

## ORIGINAL ARTICLE

# The Investigation of Suspected Fracture— a Comparison of Ultrasound With Conventional Imaging

Systematic Review and Meta-analysis

Gordian Lukas Schmid, Stefan Lippmann, Susanne Unverzagt, Christiane Hofmann, Tobias Deutsch, Thomas Frese

## SUMMARY

**Background:** Ultrasound imaging can be used to diagnose fractures in patients with acute trauma. Its main advantages over conventional imaging are the absence of radiation exposure and its greater availability.

**Methods:** A systematic search in electronic databases (Medline, Embase, Cochrane CENTRAL) was supplemented by a manual search on the Internet and in the reference lists of pertinent publications. The QUADAS-2 instrument was used to assess the quality of the individual studies retrieved. In the meta-analysis, the sensitivity and specificity of the individual studies were pooled.

**Results:** The available information on the diagnostic accuracy of ultrasound in the detection of fractures compared with that of conventional imaging (x-ray, CT, MRI) in patients with acute non-life-threatening trauma is summarized. The database search yielded 2153 hits, among which there were 48 studies that were suitable for inclusion in this review. The pooled sensitivity and specificity were 0.91 (95% confidence interval [0.90; 0.92]) and 0.94 [0.93; 0.95], although the analyzed studies were markedly heterogeneous ( $I^2$ : sensitivity 74%, specificity 81%). The sensitivity of ultrasound was higher for the detection of fractures of the humerus, the forearm, the ankle, and the long bones in general, as well as fractures in children, and lower for fractures of the short bones of the hands and feet, and in adults.

**Conclusion:** Strong evidence supports the use of ultrasound imaging for certain indications in the detection of fractures.

### ► Zitierweise

Schmid GL, Lippmann S, Unverzagt S, Hofmann C, Deutsch T, Frese T: The investigation of suspected fracture—a comparison of ultrasound with conventional imaging: systematic review and meta-analysis. *Dtsch Arztebl Int* 2017; 114: 757–64. DOI: 10.3238/arztebl.2017.0757

**S**uspected fracture following trauma is a common reason for emergency room admission. X-ray is usually used to diagnose patients with suspected fracture and is the standard procedure. The use of ultrasound to diagnose fractures has been the subject of intense discussion over the last 20 years (1–4). Viewing alterations to bone surfaces (discontinuities, displacement, subperiosteal hematomas) has been proven to be a reliable, manageable method of fracture diagnosis. One of the main advantages of ultrasound examination is the avoidance of radiation exposure; this is particularly true in children, who are more sensitive to radiation than adults (5). Ultrasound diagnosis is also easy to teach (6, e1), cheap (2), and available in emergency rooms and many medical practices. All these are arguments in favor of the use of ultrasound in diagnostic algorithms for suspected fracture, which could reduce the number of x-ray examinations performed (7, 8).

The main aim of this systematic review is to summarize the available evidence regarding the diagnostic accuracy (sensitivity and specificity) of ultrasound versus the standard imaging procedures used to diagnose acute fractures (x-ray, CT, MRI). Subgroup analyses were also performed, in order to determine accuracy for various fracture locations and age groups. This should provide a basis for discussion of the potential inclusion of ultrasound in the diagnostic algorithms of medical practices and emergency rooms and support targeted research in the future.

## Methods

### Research strategy

This systematic review was planned and performed in line with the PRISMA (preferred reporting items for systematic reviews and meta-analyses) (e2) and MOOSE (meta-analysis of observational studies in epidemiology) (e3) recommendations. The search strategy was developed by 2 physicians (GLS, TF) and one librarian (CH).

Department of General Practice and Family Medicine, Medical Faculty of the University of Leipzig: Dr. med. Schmid, Dr. rer. pol. Lippmann, Dr. rer. med. Deutsch

Institute of General Practice and Family Medicine, Martin-Luther-University Halle-Wittenberg, Halle (Saale): Dr. med. Schmid, Prof. Dr. med. Frese

Institute for Medical Epidemiology, Biometrics and Computer Science, Martin-Luther-University Halle-Wittenberg, Halle (Saale): PD Dr. rer. nat. habil. Unverzagt

Library of Medicine, University of Leipzig: Hofmann

**THE CLINICAL PERSPECTIVE**

Ultrasound might be useful as part of diagnostic algorithms, to detect or rule out fractures in patients with traumatic injuries. A final decision about appropriate treatment probably requires other information that cannot always be obtained via ultrasound and must instead be obtained via standard imaging procedures (x-ray, CT, MRI). For example, Engin et al. (9) describe a good ultrasound detection rate for sternum fractures but nevertheless find ultrasound to be inappropriate for evaluating the extent of dislocation. In contrast, Eckert et al. (24) recommend ultrasound as a viable alternative for diagnosing pediatric forearm fractures and report that it can satisfactorily identify dislocations and angulation of fractures, in addition to its high diagnostic value. Ackermann et al. (7, 18) report that ultrasound can reliably be used to evaluate axis deviation in pediatric fractures of the forearm and proximal humerus; Yesilaras et al. (32) report the same finding for metacarpal fractures. According to Kozaci et al., treatment decisions for forearm and metacarpal fractures made on the basis of ultrasound examinations alone match those made on the basis of x-ray images (34, 35). Even ultrasound-guided reduction of displaced fractures was found to be successful in 92% of pediatric forearm fractures (15) and 91% of pediatric long-bone fractures (17).

Although our findings are promising, further research must be conducted before ultrasound is incorporated into clinical diagnostic algorithms for fractures, in order to investigate its performance, safety, and cost-efficiency thoroughly.

The electronic search of the literature was performed in 3 databases (Medline, EMBASE, Cochrane CENTRAL) on March 19, 2015 (last update: September 2016). In order to take account of technical advances in ultrasound machines, only publications dating from 2000 onwards were included. The PICO Framework (e4) was used to define inclusion and exclusion criteria (*eMethods 1*). The full study protocol is provided in *eMethods 2* and was not published in advance. The methodological quality of the included studies was evaluated using QUADAS-2 (QUADAS: quality assessment of diagnostic accuracy studies) (e5) (*eMethods 3*).

**Meta-analysis and statistics**

Percent agreement and Cohen’s kappa for the 2 raters (GLS, TF) were calculated to estimate the interrater reliability of the initial screening process. Data from the included studies was pooled using a random-effects model. Heterogeneity was evaluated by means of the chi-square test and  $I^2$ , using the program MetaDiSc (version 1.4, Hospital Universitario Ramón y Cajal, Madrid, Spain) (e6).

We performed subgroup analyses for various fracture locations (humerus, forearm, ankle and foot, all long bones, short bones of the hands and feet), age groups (adults versus children and adolescents), and levels of risk of bias (low/unclear risk versus high risk) described in the studies.

Additional, detailed information on the methods can be found in *eMethods 1*.

**Results**

**Literature search**

The flow chart (*Figure 1*) illustrates the search strategy and gives the number of evaluated publications and reasons for exclusion at each step.

Substantial interrater reliability was found for the screening processes of both titles (percent agreement = 91.2%, kappa = 0.622) and abstracts (percent agreement = 89.0%, kappa = 0.685).

The 2 raters considered 78 and 94 publications respectively to be potentially relevant. In all cases in which the raters’ evaluation of an abstract differed ( $n = 17$ ), discussion with a third scientist (TD) led to consensus.

At the fulltext phase, 108 studies were analyzed using a data extraction form designed and piloted by the authors (*eMethods 2*) and either included or excluded. The 48 articles included in the review (6, 7, 9–40, e7–e20) are summarized in *eTable 1*.

