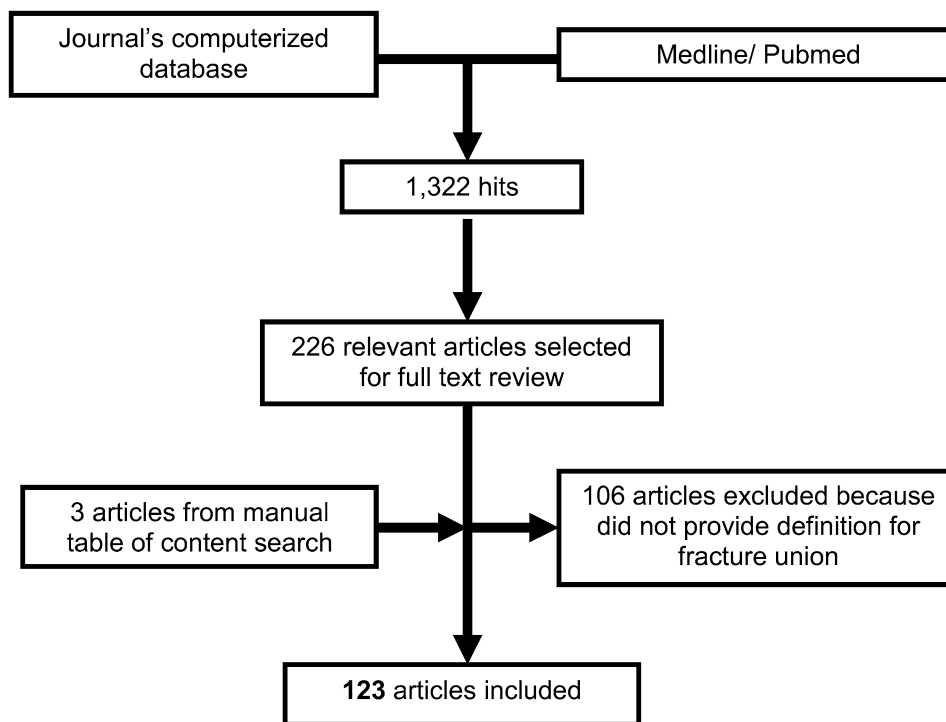


Fig. 1
Flow diagram detailing literature search.

measures (such as cortical continuity, the presence of a fracture line, and callus formation)⁷, and health-related quality-of-life measures.

We conducted a systematic review of contemporary published clinical outcomes studies pertaining to long-bone fracture-healing in order to document (1) the current definitions being used for fracture union of long bones in the orthopaedic trauma literature and (2) quantitative or qualitative reliability reporting for the radiographic assessment of fracture-healing.

Materials and Methods

Study Eligibility

Our objective was to identify therapeutic clinical studies of long-bone fractures of the appendicular skeleton (specifically, the femur, tibia, fibula, humerus, radius, and ulna) in adults in which fracture-healing was an outcome. Several different types of studies were excluded: (1) studies specifically investigating pathologic fractures, (2) cadaver studies, (3) animal studies, (4) case reports, and (5) studies of diagnostic accuracy unless a fracture-repair intervention was also being assessed. Two reviewers (S.M. and L.A.C.) independently assessed studies for inclusion eligibility. When an article was judged by a reviewer to be potentially eligible on the basis of the title or abstract, the full text of the article was reviewed, and any disagreement between the reviewers was resolved by consensus.

Search Strategy

We conducted a thorough search of MEDLINE and the computerized electronic journal databases for articles published in

three major orthopaedic journals: *The Journal of Bone and Joint Surgery* (American Volume), *The Journal of Bone and Joint Surgery* (British Volume), and the *Journal of Orthopaedic Trauma* from January 1996 through December 2006. We used keyword, title, abstract and medical subject headings in our search parameters. Finally, a manual table-of-contents search was performed to seek additional articles.

Data Abstraction

Two reviewers (S.M. and L.A.C.) identified articles and extracted data, in parallel. The main information obtained from the articles included the definition used for fracture union and the reporting of the reliability of the assessment of radiographic fracture-healing. The type of fracture care (primary fracture care, nonunion fracture care, or periprosthetic fracture care), the distribution of fractures, the outcome assessor (i.e., orthopaedic surgeon, radiologist, or resident), and outcome assessor blinding were also recorded.

Study Quality Assessments

In studies in which blinding of the outcome assessor was not reported, we determined if blinding would have been feasible on the basis of study characteristics. For the clinical assessment of fracture-healing, blinding of the observer was considered to be possible unless hardware (such as an external fixator) was visible or was otherwise detectable on physical examination. For the radiographic assessment of fracture-healing, blinding was considered to be possible unless the treatment being studied included the use of a prosthesis.

The level of evidence for all included studies was determined independently by each reviewer on the basis of the level-of-evidence rating system⁸ for articles in which the level of evidence was not specified within the published report.

Data Analysis

Data that were extracted from the studies were summarized in detailed tables with use of Microsoft Word (Microsoft, Redmond, Washington), and a database was constructed with use of Excel (Microsoft). Inter-rater reliability was calculated, adjusting for chance agreement with use of a weighted kappa statistic. A kappa of 1 indicates perfect agreement between observers, and a kappa of 0 indicates agreement equivalent to chance. Disagreements were resolved by consensus for the purposes of reporting our findings.

Results

Literature Search

The literature search yielded 226 articles for a full text review. A total of 123 studies were included in the analysis. One hundred and six studies were excluded because they provided no definition of fracture-healing, and then three additional studies were included from the manual table-of-contents search (Fig. 1). Overall, union was defined on the basis of a combination of clinical and radiographic criteria in 62% of the articles, radiographic criteria only in 37%, and clinical criteria only in 1%.

Study Characteristics

Seventy percent of the articles were related to primary fracture care, 20% were related to nonunion care, and 10% were related to periprosthetic fracture care. Forty-six percent of the articles involved fractures of the femur, 34% involved fractures of the tibia and/or fibula, 15% involved fractures of the humerus, and 9% involved fractures of the radius and/or ulna (some studies fit into more than one category). With regard to level of evidence, 5% of the studies were level I, 19% were level II, 12% were level III, and 64% were level IV. The interobserver reliability for the level of evidence classification of 107 studies was high (weighted kappa = 0.90).

Clinical Definitions of Fracture-Healing

Seventy-seven studies involved the use of clinical criteria to define fracture union. The most common clinical criteria were the absence of pain or tenderness during weight-bearing (49%), the absence of pain or tenderness on palpation or physical examination (39%), and the ability to bear weight (18%) (Table I). Clinical criteria were based on verbatim wording from the included articles, and, although there are similarities between criteria, there also are important qualitative differences such as the assessment of pain and function that differentiate them. Table II shows the most commonly reported criteria for the clinical assessment of fracture union according to the location of the fracture. None of the seventy-seven studies involved blinding of the clinical assessment of union, and yet we de-

TABLE I Criteria Used to Define Clinical Fracture Union*

Clinical Criteria Used to Define Fracture Union	Number of Articles (N = 77)
1. No pain/tenderness when bearing weight	38 (49%)
2. No pain/tenderness on palpation/examination	30 (39%)
3. Ability to bear weight	14 (18%)
4. Ability to walk/perform activities of daily living with no pain	11 (14%)
5. Ability to walk/perform activities of daily living	9 (12%)
6. No residual pain at fracture site	8 (10%)
7. No motion at fracture site on examination	4 (5%)
8. Full range of motion at adjacent joint	4 (5%)
9. "Clinically stable/asymptomatic"	2 (3%)
10. No residual warmth at fracture site	1 (1%)
11. Full range of motion at adjacent joint without pain	1 (1%)
12. Fracture stiffness measured mechanically†	1 (1%)

*The clinical criteria were grouped into twelve similar categories and were arranged in order of most to least common use. †A fracture stiffness of >15 Nm/deg in two orthogonal planes was reported to indicate sufficient healing for external fixator removal in the case of a tibial fracture⁹.

termined that it would have been possible to do so in 95% of those studies. Only one study used a strictly clinical definition for fracture union⁹; in that study, mechanical parameters were used to measure the stiffness of tibial shaft fractures that had been treated with external fixation.