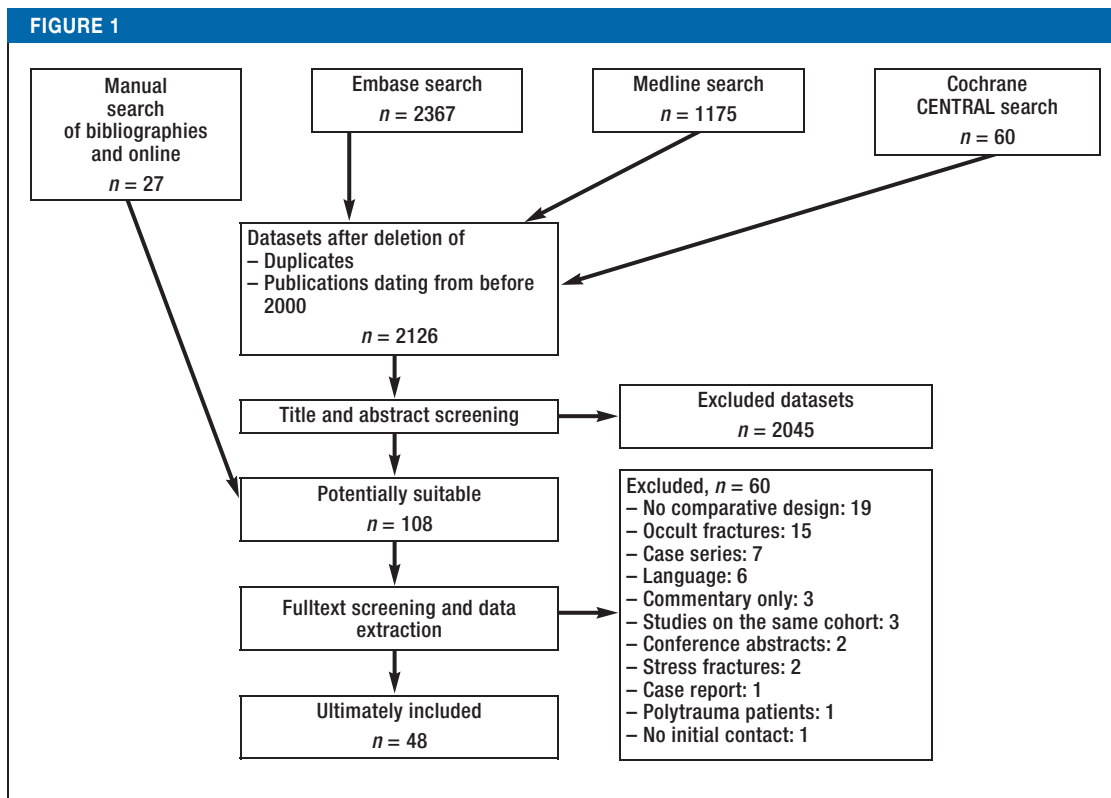
**Characteristics of the selected studies**

With the exception of one publication (16), all the studies were conducted in emergency rooms. Thirty-four studies compared ultrasound findings with conventional x-ray images, 3 used CT or MRI as the reference standard, 4 combined x-ray and CT, and 7 used clinical consensus or follow-up examinations to make a correct diagnosis.

The 48 included studies covered a total of 4427 patients and 5436 examinations. They report on the sensitivity and specificity of ultrasound compared to the respective reference imaging method used (*Figures 2 and 3*). Pooled sensitivity was 0.91 (95% confidence interval: [0.90; 0.92]), and pooled specificity 0.94 [0.93; 0.95]. In the pooled population of all included studies, radiation exposure would have been avoided for 56 of 100 ultrasound examinations (true negatives) (median: 48%, range: 7 to 84%), but 4 of the 100 examinations would have led to a false negative (median: 2%, range: 0 to 29%).

In 15 studies findings were calculated on the basis of the number of investigated bones rather than that of injured patients. Excluding these studies, sensitivity was the same and specificity was slightly lower, at 0.92 (0.90; 0.93). This indicates that these 15 studies slightly overestimated the accuracy of ultrasound.

Nine publications reported on pain caused by ultrasound examination. In 5 of these there were no differences between ultrasound and x-ray (15, 19–21, e12), while 4 found that ultrasound diagnosis was less



painful because a pain-relieving position is permitted and the ultrasound gel is cooling (18, 36, e15, e16).

The quality of the included studies improved over the last 5 years, as illustrated by a lower risk of bias and more accurate description of study design (eMethods 3). On the basis of the funnel plots shown in eMethods 4, the risk of significant publication bias was estimated as low.

The  $I^2$  values calculated to estimate the heterogeneity of the studies included in the meta-analysis were 74% for sensitivity and 81% for specificity. This indicates that heterogeneity is substantial. Subgroups for different fracture locations, age groups, and risk of bias levels were investigated to identify possible sources of this heterogeneity.

**Subgroup meta-analysis**

Meta-analyses were performed for the most common fractures: those of the forearm, ankle, all long bones, humerus, and bones of the hand and foot. Studies involving only pediatric patients were also summarized. Detailed findings are shown in eTable 2.

A total of 14 studies investigated forearm fractures. Unlike an earlier review (1), we did not include the data of Moritz et al. (e21) or Sinha et al. (e22). This was due to unsuitable study design in the first case, and because the same cohort was described in a more recent publication (22) in the second.

Of the 6 studies investigating fractures of the ankle and foot, 4 used the Ottawa Foot and Ankle Rules (OFAR) as a clinical test to increase pretest probability.

To generate data on common ankle injuries we excluded studies on injuries to the lower leg not explicitly described as located in the ankle region.

Eight studies investigated the potential of ultrasound to diagnose humerus fractures. Five investigated only pediatric patients. Two investigated only supracondylar fractures of the humerus, and 2 only proximal fractures of the humerus.

Some authors (6, 10, 11, 14, e12) stated that diagnosis of diaphyseal long-bone fractures was more reliable than that of short-bone fractures and fractures close to joints. We compared the findings of ultrasound imaging of long bones, including the femur, humerus, forearm, and lower leg, with those of short bones (hands, feet, patella). For long bones, sensitivity was substantially higher and specificity slightly higher.

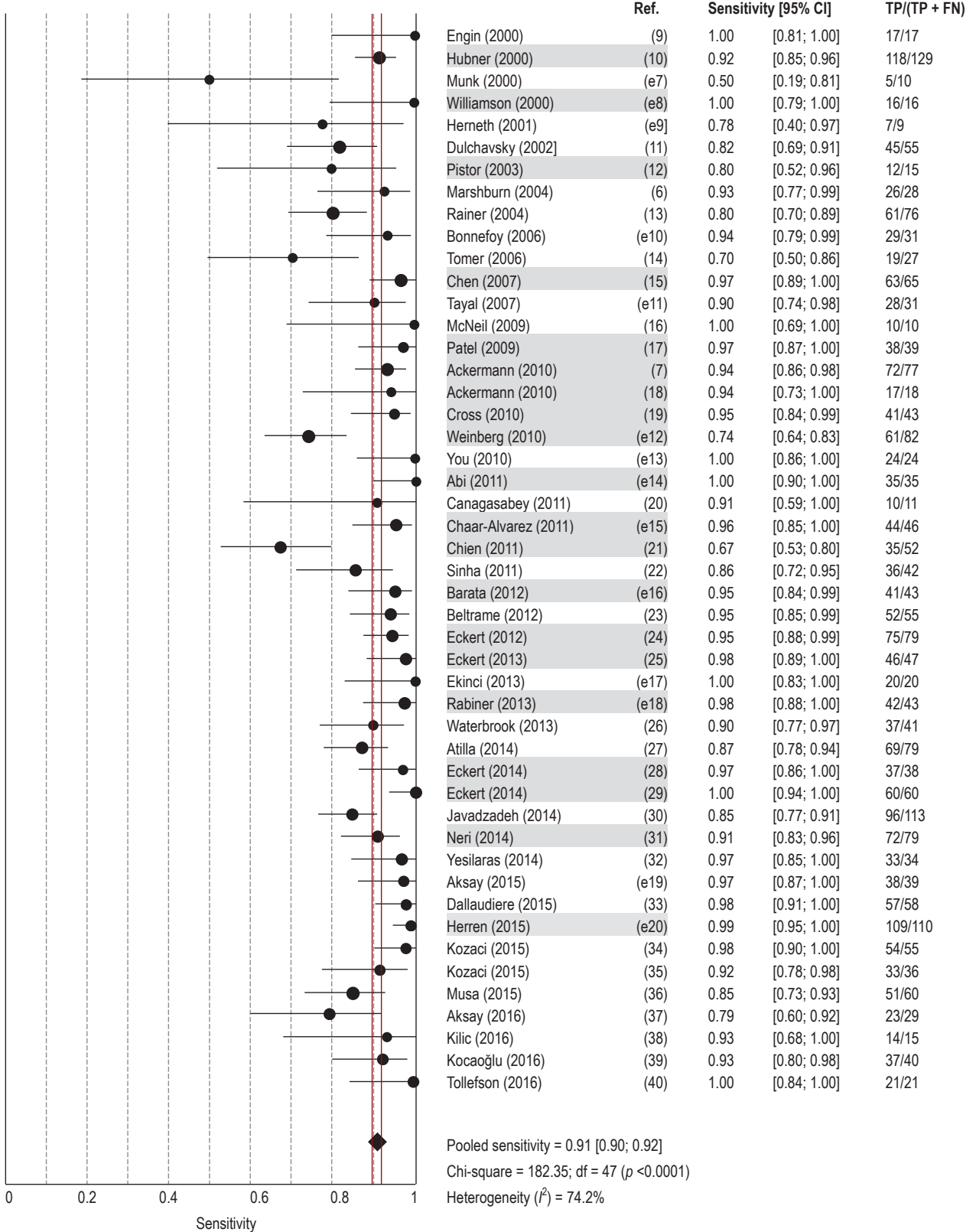
Reduced radiation exposure is particularly important for pregnant patients and children. Meta-analysis of the 20 studies conducted only in children and adolescents yielded higher sensitivity, but lower specificity, than for studies investigating only adult patients.