Radiographic Definitions of Fracture-Healing

One hundred and twenty-two articles used radiographic criteria to define fracture union. Three radiographic modalities were used: plain radiographs (employed in 98% of the studies), computed tomography (1%), and ultrasound (1%). The most common radiographic definitions of fracture-healing in studies involving the use of plain radiographs were bridging of the fracture site by callus, trabeculae, or bone (53%); bridging of the fracture site at three cortices (27%); and obliteration of the fracture line or cortical continuity (18%) (Table III). Table II shows the most commonly reported criteria for radiographic assessment of fracture union according to the location of the fracture.

Two studies did not involve the use of plain radiographs to assess fracture-healing. In the study in which computed tomography was used, union was defined as bridging of >25% of the cross-sectional area at the fracture site¹⁰. In the study in which ultrasound was used, union was defined as the complete

TABLE II Most Common Criteria for Assessing Fracture Union According to Location of Fracture Based on Articles Reviewed*

	Femur	Tibia	Humerus	Forearm	Radius
Proximal fracture					
Radiographic assessment	Bridging of fracture site	Obliteration of fracture line	Obliteration of fracture line or bridging of fracture site	—	—
Clinical assessment	No pain during weight-bearing	Ability to bear weight	Ability to perform activities of daily living without pain, or no residual pain at fracture site, or full range of motion at adjacent joint	—	—
Shaft fracture					
Radiographic assessment	Bridging of fracture site	Bridging of fracture site	Bridging of fracture site	Obliteration of fracture line	—
Clinical assessment	No pain during weight-bearing	No pain during weight-bearing	No pain on palpation/examination	No pain on palpation/examination	—
Distal fracture					
Radiographic assessment	Bridging of fracture site	Bridging of fracture site	Bridging of fracture site	—	Obliteration of fracture line or calcification of callus
Clinical assessment	No pain during weight-bearing	No pain during weight-bearing or ability to bear weight	No pain on palpation/examination	—	No residual pain at fracture site

*The most common radiographic criteria (based on plain radiographs) and clinical definitions for fracture union are given. If more than one definition was used equally for a given region, then all are listed. A dash indicates that none of the reviewed studies provided a definition for fracture union for that body region.

disappearance of the intramedullary nail on ultrasound imaging at six weeks or progressive disappearance of the intramedullary nail with the formation of periosteal callus between six and nine weeks following treatment¹¹.

Blinding of the observers assessing radiographic union was not reported in 93% of the 122 articles reviewed. We determined that blinding would have been feasible in only thirteen studies. Eight articles indicated whether or not blinding had taken place, and, of these, seven indicated that it had.

Seventy-four percent of the 122 studies did not indicate who functioned as the observer during the assessment of radiographic union; in the remaining studies, the reported assessors were orthopaedic surgeons (19%), orthopaedic surgeons and radiologists (4%), radiologists (2%), and orthopaedic surgeons and residents in training (1%). In three of the five studies in which both orthopaedic surgeons and radiologists were used as assessors, there were reported differences in the union rate and the time to union, with radiologists consistently giving more conservative measurements. Kristiansen et al. reported an average difference of eleven days in the time to union between surgeon and radiologist¹², and Emami et al. reported an average difference of twenty-two days¹³, with radiologists declaring union later in both studies. Jones et al. reported a difference between radiologists and orthopaedic surgeons with regard to

the union rate in the treatment group at twelve months, with the radiologists determining a rate of 75% to 78% and the orthopaedic surgeons determining a rate of 90% to 92%¹⁴.

Reliability Reporting

A quantitative measure of the reliability of the assessment of radiographic fracture union was reported in only two of the studies that were reviewed. Bhandari et al. evaluated chance-adjusted agreement for cortical continuity at two cortices and reported a kappa of 0.75¹⁵. Govender et al. reported that the surgeons' assessments of fracture-healing agreed with radiologists' assessments in 91% of cases¹⁶. Eleven studies qualitatively confirmed the reliability of the assessment of radiographic union either by using an independent panel of radiologists to corroborate findings or by achieving assessor consensus^{12-15, 17-23}.

Discussion

We performed a systematic review of studies of fracture repair and found variability in the definition of fracture union. The majority of studies involved the use of both radiographic and clinical parameters to define fracture-healing. The observers who assessed fracture-healing radiographically, the blinding of outcome assessment, and the reliability of the radiographic assessment of fracture union were not commonly reported.

TABLE III Criteria Used to Define Radiographic Fracture Union on Plain Radiographs*

Radiographic Criteria Used to Define Fracture Union	Number of Articles (N = 120)
1. Bridging of fracture by callus/bone/trabeculae or osseous bone	63 (53%)
2. Bridging of fracture by callus, trabeculae, or bone	32 (27%)
3. Obliteration of fracture line/cortical continuity	22 (18%)
4. Bridging of fracture at 2 cortices	7 (6%)
5. Bridging of fracture at 1 cortex	4 (3%)
6. Absence of fracture displacement†	4 (3%)
7. Absence of hardware failure/loosening	3 (2.5%)
8. Absence of osteonecrosis	2 (1.7%)
9. Calcification of callus	2 (1.7%)
10. Bridging of fracture at 4 cortices	1 (1%)
11. Presence of callus	1 (1%)

*The radiographic criteria were grouped into eleven similar categories and were arranged in order of most to least common use.
†Assessed at one to three months following treatment.

The array of definitions for fracture union found in the current orthopaedic trauma literature reflects the lack of consensus that exists in the orthopaedic community³. The field of orthopaedic surgery has strived to standardize and improve outcomes assessment²⁴. However, there is a need for the standardization of guidelines used for the assessment of fracture union. The ability to accurately and reliably measure outcomes is fundamental to proper study design, reduces bias, and benefits patient care.

Fracture-healing is a dynamic process and not a single event, but without standardized assessment, comparing outcomes between studies must be done with caution. For example, in the present review, two randomized prospective trials comparing antegrade with retrograde reamed intramedullary fixation of femoral shaft fractures demonstrated different results^{25,26}. One study demonstrated a faster healing time in association with antegrade nail fixation²⁵, whereas the second study found no difference²⁶. The definition of fracture union differed both clinically and radiographically between these two studies, and, while there were similarities, the parameters used to define union could be interpreted differently. Therefore, in studies of patients with a similar injury receiving similar treatment, reported differences in time to healing could be due to differing definitions of fracture union and not to the natural history of the disease or treatment.

Plain radiography is the most common way in which fracture union is assessed, and a substantial number of studies defined fracture union on the basis of radiographic parameters alone. Hammer et al. combined cortical continuity, the loss of a

visible fracture line, and callus size in a scale to assess fracture-healing radiographically but found conventional radiographic examination difficult to correlate with fracture stability and could not conclusively determine the state of union⁷. In animal models, cortical continuity is a good predictor of fracture torsional strength, whereas callus area is not²⁷. Also, clinicians cannot reliably determine the strength of a healing fracture on the basis of a single set of radiographs²⁸ and are unable to rank radiographs of healing fractures in order of strength²⁹. Therefore, we rely heavily on a radiographic method without proven validity for predicting bone strength in the assessment of fracture union.

There is very little information in the literature on the reliability of commonly used radiographic measures for fracture union, and the vast majority of studies do not address the reliability of the methods used. Interobserver chance-corrected agreement for radiographic assessment of fracture-healing following intramedullary fixation of tibial shaft fractures is low³⁰, and a surgeon's ability to rank fracture union using radiographs chronologically following internal fixation is approximately 70%³¹.

The type of outcome assessor who interprets the radiographic findings during fracture-healing is important and can contribute to bias. We found that radiologists were more conservative in determining fracture-healing than were orthopaedic surgeons. This difference is important because the type of assessor was not reported in many of the studies that we reviewed, and thus differences reported between studies may be due to assessor bias and not to a treatment effect.