To estimate the effect of bias on study findings, meta-analyses were conducted of studies with a low/unclear risk of bias versus those with a high risk of bias. Sensitivity and specificity were higher in studies with a low/unclear risk of bias.

**Subpopulations:**

Twelve studies (6, 10–12, 17, 23, 26, 27, 30, 33, 36, e12) yielded clearly delineated data on multiple

FIGURE 2



Sensitivity of ultrasound imaging versus reference standard in the diagnosis of fractures in the included studies

Studies shaded in gray include examinations of children and adolescents only.

95% CI: 95% confidence interval; FN: Number of false negatives; Ref.: Reference; TP: Number of true positives

subpopulations (e.g. some patients with suspected forearm fractures and others with ankle fractures in the same publication). These subpopulations were included separately in meta-analyses of the relevant subgroups.

## Discussion

This systematic review summarizes data on the accuracy of ultrasound imaging in diagnosing fractures following acute trauma. Meta-analysis of 48 suitable studies found 91% sensitivity and 94% specificity. These findings show clearly that ultrasound of bone surfaces can be used as a suitable method for initial evaluation of acute injuries.

Ultrasound could be included in a diagnostic algorithm as the first machine-based method of diagnosis, instead of x-ray, following physical examination. All patients with true negative findings (corresponding to 56% of the included examinations) benefit from reduced radiation exposure. However, suitable treatment for patients with false negative findings (corresponding to 4%) might be delayed. High sensitivity is therefore vitally important in identifying fractures correctly and in ruling them out when findings are negative.

Fractures of the humerus, forearm, ankle, and long bones in general can be diagnosed with very high sensitivity and specificity. Fractures to bones of the hands and feet are more likely to be overlooked on ultrasound and should therefore be ruled out using another imaging procedure. In addition, our analyses showed ultrasound to be of greater diagnostic value in children and adolescents than in adults.

These findings are in line with those of previously published systematic reviews. Joshi et al. (3) summarized the findings of 8 relevant studies on the diagnostic value of ultrasound for fractures to the extremities and proposed the use of ultrasound in addition to x-ray. Katzer et al. (2) reported on 8 studies that investigated forearm fractures in children only. They found ultrasound to be of comparable diagnostic value to x-ray and superior to it in terms of patient comfort, time required, and cost-efficiency. Douma-den Hamer et al. (1) included 16 studies in a meta-analysis that investigated the accuracy of ultrasound in distal forearm fractures. Their calculations found sensitivity and specificity of 97% and 95% respectively, which broadly matches our findings (eTable 2). Their findings in pediatric forearm fractures only were even better.

The potential of ultrasound to diagnose fractures of the ribs and sternum should be evaluated separately. Most publications on this subject were excluded from our review due to unsuitable study design (4, e23–e29). X-ray cannot be considered the gold standard for these fractures, as some studies found ultrasound to be superior to it in diagnosing fractures of the ribs and sternum (13, e24, e25). Future studies on the accuracy of ultrasound in diagnosing fractures of the ribs and sternum should therefore select a reference standard that can establish the outcome correctly, such as MRI or strict follow-up examinations to detect callus formation, as proposed by Rainer et al. (13).

As noted by Douma den Hamer et al., the fact that occult and other fractures that were detected on ultrasound but not on x-ray were counted as false positives may result in systematic underestimation of the accuracy of ultrasound for other fracture locations too, such as those of the forearm (1).

With one exception, all the studies examined here were conducted in emergency rooms. Although the data reported supports the use of ultrasound as part of routine diagnostic algorithms for suspected fractures, it seems unlikely that current standards in emergency rooms will be changed without further incentives. Katzer et al. (2) calculated the costs of ultrasound and conventional x-ray in diagnosing forearm fractures in children as €20.54 for ultrasound and €26.60 for conventional x-ray. They reported that the greater efficiency of working processes when using ultrasound diagnostics played a vital role; often, a physician in an emergency room can make a diagnosis on the basis of history, clinical examination, and ultrasound, without needing to involve a radiologist or other staff. This factor would also be of particular interest in outpatient care, where the use of ultrasound in diagnosing fractures when possible could save resources and costs.

## Strengths and limitations

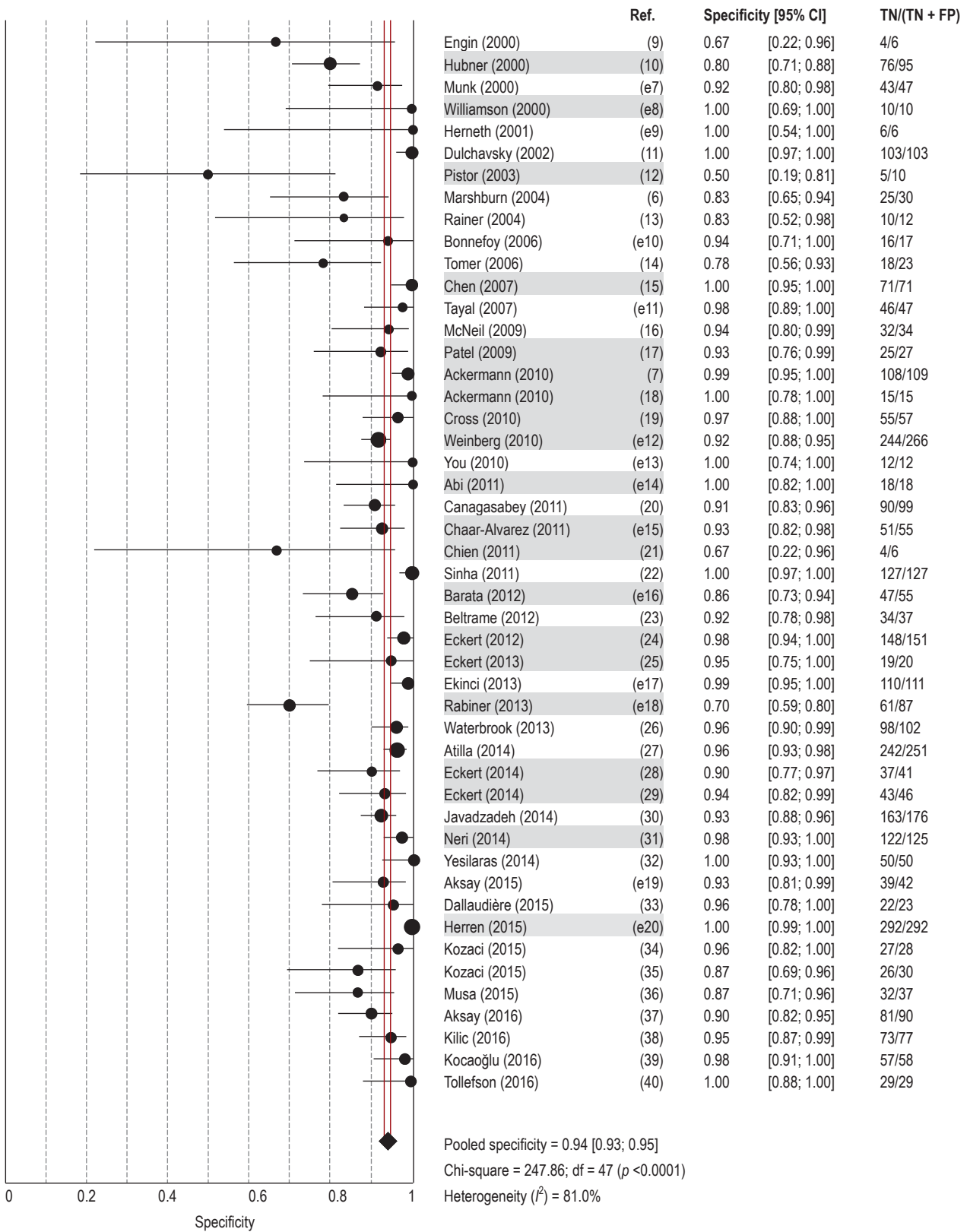
This review evaluates ultrasound imaging for a broad spectrum of fracture locations and thus provides an overview of potential uses of this method of diagnosis. We were able to include a substantial number of studies, which indicates that the amount of data on the sensitivity and specificity of ultrasound in diagnosing bone fractures has increased. Subgroup meta-analyses provide data for sophisticated discussion of the potential and limitations of this method.