The blinding of outcome assessment is necessary in order to accurately measure the effect of a treatment but it can be difficult in nonpharmacologic treatment trials³² and was not routinely reported in the studies that we reviewed, even when potentially feasible. For studies assessing implant choice, complete blinding may require manipulation of radiographs, whereas for studies assessing therapies not involving an implant or a surgical approach, blinding of an assessor may be accomplished with more conventional means. Unblinded assessment of outcomes, both continuous and dichotomous, is associated with a potential for exaggeration of the effectiveness of treatment³³. The lack of reporting that we found is not unusual and may not mean that bias-reducing procedures did not indeed occur³⁴. Nonetheless, care needs to be taken in blinding assessors, when possible, as this can affect the estimated efficacy of treatment of long-bone fractures.

Future efforts will need to address how to standardize the assessment of fracture union and will need to explore new methods that are quantifiable and reliable in order to improve research and clinical practice. Dichotomizing fracture-healing is not possible without introducing bias or losing information about this continuous biological process. Still, study results need to convey clinically recognizable and relevant outcomes that are consistently reported by injury type. Advances in imaging and molecular biology ultimately may give us the ability to follow the healing process over time with a quantifiable measure that accurately conveys the probability of union. Moreover, a

patient's input must not be ignored or marginalized when determining fracture union. We have few validated quality-of-life instruments tailored for patients with long-bone fractures. Instruments that measure patient-important outcomes, or factors that matter to the subject undergoing fracture repair, have to be included in the development of a comprehensive definition of healing if such a definition is to be widely applicable. A composite definition that includes improved radiographic modalities, clinical examination, and patient-important outcomes will provide a better measurement of fracture repair in the future. Until this is accomplished, we must be aware of the limitations of the methods that we currently use.

The present systematic review had several possible limitations. One such limitation is the possibility of selection bias and the obvious omission of comparable studies not published in the three journals that we searched. While there is always the possibility of selection bias, this risk was minimized with a comprehensive literature search involving the use of multiple electronic databases and a manual table-of-contents search, both performed in duplicate. Although our review was limited to reports that had been published in three journals, we believed that these three journals have the highest impact among those regularly publishing results of fracture care and that they provide an adequate and representative sample of studies to reflect current practice in defining fracture union.

In conclusion, we found a wide variety in, and a lack of consensus about, the definition of fracture union in the current orthopaedic trauma literature. Without clear guidelines to assess fracture-healing, and with our reliance on unreliable

measurements of fracture union, there is an obvious potential for bias. It is important to use caution when comparing studies of fracture treatment that assess union and to be aware of the potential for bias resulting from these imperfect methods of measurement. Increasing the use of blinded assessment of outcomes and improved reporting of reliability of subjective end points will improve the quality of inferences derived from clinical studies. In the future, there will be a need to incorporate objective, quality-of-life, and functional parameters into the development of a more standardized definition for fracture union that better characterizes the chronological process of healing. ■

Luis A. Corrales, MD

Department of Orthopaedic Surgery, University of California at San Francisco School of Medicine, 500 Parnassus Avenue, MU-320W, San Francisco, CA 94143

Mohit Bhandari, MD, MSc, FRCSC

Hamilton General Hospital, 7 North, Suite 727, 237 Barton Street East, Hamilton, ON L8L 2X2, Canada

Saam Morshed, MD, MPH

Theodore Miclau III, MD

Department of Orthopaedic Surgery, University of California at San Francisco, San Francisco General Hospital, 1001 Potrero Avenue, Room 3A-36, San Francisco, CA 94110. E-mail address for T. Miclau III: miclaut@orthosurg.ucsf.edu

References

- Einhorn TA. Enhancement of fracture-healing. *J Bone Joint Surg Am.* 1995;77:940-56.
- Finkelstein EA, Corso PS, Miller T. The incidence and economic burden of injuries in the United States. New York: Oxford University Press; 2006.
- Bhandari M, Guyatt GH, Swiontkowski MF, Tornetta P 3rd, Sprague S, Schemitsch EH. A lack of consensus in the assessment of fracture healing among orthopaedic surgeons. *J Orthop Trauma.* 2002;16:562-6.
- Sarmiento A, Sharpe FE, Ebramzadeh E, Normand P, Shankwiler J. Factors influencing the outcome of closed tibial fractures treated with functional bracing. *Clin Orthop Relat Res.* 1995;315:8-24.
- Matthews LS, Kaufer H, Sonstegard DA. Manual sensing of fracture stability: a biomechanical study. *Acta Orthop Scand.* 1974;45:373-81.
- Webb J, Herling G, Gardner T, Kenwright J, Simpson AH. Manual assessment of fracture stiffness. *Injury.* 1996;27:319-20.
- Hammer RR, Hammerby S, Lindholm B. Accuracy of radiologic assessment of tibial shaft fracture union in humans. *Clin Orthop Relat Res.* 1985;199:233-8.
- Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am.* 2003;85:1-3.
- Wade RH, Moorcroft CI, Thomas PBM. Fracture stiffness as a guide to the management of tibial fractures. *J Bone Joint Surg Br.* 2001;83:533-5.
- Brinker MR, O'Connor DP. Ilizarov compression over a nail for aseptic femoral nonunions that have failed exchange nailing: a report of five cases. *J Orthop Trauma.* 2003;17:668-76.
- Moed BR, Subramanian S, van Holsbeeck M, Watson JT, Cramer KE, Karges DE, Craig JG, Bouffard JA. Ultrasound for the early diagnosis of tibial fracture healing after static interlocked nailing without reaming: clinical results. *J Orthop Trauma.* 1998;12:206-13.
- Kristiansen TK, Ryaby JP, McCabe J, Frey JJ, Roe LR. Accelerated healing of distal radial fractures with the use of specific, low-intensity ultrasound. A multicenter, prospective, randomized, double-blind, placebo-controlled study. *J Bone Joint Surg Am.* 1997;79:961-73.
- Emami A, Petré-Mallmin M, Larsson S. No effect of low-intensity ultrasound on healing time of intramedullary fixed tibial fractures. *J Orthop Trauma.* 1999;13:252-7.
- Jones AL, Bucholz RW, Bosse MJ, Mirza SK, Lyon TR, Webb LX, Pollak AN, Golden JD, Valentin-Opran A; BMP-2 Evaluation in Surgery for Tibial Trauma-Allograft (BESTT-ALL) Study Group. Recombinant human BMP-2 and allograft compared with autogenous bone graft for reconstruction of diaphyseal tibial fractures with cortical defects. A randomized, controlled trial. *J Bone Joint Surg Am.* 2006; 88:1431-41.
- Bhandari M, Tornetta P 3rd, Sprague S, Najibi S, Petrisor B, Griffith L, Guyatt GH. Predictors of reoperation following operative management of fractures of the tibial shaft. *J Orthop Trauma.* 2003;17:353-61.
- Govender S, Csimma C, Genant HK, Valentin-Opran A, Amit Y, Arbel R, Aro H, Atar D, Bishay M, Börner MG, Chiron P, Choong P, Cinats J, Courtenay B, Feibel R, Geulette B, Gravel C, Haas N, Raschke M, Hammacher E, van der Velde D, Hardy P, Holt M, Josten C, Ketterl RL, Lindeque B, Lob G, Mathevon H, McCoy G, Marsh D, Miller R, Munting E, Oevre S, Nordsletten L, Patel A, Pohl A, Rennie W, Reynders P, Rommens PM, Rondia J, Rossouw WC, Daneel PJ, Ruff S, Rüter A, Santavirta S, Schildhauer TA, Gekle C, Schnettler R, Segal D, Seiler H, Snowdowne RB, Stapert J, Taglang G, Verdonk R, Vogels L, Weckbach A, Wentzensen A, Wisniewski T; BMP-2 Evaluation in Surgery for Tibial Trauma (BESTT) Study Group. Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: a prospective, controlled, randomized study of four hundred and fifty patients. *J Bone Joint Surg Am.* 2002;84:2123-34.
- Friedlaender GE, Perry CR, Cole JD, Cook SD, Cierny G, Muschler GF, Zych GA, Calhoun JH, LaForte AJ, Yin S. Osteogenic protein-1 (bone morphogenetic protein-7)

in the treatment of tibial nonunions. *J Bone Joint Surg Am.* 2001;83 Suppl 1(Pt 2):S151-8.