Despite our best efforts, the findings of the meta-analysis must be interpreted critically, as there was substantial heterogeneity between the included studies. Although it was essentially shown that most study findings varied only slightly (Figures 2 and 3), the very high  $I^2$  values nevertheless show that the variability found is probably caused by genuine differences (e.g. different populations, fracture locations, experience of those performing examinations, etc.), not by chance alone. Unfortunately, subgroup analysis could not reveal all sources of heterogeneity. On the other hand, Rucker et al. (e30) emphasized that  $I^2$  values are susceptible to systematic error, and that ultimately the clinical relevance is more important. We therefore consider our meta-analysis valid and helpful in clinical decision-making.

The fracture locations included were limited to the limbs, ribs, sternum, clavicle, and scapula in order to focus the review on fractures that might also be treated in nonhospital settings.

The study protocol was not published before the review was performed. Language-related bias cannot be completely ruled out because the included studies were limited to publications in either English or German. However, only 6 titles that were potentially of

FIGURE 3



Specificity of ultrasound imaging versus reference standard in the diagnosis of fractures in the included studies

Studies shaded in gray include examinations of children and adolescents only.

95% CI: 95% confidence interval; FP: Number of false positives; Ref.: Reference; TN: Number of true negatives

**KEY MESSAGES**

- Ultrasound imaging can be used to view alterations to bone surfaces and thus identify fractures.
- The advantages of ultrasound over conventional x-ray are lack of radiation exposure, low levels of pain caused, and better availability, particularly in the outpatient sector and isolated areas such as developing countries.
- Many studies report high sensitivity and specificity for ultrasound in diagnosing fractures.
- Long bones and superficial structures can be evaluated more reliably than short bones in deep locations and bony structures close to joints.
- Sensitivity in fracture diagnosis is higher in children than in adults.

interest after title and abstract screening were excluded, suggesting that the potential effect on the final conclusions of the review was small.

We excluded absolutely all studies in which not all patients received ultrasound imaging and a reference standard. Although it is to be expected that this decision had a positive effect on the quality of the meta-analysis, it has resulted in the exclusion of some studies that make important contributions on diagnostic accuracy, practicability, and safety.

**Conclusion**

This review shows that there is already considerable evidence that the sensitivity and specificity of ultrasound imaging in diagnosing fractures following acute trauma are high. The use of ultrasound could be of greatest benefit to patients with suspected fractures of the long bones of the limbs or of superficial bones, and children and adolescents in general. Randomized controlled trials should be performed in the future in order to evaluate the potential impact of revised diagnostic algorithms on patient safety and on direct and indirect health costs.

**Acknowledgement**

We would like to thank Anja Heuser, Marie-Luise Meja, and Britt Häusler for their organizational support.

**Conflict of interest statement**

The authors declare that no conflict of interest exists.

Manuscript received on 15 March 2017, revised version accepted on 10 July 2017.

Translated from the original German by Caroline Shimakawa-Devitt, M.A.

**REFERENCES**

1. Douma-den Hamer D, Blanker MH, Edens MA, et al.: Ultrasound for distal forearm fracture: a systematic review and diagnostic meta-analysis. *PLoS One* 2016; 11: e0155659.
2. Katzer C, Wasem J, Eckert K, Ackermann O, Buchberger B: Ultrasound in the diagnostics of metaphyseal forearm fractures in children: a systematic review and cost calculation. *Pediatr Emerg Care* 2016; 32: 401–7.
3. Joshi N, Lira A, Mehta N, Paladino L, Sinert R: Diagnostic accuracy of history, physical examination, and bedside ultrasound for diagnosis of extremity fractures in the emergency department: a systematic review. *Acad Emerg Med* 2013; 20: 1–15.
4. Chan SSW: Emergency bedside ultrasound for the diagnosis of rib fractures. *Am J Emerg Med* 2009; 27: 617–20.
5. Linet MS, Kim KP, Rajaraman P: Children's exposure to diagnostic medical radiation and cancer risk: epidemiologic and dosimetric considerations. *Pediatr Radiol* 2009; Sonderheft 1: 4–26.
6. Marshburn TH, Legome E, Sargsyan A, et al.: Goal-directed ultrasound in the detection of long-bone fractures. *J Trauma* 2004; 57: 329–32.
7. Ackermann O, Liedgens P, Eckert K, et al.: Ultrasound diagnosis of juvenile forearm fractures. *J Med Ultrason* 2010; 37: 123–7.
8. Ackermann O, Eckert K, Schulze, Pellengahr C: Ultrasound diagnosis of forearm fractures in the growing age. No more x-ray. *Ultraschall in der Medizin, Supplement* 2013; Conference Ultraschall 2013 – Euroson 2013, Stuttgart, Germany, Conference Publication: DOI 10.1055/s-0033-1354901.
9. Engin G, Yekeler E, Guloglu R, Acunas B, Acunas G: US versus conventional radiography in the diagnosis of sternal fractures. *Acta Radiol* 2000; 41: 296–9.
10. Hubner U, Schlicht W, Outzen S, Barthel M, Halsband H: Ultrasound in the diagnosis of fractures in children. *J Bone Joint Surg Br* 2000; 82: 1170–3.
11. Dulchavsky SA, Henry SE, Moed BR, et al.: Advanced ultrasonic diagnosis of extremity trauma: the FASTER examination. *J Trauma* 2002; 53: 28–32.
12. Pistor G, Graffstadt H: Sonographic diagnosis of supracondylar fractures of the humerus. *Ultraschall Med* 2003; 24: 331–9.
13. Rainer TH, Griffith JF, Lam E, Lam PK, Metreweli C: Comparison of thoracic ultrasound, clinical acumen, and radiography in patients with minor chest injury. *J Trauma* 2004; 56: 1211–3.
14. Tomer K, Kleinbaum Y, Heyman Z, Dudkiewicz I, Blankstein A: [Ultrasound diagnosis of fractures in adults]. *Akt Traumatol* 2006; 36: 171–4.
15. Chen L, Kim Y, Moore CL: Diagnosis and guided reduction of forearm fractures in children using bedside ultrasound. *Pediatr Emerg Care* 2007; 23: 528–31.
16. McNeil CR, McManus J, Mehta S: The accuracy of portable ultrasonography to diagnose fractures in an austere environment. *Prehosp Emerg Care* 2009; 13: 50–2.
17. Patel DD, Blumberg SM, Crain EF: The utility of bedside ultrasonography in identifying fractures and guiding fracture reduction in children. *Pediatr Emerg Care* 2009; 25: 221–5.
18. Ackermann O, Sesia S, Berberich T, et al.: Sonographic diagnostics of proximal humerus fractures in juveniles. *Unfallchirurg* 2010; 113: 839.
19. Cross KP, Warkentine FH, Kim IK, Gracely E, Paul RI: Bedside ultrasound diagnosis of clavicle fractures in the pediatric emergency department. *Acad Emerg Med* 2010; 17: 687–93.

20. Canagasabay MD, Callaghan MJ, Carley S: The sonographic Ottawa Foot and Ankle Rules study (the SOFAR study). *Emerg Med J* 2011; 28: 838–40.
21. Chien M, Bulloch B, Garcia-Filion P, Youssef M, Shrader MW, Segal LS: Bedside ultrasound in the diagnosis of pediatric clavicle fractures. *Pediatr Emerg Care* 2011; 27: 1038–41.
22. Sinha TP, Kumar S, Bhoi S, et al.: Accuracy of point-of-care ultrasound for identifying fractures in patients with orthopaedic trauma presenting to emergency department of the All India Institute of Medical Sciences, level 1 trauma centre. *Crit Ultrasound J* 2011; 3: 67–70.
23. Beltrame V, Stramare R, Rebellato N, Angelini F, Frigo AC, Rubattelli L: Sonographic evaluation of bone fractures: a reliable alternative in clinical practice? *Clin Imaging* 2012; 36: 203–8.
24. Eckert K, Ackermann O, Schweiger B, Radeloff E, Liedgens P: Ultrasound as a viable alternative to standard X-rays for the diagnosis of distal forearm fractures in children. *Z Orthop Unfall* 2012; 150: 409–14.
25. Eckert K, Ackermann O, Schweiger B, Radeloff E, Liedgens P: Ultrasound evaluation of elbow fractures in children. *J Med Ultrason* 2013; 40: 443–51.
26. Waterbrook AL, Adhikari S, Stolz U, Adrion C: The accuracy of point-of-care ultrasound to diagnose long bone fractures in the ED. *Am J Emerg Med* 2013; 31: 1352–6.
27. Atilla OD, Yesilaras M, Kilic TY, et al.: The accuracy of bedside ultrasonography as a diagnostic tool for fractures in the ankle and foot. *Acad Emerg Med* 2014; 21: 1058–61.
28. Eckert K, Ackermann O, Janssen N, Schweiger B, Radeloff E, Liedgens P: Accuracy of the sonographic fat pad sign for primary screening of pediatric elbow fractures: a preliminary study. *J Med Ultrason* 2014; 41: 473–80.
29. Eckert K, Janssen N, Ackermann O, Schweiger B, Radeloff E, Liedgens P: Ultrasound diagnosis of supracondylar fractures in children. *Eur J Trauma Emerg Surg* 2014; 40: 159–68.
30. Javadzadeh HR, Davoudi A, Davoudi F, et al.: Diagnostic value of „bedside ultrasonography“ and the „water bath technique“ in distal forearm, wrist, and hand bone fractures. *Emerg Radiol* 2014; 21: 1–4.
31. Neri E, Barbi E, Rabach I, et al.: Diagnostic accuracy of ultrasonography for hand bony fractures in paediatric patients. *Archives of Disease in Childhood: Education and Practice Edition* 2014; 99: 1087–90.
32. Yesilaras M, Aksay E, Atilla OD, Sever M, Kalenderer O: The accuracy of bedside ultrasonography as a diagnostic tool for the fifth metatarsal fractures. *American Journal of Emergency Medicine* 2014; 32: 171–4.
33. Dallaudière B, Larbi A, Lefere M, et al.: Musculoskeletal injuries in a resource-constrained environment: comparing diagnostic accuracy of on-the-spot ultrasonography and conventional radiography for bone fracture screening during the Paris-Dakar rally raid. *Acta Radiol Open* 2015; 4: 2058460115577566.
34. Kozaci N, Ay MO, Akcimen M, et al.: Evaluation of the effectiveness of bedside point-of-care ultrasound in the diagnosis and management of distal radius fractures. *Am J Emerg Med* 2015; 33: 67–71.
35. Kozaci N, Ay MO, Akcimen M, Sasmaz I, Turhan G, Boz A: The effectiveness of bedside point-of-care ultrasonography in the diagnosis and management of metacarpal fractures. *Am J Emerg Med* 2015; 33: 1468–72.
36. Musa S, Wilson P: Ultrasonography and radiography: a comparison. *Emerg Nurse* 2015; 23: 34–7.
37. Aksay E, Kilic TY, Yesilaras M, Tur FC, Sever M, Kalenderer O: Accuracy of bedside ultrasonography for the diagnosis of finger fractures. *Am J Emerg Med* 2016; 34: 809–12.
38. Kilic TY, Yesilaras M, Atilla OD, Turgut A: The accuracy of point-of-care ultrasound as a diagnostic tool for patella fractures. *Am J Emerg Med* 2016; 34: 1576–8.
39. Kocaoğlu S, Özhasenekler A, İçme F, Pamukçu Günaydin G, Şener A, Gökhan Ş: The role of ultrasonography in the diagnosis of metacarpal fractures. *Am J Emerg Med* 2016; 34: 1868–71.
40. Tollefson B, Nichols J, Fromang S, Summers RL: Validation of the Sonographic Ottawa Foot and Ankle Rules (SOFAR) study in a large urban trauma center. *J Miss State Med Assoc* 2016; 57: 35–8.

**Corresponding author:**

Dr. med. Gordian Lukas Schmid  
 Independent Department for General Medicine  
 University of Leipzig  
 Philipp-Rosenthal-Str. 55  
 04103 Leipzig, Germany  
 Gordian.schmid@medizin.uni-leipzig.de

**Supplementary material**

For eReferences please refer to:  
[www.aerzteblatt-international.de/ref4517](http://www.aerzteblatt-international.de/ref4517)

eSupplement:  
[www.aerzteblatt-international.de/17m0757](http://www.aerzteblatt-international.de/17m0757)



Supplementary material to:

**The Investigation of Suspected Fracture—A Comparison of Ultrasound With Conventional Imaging Systematic Review and Meta-analysis**

by Gordian Lukas Schmid, Stefan Lippmann, Susanne Unverzagt, Christiane Hofmann, Tobias Deutsch, and Thomas Frese

Dtsch Arztebl Int 2017; 114: 757–64. DOI: 10.3238/arztebl.2017.0757

**eREFERENCES**

e1. Heiner JD, Baker BL, McArthur TJ: The ultrasound detection of simulated long bone fractures by U.S. Army Special Forces Medics. *J Spec Oper Med* 2010; 10: 7–10.

e2. Moher D, Liberati A, Tetzlaff J, Altman DG: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010; 8: 336–41.

e3. Stroup DF, Berlin JA, Morton SC, et al.: Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; 283: 2008–12.

e4. Huang X, Lin J, Demner-Fushman D: Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annu Symp Proc* 2006; 359–63.

e5. Whiting PF, Rutjes AWS, Westwood ME, et al.: QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–36.

e6. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A: Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006; 6: 31.

e7. Munk B, Bolvig L, Kroner K, Christiansen T, Borris L, Boe S: Ultrasound for diagnosis of scaphoid fractures. *J Hand Surg Br* 2000; 25: 369–71.

e8. Williamson D, Watura R, Cobby M: Ultrasound imaging of forearm fractures in children: a viable alternative? *J Accid Emerg Med* 2000; 17: 22–4.

e9. Herneth AM, Siegmeth A, Bader TR, et al.: Scaphoid fractures: evaluation with high-spatial-resolution US initial results. *Radiology* 2001; 220: 231–5.

e10. Bonnefoy O, Diris B, Moïnard M, Aunoble S, Diard F, Hauger O: Acute knee trauma: role of ultrasound. *Eur Radiol* 2006; 16: 2542–8.

e11. Tayal VS, Antoniazzi J, Pariyadath M, Norton HJ: Prospective use of ultrasound imaging to detect bony hand injuries in adults. *J Ultrasound Med* 2007; 26: 1143–8.

e12. Weinberg ER, Tunik MG, Tsung JW: Accuracy of clinician-performed point-of-care ultrasound for the diagnosis of fractures in children and young adults. *Injury* 2010; 41: 862–8.

e13. You JS, Chung YE, Kim D, Park S, Chung SP: Role of sonography in the emergency room to diagnose sternal fractures. *J Clin Ultrasound* 2010; 38: 135–7.

e14. Abi KS, Haddad-Zebouni S, Roukoz S, et al.: Ultrasound as an adjunct to radiography in minor musculoskeletal pediatric trauma. *J Med Liban* 2011; 59: 70–4.

e15. Chaar-Alvarez FM, Warkentine F, Cross K, Herr S, Paul RI: Bedside ultrasound diagnosis of nonangulated distal forearm fractures in the pediatric emergency department. *Pediatr Emerg Care* 2011; 27: 1027–32.

e16. Barata I, Spencer R, Suppiah A, Raio C, Ward MF, Sama A: Emergency ultrasound in the detection of pediatric long-bone fractures. *Pediatr Emerg Care* 2012; 28: 1154–7.

e17. Ekinci S, Polat O, Günalp M, Demirkan A, Koca A: The accuracy of ultrasound evaluation in foot and ankle trauma. *Am J Emerg Med* 2013; 31: 1551–5.

e18. Rabiner JE, Khine H, Avner JR, Friedman LM, Tsung JW: Accuracy of point-of-care ultrasonography for diagnosis of elbow fractures in children. *Ann Emerg Med* 2013; 61:9–17.

e19. Aksay E, Yesilaras M, Kilic TY, Tur FC, Sever M, Kaya A: Sensitivity and specificity of bedside ultrasonography in the diagnosis of fractures of the fifth metacarpal. *Emerg Med J* 2015; 32: 221–5.

e20. Herren C, Sobottke R, Ringe MJ, et al.: Ultrasound-guided diagnosis of fractures of the distal forearm in children. *Orthop Traumatol Surg Res* 2015; 101: 501–5.

e21. Moritz JD, Berthold LD, Soenksen SF, Alzen GF: Ultrasound in diagnosis of fractures in children: unnecessary harassment or useful addition to X-ray? *Ultraschall Med* 2008; 29: 267–74.

e22. Sinha TP, Bhoi S, Kumar S, et al.: Diagnostic accuracy of bedside emergency ultrasound screening for fractures in pediatric trauma patients. *J Emerg Trauma Shock* 2011; 4: 443–5.

e23. Racine S, Drake D: BET 3: Bedside ultrasound for the diagnosis of sternal fracture. *Emerg Med J* 2015; 32: 971–2.

e24. Wustner A, Gehmacher O, Hammerle S, Schenkenbach C, Hafele H, Mathis G: [Ultrasound diagnosis in blunt thoracic trauma]. *Ultraschall Med* 2005; 26: 285–90.

e25. Jin W, Yang DM, Kim HC, Ryu KN: Diagnostic values of sonography for assessment of sternal fractures compared with conventional radiography and bone scans. *Journal of Ultrasound in Medicine* 2006; 25: 1263–8.

e26. Hurley ME, Keye GD, Hamilton S: Is ultrasound really helpful in the detection of rib fractures? *Injury* 2004; 35: 562–6.

e27. Mahlfeld A, Franke J, Mahlfeld K: [Ultrasound diagnosis of sternal fractures]. *Zentralbl Chir* 2001; 126: 62–4.