18. Henley MB, Chapman JR, Agel J, Harvey EJ, Whorton AM, Swiontkowski MF. Treatment of type II, IIIA, and IIIB open fractures of the tibial shaft: a prospective comparison of unreamed interlocking intramedullary nails and half-pin external fixators. *J Orthop Trauma.* 1998;12:1-7.

19. Kreder HJ, Agel J, McKee MD, Schemitsch EH, Stephen D, Hanel DP. A randomized, controlled trial of distal radius fractures with metaphyseal displacement but without joint incongruity: closed reduction and casting versus closed reduction, spanning external fixation, and optional percutaneous K-wires. *J Orthop Trauma.* 2006;20:115-21.

20. Nork SE, Barei DP, Schildhauer TA, Agel J, Holt SK, Schrick JL, Sangeorzan BJ. Intramedullary nailing of proximal quarter tibial fractures. *J Orthop Trauma.* 2006;20:523-8.

21. Stannard JP, Harris HW, McGwin G Jr, Volgas DA, Alonso JE. Intramedullary nailing of humeral shaft fractures with a locking flexible nail. *J Bone Joint Surg Am.* 2003;85:2103-10.

22. Stannard JP, Wilson TC, Volgas DA, Alonso JE. The less invasive stabilization system in the treatment of complex fractures of the tibial plateau: short-term results. *J Orthop Trauma.* 2004;18:552-8.

23. Tang P, Gates C, Hawes J, Vogt M, Prayson MJ. Does open reduction increase the chance of infection during intramedullary nailing of closed tibial shaft fractures? *J Orthop Trauma.* 2006;20:317-22.

24. Swiontkowski MF, Buckwalter JA, Keller RB, Haralson R. The outcomes movement in orthopaedic surgery: where we are and where we should go. *J Bone Joint Surg Am.* 1999;81:732-40.

25. Ostrum RF, Agarwal A, Lakatos R, Poka A. Prospective comparison of retrograde and antegrade femoral intramedullary nailing. *J Orthop Trauma.* 2000;14:496-501.

26. Tornetta P 3rd, Tiburzi D. Antegrade or retrograde reamed femoral nailing. A prospective, randomised trial. *J Bone Joint Surg Br.* 2000;82:652-4.

27. Panjabi MM, Walter SD, Karuda M, White AA, Lawson JP. Correlations of radiographic analysis of healing fractures with strength: a statistical analysis of experimental osteotomies. *J Orthop Res.* 1985;3:212-8.

28. Panjabi MM, Lindsey RW, Walter SD, White AA 3rd. The clinician's ability to evaluate the strength of healing fractures from plain radiographs. *J Orthop Trauma.* 1989;3:29-32.

29. Nicholls PJ, Berg E, Bliven FE Jr, Kling JM. X-ray diagnosis of healing fractures in rabbits. *Clin Orthop Relat Res.* 1979;142:234-6.

30. Whelan DB, Bhandari M, McKee MD, Guyatt GH, Kreder HJ, Stephen D, Schemitsch EH. Interobserver and intraobserver variation in the assessment of the healing of tibial fractures after intramedullary fixation. *J Bone Joint Surg Br.* 2002;84:15-8.

31. Davis BJ, Roberts PJ, Moorcroft CI, Brown MF, Thomas PB, Wade RH. Reliability of radiographs in defining union of internally fixed fractures. *Injury.* 2004;35:557-61.

32. Boutron I, Tubach F, Giraudeau B, Ravaud P. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *J Clin Epidemiol.* 2004;57:543-50.

33. Poolman RW, Struijs PA, Krips R, Sierevelt IN, Marti RK, Farrokhyar F, Bhandari M. Reporting of outcomes in orthopaedic randomized trials: does blinding of outcome assessors matter? *J Bone Joint Surg Am.* 2007;89:550-8.

34. Devereaux PJ, Choi PT, El-Dika S, Bhandari M, Montori VM, Schünemann HJ, Garg AX, Busse JW, Heels-Ansdell D, Ghali WA, Manns BJ, Guyatt GH. An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. *J Clin Epidemiol.* 2004;57:1232-6.