e28. Uzun M, Beksac B, Kara A, Kucukdurmaz F, Kircali BA, Tetik C: Ultrasonography as a better diagnostic efficiency in rib fracture. *Journal of Experimental and Clinical Medicine (Turkey)* 2013; 30: 133–5.

e29. Mattox R, Reckelhoff KE, Welk AB, Kettner NW: Sonography of occult rib and costal cartilage fractures: a case series. *J Chiropr Med* 2014; 13: 139–43.

e30. Rucker G, Schwarzer G, Carpenter JR, Schumacher M: Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008; 8: 79.

e31. Higgins JPT GS: *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0.* www.handbook.cochrane.org.

eTABLE 1

**Characteristics of studies and patients included in meta-analysis**

Main author (year)	Ref.	Fracture location	Reference standard	No. of patients	No. of bones examined	Fractures, % (no.)	Age in years, mean ± SD (age range) <sup>a</sup>	Female/male, %
Engin (2000)	(9)	Sternum (adults)	X-ray	23	23	73.9 (17)	35.4 (18 to 75)	17/83
Hubner (2000)	(10)	Mixed, extremities, skull, clavicle (children)	X-ray	163	224	57.6 (129)	N/A	N/A
Munk (2000)	(e7)	Scaphoid (adults)	X-ray + radiological follow-up examination	57	57	17.5 (10)	38 (13 to 82)	46/54
Williamson (2000)	(e8)	Forearm (children)	X-ray	26	26	61.5 (16)	8 (2 to 13)	38/62
Hemeth (2001)	(e9)	Scaphoid (adults)	MRI	15	15	60.0 (9)	23.5 (15.8 to 55.2)	53/47
Dulchavsky (2002)	(11)	Mixed, extremities (adults)	X-ray	95	158	34.8 (55)	N/A	18/82
Pistor (2003)	(12)	Elbow, humerus (children)	X-ray	25	25	60.0 (15)	N/A	N/A
Marshburn (2004)	(6)	Humerus, femur (adults)	X-ray or CT	58	58	48.3 (28)	79 (28 to 97)	N/A
Rainer (2004)	(13)	Sternum, ribs	Radiological follow-up examination	88	88	86.4 (76)	51 ± 19	41/59
Bonnefoy (2006)	(e10)	Knee (adults)	CT (+ MRI in one case)	48	48	64.6 (31)	36.3 (14 to 88)	23/77
Toner (2006)	(14)	Mixed, extremities (adults)	X-ray	50	50	54.0 (27)	44 ± 20	N/A
Chen (2007)	(15)	Forearm (children)	X-ray	68	136	47.8 (65)	10 (2 to 21)	40/60
Tayal (2007)	(e11)	Hand (adults)	X-ray or intraoperative findings	78	78	39.7 (31)	34 ± 14	45/55
McNeil (2009)	(16)	Mixed, extremities (adults)	X-ray or clinical follow-up examination	44	44	22.7 (10)	N/A	N/A
Patel (2009)	(17)	Mixed, extremities (children)	X-ray	33	66	59.1 (39)	9.1 ± 3.1	67/33
Ackermann (2010)	(7)	Forearm (children)	X-ray	93	186	41.4 (77)	8.1 (0 to 12)	53/47
Ackermann (2010)	(18)	Humerus (children)	X-ray	33	33	54.5 (18)	7.6 (0 to 12)	58/42
Cross (2010)	(19)	Clavicle (children)	X-ray	100	100	43.0 (43)	9.1 ± 3.1	25/75
Weinberg (2010)	(e12)	Mixed, extremities, skull, sternum, ribs, clavicle (children)	X-ray (+ CT for skull)	212	348	23.6 (82)	13 (IQR 10)	N/A
You (2010)	(e13)	Sternum (adults)	Clinical consensus of review board	36	36	66.7 (24)	43 (8 to 78)	47/53
Abi (2011)	(e14)	Mixed, extremities (children)	Clinical consensus + radiological follow-up examination	50	53	66.0 (35)	N/A	N/A
Canagasabay (2011)	(20)	Foot and ankle (adults)	X-ray	110	110	10.0 (11)	31.8 ± 12.7	41/59
Chaar-Alvarez (2011)	(e15)	Forearm (children)	X-ray	101	101	45.5 (46)	10.3 ± 4.3	43/57
Chien (2011)	(21)	Clavicle (children)	X-ray	58	58	89.7 (52)	7.4 (0.25 to 16)	34/66
Sinha (2011)	(22)	Mixed, extremities	X-ray	133	169	24.9 (42)	29.8 (7 to 70)	28/72
Barata (2012)	(e16)	Mixed, extremities, clavicle (children)	X-ray	53	98	43.9 (43)	10.2 ± 3.8	43/57

Main author (year)	Ref.	Fracture location	Reference standard	No. of patients	No. of bones examined	Fractures, % (no.)	Age in years, mean ± SD (age range)*	Female/male, %
Beltrame (2012)	(23)	Mixed: extremities, skull, sternum, ribs, clavicle	X-ray	86	86	59.8 (55)	53 (7 to 88)	53/47
Eckert (2012)	(24)	Forearm (children)	X-ray	115	230	34.3 (79)	9.1 (2 to 14)	44/56
Eckert (2013)	(25)	Elbow (children)	X-ray	67	67	70.1 (47)	6 (1 to 13)	48/52
Ekinci (2013)	(e17)	Ankle and foot (adults)	X-ray	131	131	15.3 (20)	37.2 ± 15.44	51/49
Rabiner (2013)	(e18)	Elbow (children)	X-ray + radiological follow-up examination in unclear cases	130	130	33.1 (43)	7.5 ± 5.4	45/55
Waterbrook (2013)	(26)	Mixed, extremities, clavicle	X-ray	103	143	28.7 (41)	34 ± 20	50/50
Atilla (2014)	(27)	Ankle and foot (adults)	X-ray + CT (if requested)	246	330	23.9 (79)	37 (IQR 19)	57/43
Eckert (2014)	(28)	Elbow (children)	X-ray + radiological or clinical follow-up examination in unclear cases	79	79	48.1 (38)	6.5 (1 to 13)	N/A
Eckert (2014)	(29)	Humerus (children)	X-ray	106	106	56.6 (60)	5.9 (1 to 13)	44/56
Javadzadeh (2014)	(30)	Forearm, hand (adults)	X-ray	260	289	39.1 (113)	42.63 ± 12.7	43/57
Neri (2014)	(31)	Hand (children)	X-ray	204	204	38.7 (79)	12 ± 3	35/65
Yesilaras (2014)	(32)	Foot (adults)	X-ray	84	84	40.5 (34)	36 ± 15	57/43
Aksay (2015)	(e19)	Hand (adults)	X-ray	81	81	48.1 (39)	28 (IQR 19 to 35)	14/86
Dallaudière (2015)	(33)	Mixed, extremities, ribs, scapula, clavicle (adults)	X-ray	81	81	71.6 (58)	40.6 ± 9.9	2/98
Herren (2015)	(e20)	Forearm (children)	X-ray	201	402	27.4 (110)	N/A	34/66
Kozaci (2015)	(34)	Forearm	X-ray	83	83	66.3 (55)	M: 13 ± 14/F: 15 ± 13	22/78
Kozaci (2015)	(35)	Hand	X-ray	66	66	54.5 (36)	24 ± 10	11/89
Musa (2015)	(36)	Mixed: extremities	X-ray	97	97	61.9 (60)	N/A	N/A
Aksay (2016)	(37)	Hand (adults)	X-ray	119	119	24.4 (29)	27 (IQR 16 to 36)	39/61
Kilic (2016)	(38)	Patella (adults)	X-ray + CT (if findings are unclear)	92	92	16.3 (15)	37.5 (IQR 29.5)	40/60
Kocaoğlu (2016)	(39)	Hand (adults)	X-ray	96	98	40.8 (40)	30.1 ± 11.8	N/A
Tollefson (2016)	(40)	Foot and ankle (adults)	X-ray	50	50	42.0 (21)	35	46/54

CT: Computed tomography; (adults): Studies involving adult patients only; IQR: Interquartile range; N/A: Information not available in publication; (children): Studies involving children and adolescents only; X-ray: X-ray images taken in various planes and using various techniques for various fracture locations; MRI: Magnetic resonance imaging; SD: Standard deviation  
 \*Age range stated where either IQR or standard deviation was unavailable.

eTABLE 2

**Subgroup analyses**

	No. of studies	No. of patients	No. of ultrasound examinations, <i>n</i>	Sensitivity [95% CI]	Specificity [95% CI]
All studies	48	4427	5436	0.91 [0.90; 0.92]	0.94 [0.93; 0.95]
<b>Population:</b>					
Adults only	21	1854	2032	0.90 [0.87; 0.92]	0.95 [0.93; 0.96]
Children and adolescents only	20	1917	2672	0.93 [0.91; 0.94]	0.93 [0.92; 0.94]
<b>Fracture location:</b>					
Long bones	22 (incl. subpopulations)		2700	0.93 [0.91; 0.94]	0.96 [0.95; 0.97]
Hand, foot	14 (incl. subpopulations)		1219	0.86 [0.83; 0.89]	0.94 [0.92; 0.96]
Forearm	14 (incl. subpopulations)		1658	0.95 [0.93; 0.97]	0.96 [0.95; 0.98]
Ankle, foot	6 (incl. subpopulations)		722	0.93 [0.88; 0.96]	0.96 [0.95; 0.98]
Humerus	8 (incl. subpopulations)		233	0.98 [0.93; 1.00]	0.92 [0.86; 0.96]
<b>Risk of bias</b>					
Low/unclear	21	1875	2233	0.93 [0.91; 0.94]	0.95 [0.94; 0.96]
High	27	2552	3203	0.90 [0.88; 0.92]	0.94 [0.92; 0.95]

Meta-analyses of fracture locations also include subpopulations of multiple studies. 95% CI: 95% confidence interval

### Research strategy

This systematic review was planned and performed in line with the PRISMA (e2) and MOOSE (e3) recommendations. The search strategy was developed by 2 physicians (GLS, TF) and one librarian (CH).

The electronic search of the literature involved 3 databases (Medline, EMBASE, Cochrane Central Register of Controlled Trials [CENTRAL]) and was performed on March 19, 2015. We used a combination of MESH/EMTREE terms and search terms.

#### Medline via PubMed:

((„fractures, bone“[mesh] OR fracture\*[tiab] OR „bone injuries“[tiab] OR „bone injury“[tiab] OR „broken bones“[tiab] OR „broken bone“[tiab]) AND (ultrasonography[mesh] OR ultrasound[tiab] or echogram[tiab] OR ultrasonic[tiab]) AND diagnos\*[tiab])

Each time we came across conference abstracts with relevant content but without traceable original data, we searched manually for other publications by the same authors that did contain the data in question.

The datasets resulting from the database search were imported into a file using Citavi 4 (Swiss Academic Software, Wädenswil, Switzerland). Duplicates and studies published before the year 2000 were deleted. Next, publications' titles and then abstracts were screened; both processes were performed by 2 researchers (GLS, TF). The review was limited to publications dating from 2000 onwards in order to take account of technical advances in ultrasound machines (higher frequencies and resolutions, better display methods). Older studies might underestimate the potential of ultrasound to diagnose fractures. Searches in fulltext bibliographies and online were also performed by 2 researchers (GLS, TF), as was data extraction. These individuals also decided whether publications were to be included or excluded.

The database search was last updated in September 2016. The study protocol is shown in *eMethods 2* and was not published in advance of the review. The methodological quality of the included studies was evaluated by 2 researchers (GLS, SL) using the QUADAS-2 tool (QUADAS: quality assessment of diagnostic accuracy studies) (e5). Further, detailed information on this can be found in *eMethods 3*.

#### eBOX 1

### Embase via Ovid:

#### # Search terms

1. fracture/ or exp arm fracture/ or exp leg fracture/ or exp limb fracture/ or exp rib fracture/ or („bone injury“ or broken bone\* or fracture\*).ti,ab.
2. exp echography/ or (ultrasound or sonography or ultrasonography or sonogram or ultrasonic).ti,ab.
3. diagnos\*.ti,ab.
4. #1 AND #2 AND #3

### Inclusion and exclusion criteria

The PICO Framework (e4) was used to define the inclusion and exclusion criteria:

**Population:** All human patients of any age group presenting after acute trauma; pretests: history and physical examination (also clinical decision-making aids, e.g. the Ottawa Ankle Rules, if stated); suspected fracture of limbs, ribs, sternum, scapula, or clavicle; exclusion of studies reporting only on birth injuries, pathological fractures, stress fractures, or occult fractures. In addition, polytrauma patients and patients with fractures of the skull, vertebrae, pelvis, or hip were excluded due to the need for hospital care and the dangerous nature of the associated health conditions.

**Intervention:** Ultrasound imaging as diagnostic procedure (the index test).

**Comparison:** “Other diagnostic imaging procedures” as the reference standard (conventional x-ray, CT, MRI, or consensus of treating physicians taking all available findings into account).

**Outcome:** Test accuracy (sensitivity and specificity), 4-field tables (true positive, true negative, false positive, false negative).

**Study types:** Studies on diagnostic accuracy with a comparative design (**all** patients undergoing an index test **and** a reference standard); case reports, case series, letters to the editor, and conference abstracts with no publication of original data were excluded. In addition, we included only articles written in German or English.

#### eBOX 2

### Cochrane Central Register of Controlled Trials:

#### # Search terms

1. fracture/ or exp arm fracture/ or exp leg fracture/ or exp limb fracture/ or exp rib fracture/ or („bone injury“ or broken bone\* or fracture\*).ti,ab.
2. exp echography/ or (ultrasound or sonography or ultrasonography or sonogram or ultrasonic).ti,ab.
3. diagnos\*.ti,ab,kw
4. #1 AND #2 AND #3

### Meta-analysis and statistics

Percent agreement and Cohen's kappa for the 2 raters (GLS, TF) were calculated to estimate the interrater reliability of the initial screening process. Data from the included studies was pooled using a random-effects model. Heterogeneity was evaluated by means of the chi-square test and  $I^2$ , using the program MetaDiSc (version 1.4, Hospital Universitario Ramón y Cajal, Madrid, Spain) (e6).

Sensitivity and specificity data is given in the format [number (95% confidence interval)].

Heterogeneity was assessed using the  $I^2$  test. Interpretation of this depends very much on the researcher.  $I^2$  findings were interpreted using the Cochrane Hand-

book (e31).  $I^2$  values between 0 and 40% can be considered as representing unimportant heterogeneity, those between 40 and 60% moderate heterogeneity, between 60 and 80% substantial heterogeneity, and between 80 and 100% considerable heterogeneity. We suspected that differing fracture locations and the varying age structures of the study populations were potential sources of heterogeneity. We performed subgroup analyses for various fracture locations (humerus, forearm, ankle and foot, long bones of the extremities, short bones of the hands and feet), age groups (adults versus children and adolescents), and risks of bias (low/unclear versus high risk).

**Research question**

- Is ultrasound a suitable tool for diagnosing fractures in outpatient settings?
- Accuracy of ultrasound versus conventional x-ray/CT/MRI
- Parts of body: arm, leg, chest, clavicle, scapula; excluded: skull, pelvis, vertebrae
- PICO Framework:
  - Population:
    - All available patients with injuries that may have led to fractures
    - Clinical presentation with at least one unclear sign of fracture
    - Primary and secondary care
    - Pretest: history and physical examination
    - Children included
    - Excluded: Birth injuries/neonates, occult fractures according to x-ray
  - Intervention: Ultrasound as diagnostic procedure = index test
  - Comparison: X-ray, CT, MRI, best clinical consensus
  - Outcome: Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV)

**Protocol:**

**Databases**

- Medline via PubMed
- Embase
- Cochrane Controlled Trials Register
- Google Scholar
- Bibliographies of publications identified as relevant

**Inclusion criteria**

- Study type = diagnostic accuracy studies, cross-sectional studies, cross-sectional studies with follow-up examinations, prognostic accuracy studies
- Ultrasound AND reference standard used
- Languages: English, German

**Exclusion criteria**

- No fulltext available
- Reports on individual cases and case series
- Systematic reviews
- Intensive care or surgery required; open fractures, clear signs of fracture
- Studies in neonates only

**eTABLE 3**

<b>Inclusion and exclusion criteria</b>		
<b>Criterion</b>	<b>Inclusion</b>	<b>Exclusion</b>
Study type	Ultrasound versus reference standard	Case reports, case series, no original data (e.g. reviews, commentaries), conference abstracts only
Patients	All age groups; fracture suspected on clinical grounds; pretest: clinical examination, Ottawa Foot and Ankle Rules, or similar	Only neonates or birth injuries; open fractures; intensive care needed; polytrauma; pretest: negative x-ray finding, occult fractures
Languages	English, German	Others
Bones	Limbs, sternum, clavicle, ribs	Skull, hip, pelvis, vertebrae
Comparison	Ultrasound vs. x-ray, CT, MRI, bone scintigraphy, follow-up examination, or clinical consensus	Studies showing ultrasound to be superior to x-ray; no comparator method to confirm diagnosis, no reference standard

## Procedure

### Screening

- One researcher performs database search, extraction of all suitable findings, deletion of duplicates
- Screening (of titles and abstracts) by 2 researchers, Citavi and Excel spreadsheet
- Extracted publications: fulltexts obtained, excluded if no fulltext available
- Two researchers: screening of bibliographies of selected publications
- Manual search using Google and other online sources
- Comparison of lists; where differences: discussion and advice from third rater

### Fulltext phase

- Screening of fulltexts by 2 raters leading to possible exclusion
- Data extracted to extraction file
- Final decision on inclusion/exclusion: every exclusion discussed by 2 raters
- Quality evaluated by 2 raters using the QUADAS-2 tool; discussion and advice from third rater in the event of disagreement

### Data extraction file

- ID number: Every study was allocated a unique ID number.
- Bibliographical information: Main author, year of publication, institution
- Aim of study
- Examined part of body
- Is data on subpopulations for meta-analysis provided?
- Study type (prospective, retrospective, controlled, case series)
- Type of patient sample (random, sequential, randomized)
- Number of patients, number of examined bones
- Demographic characteristics of patients: Age (mean, median), age range, sex
- Index test
- Details of ultrasound machine (manufacturer, frequencies used)
- Reference standard (which? for all patients? the same for all patients?)
- Measured outcomes (primary, secondary)
- Four-field tables wherever possible (based on sensitivity and specificity where applicable)
- Time from injury to examination
- Pretests (e.g. physical examination, particular clinical decision-making aids)
- Is a definition of fracture given (e.g. fracture line)?
- Clinical setting (emergency room, general medical practice, etc.)
- Were those who performed ultrasound examination involved in clinical treatment?
- Who performed ultrasound examination? Who performed the reference standard?
- Was test evaluation blinded?
- Was specific ultrasound training for bones given?
- Were any patients excluded inappropriately?
- Reason for exclusion if excluded (language, no original data, abstract only, no reference standard)

The extracted data was used for meta-analysis and subgroup analysis following quality evaluation.

This study protocol was not published before the systematic review was performed.



We used QUADAS-2 (e5) to evaluate the quality of the included studies. For this we used a Microsoft Access™ database provided by the University of Bristol (<http://www.bristol.ac.uk/social-community-medicine/projects/quadas/resources/>). We tailored the signaling questions to our review. We used the following standardized rules to evaluate the risk of bias:

- If all questions can be answered with “YES” = low risk of bias
- If the answer to one or more question is “NO” = high risk of bias
- If the answer to one or more question is “UNCLEAR” and the answer to the others is “YES” = unclear risk of bias

We used the following questions:

**Domain 1: Patient selection**

- Was a consecutive or random sample of patients enrolled?
- Did the study avoid inappropriate exclusions?
- Was pretesting limited to case history and physical examination?

**Domain 2: Index test**

- Were the index test results interpreted without knowledge of the results of the reference standard?
- Were the investigators specially trained for fracture diagnosis?
- Were the criteria for consideration as “fracture positive” pre-specified?

**Domain 3: Reference standard**

- Is the reference standard likely to correctly classify the target condition (fracture)?
- Was the reference standard independent of ultrasound results? (e.g. clinical consensus)
- Were the reference standard results interpreted without knowledge of the results of the index test (blinded)?

**Domain 4: Flow and timing**

- Was there an appropriate interval between index test(s) and reference standard?
- Did all patients receive a reference standard?
- Did all patients receive the same reference standard?
- Were all patients included in the analysis?

eTABLE 4

**Results of the quality assessment**

Main author (reference)	Year	Could the selection of patients have introduced bias?	Are there concerns that the included patients do not match the review question?	Could the conduct or interpretation of the index test have introduced bias?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	Could the patient flow have introduced bias?
Engin (9)	2000	0	+	0	+	–	+	0
Hübner (10)	2000	0	+	–	+	–	+	+
Munk (e7)	2000	+	+	0	+	+	+	+
Williamson (e8)	2000	0	+	–	+	0	+	0
Herneth (e9)	2001	+	+	0	+	0	+	+
Dulchavsky (11)	2002	0	–	+	–	–	–	+
Pistor (12)	2003	0	–	0	0	0	+	0
Marshburn (6)	2004	–	0	+	+	0	+	0
Rainer (13)	2004	+	+	0	+	0	–	–
Bonnefoy (e10)	2006	+	+	0	+	0	+	–
Tomer (14)	2006	0	–	+	+	–	+	0
Chen (15)	2007	–	+	–	+	0	+	0

Main author (reference)	Year	Could the selection of patients have introduced bias?	Are there concerns that the included patients do not match the review question?	Could the conduct or interpretation of the index test have introduced bias?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Are there concerns that the target condition as defined by the reference standard does not match the review question?	Could the patient flow have introduced bias?
Tayal (e11)	2007	0	+	+	+	+	+	0
McNeil (16)	2009	-	+	0	+	-	+	-
Patel (17)	2009	0	+	+	+	+	+	0
Ackermann (7)	2010	0	+	0	+	0	+	0
Ackermann (18)	2010	0	+	0	+	0	+	0
Cross (19)	2010	-	+	-	+	+	+	+
Weinberg (e12)	2010	-	+	+	+	+	+	-
You (e13)	2010	0	+	0	+	-	+	+
Abi (e14)	2011	-	-	-	-	-	-	-
Canagasabay (20)	2011	0	+	+	+	0	+	0
Chaar-Alvarez (e15)	2011	-	+	-	+	+	+	-
Chien (21)	2011	-	+	+	+	+	+	+
Sinha (22)	2011	-	+	+	+	+	+	0
Beltrame (23)	2012	-	+	0	+	+	+	0
Eckert (24)	2012	0	+	-	+	+	+	0
Barata (e16)	2012	-	+	+	+	+	+	0
Eckert (25)	2013	0	+	0	+	+	+	0
Rabiner (e18)	2013	-	+	+	+	+	+	0
Waterbrook (26)	2013	-	+	+	+	+	+	0
Ekinci (e17)	2013	0	+	-	+	0	+	0
Atila (27)	2014	+	+	+	+	+	+	-
Javadzadeh (30)	2014	0	+	+	+	+	+	0
Neri (31)	2014	-	+	+	+	+	+	0
Yesilaras (32)	2014	+	+	+	+	+	+	0
Eckert (28)	2014	0	+	+	+	+	+	0
Eckert (29)	2014	0	+	+	+	+	+	0
Aksay (e19)	2015	+	+	+	+	+	+	0
Kozaci (34)	2015	0	+	+	+	+	+	0
Herren (e20)	2015	0	+	+	+	+	+	0
Musa (36)	2015	0	+	0	+	+	+	0
Dallaudiere (33)	2015	+	+	-	+	+	+	0
Kozaci (35)	2015	0	-	0	+	+	+	0
Tollefson (40)	2016	-	+	0	+	0	+	0
Kocaoğlu (39)	2016	0	+	0	+	+	+	0
Kilic (38)	2016	+	+	+	+	+	+	0
Aksay (37)	2016	+	+	+	+	+	+	0

+: Low risk of bias/no concerns regarding applicability; -: High risk of bias/concerns regarding applicability; 0: Unclear risk of bias/unclear whether there are concerns regarding applicability

### Discussion of method

There were 2 problems concerning the signaling questions during quality evaluation. Firstly, the question “Were the criteria for consideration as “fracture positive” pre-specified?” was worded in such a way that the answer “NO” had to be selected if there was insufficient information, whereas the answer “UNCLEAR” could be selected for other questions (e.g. “Were the investigators specially trained for fracture diagnosis?”) if the relevant information was not stated in the text. These questions may therefore be overrepresented in the evaluation of risk of bias.

The second problem arose from the unforeseen fact that, in many studies conducted in an emergency room, ultrasound was integrated into standard working procedures, and the time interval from admission to examination was not explicitly stated. Although one might assume that ultrasound examination was performed appropriately swiftly, we made a formal decision to classify the risk of bias as “UNCLEAR” if manuscripts did not report clear information on timeframes. This may have led to an underestimate of the quality of some of the included studies.

**Method**

We used funnel plots to assess the risk of publication bias (eFigures 1–4). 1 – standard error and the number of examinations described per study were plotted separately against sensitivity and specificity. Standard errors were calculated as follows:

$$SE \text{ (standard error)} = \text{sensitivity} \times (1 - \text{sensitivity}) / \text{number of examinations}$$

$$SE = \text{specificity} \times (1 - \text{specificity}) / \text{number of examinations}$$

**Findings**

The funnel plots show an increase in spread as standard error increases. The findings of studies with few examinations deviate more from the pooled value, for both sensitivity and specificity.

Our visual assessment of the funnel plots is that they are symmetrical. We consider the risk of significant publication bias to be low.

Because the pooled results for sensitivity and specificity are close to 1 and the findings of individual studies cannot be higher, it must be assumed that—even with normal distribution—only one half of the expected funnel shape, the left-hand side, will be present.

