Low intensity pulsed ultrasound for bone healing: a systematic review of randomised controlled trials

Journal:	*BMJ*
Manuscript ID	BMJ.2016.036958.R1
Article Type:	Research
BMJ Journal:	BMJ
Date Submitted by the Author:	23-Jan-2017
Complete List of Authors:	Schandelmaier, Stefan; McMaster University Faculty of Health Sciences, Department of Clinical Epidemiology and Biostatistics; University Hospital Basel, Intitute for Clinical Epidemiology and Biostatistics Kaushal, Alka; McMaster University, Clinical Epidemiology and Biostatistics Lytvyn, Lyubov; Oslo Universitetssykehus Heels-Ansdell, Diane; McMaster University, Clinical Epidemiology and Biostatistics Siemieniuk, Reed; McMaster University, Clinical Epidemiology & Biostatistics; University of Toronto, Medicine Agoritsas, Thomas; McMaster University, Department of Clinical Epidemiology and Biostatistics; University Hospitals of Geneva, Division of General Internal Medicine & Division of Epidemiology Guyatt, Gordon; McMaster University, Vandvik, Per; Innlandet Hospital Trust-divisjon Gjøvik, Department of Medicine Couban, Rachel; McMaster University, Michael G. DeGroote Institute for Pain Research and Care Mollon, Brent; Orillia Soldiers' Memorial Hospital Busse, Jason; McMaster University, Anesthesia
Keywords:	systematic review, meta-analysis, randomised controlled trials, fracture healing, orthopaedic procedures, ultrasonic waves, ultrasonic therapy, patient-reported outcomes

Low intensity pulsed ultrasound for bone healing: a systematic review of randomised controlled trials

Stefan Schandelmaier, *methodologist*[1,2]; Alka Kaushal, *physician*[1,3]; Lyubov Lytvyn, *methodologist*[4]; Diane Heels-Ansdell, *biostatistician*[1]; Reed A.C. Siemieniuk, *methodologist*[1,5]; Thomas Agoritsa s, *assistant professor*[1,6]; Gordon H Guyatt, *distinguished professor*[1,7]; Per O Vandvik, *associate professor*[8,9]; Rachel Couban, *medical librarian*[3]; Brent Mollon, *orthopedic surgeon*[10]; Jason W. Busse, *associate professor*[1,3,11]

1. Department of Clinical Epidemiology & Biostatistics, McMaster University, 1280 Main St West, Hamilton, Ontario, Canada L8S 4L8
2. Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Spitalstrasse 12, CH-4031 Basel, Switzerland
3. Michael G. DeGroote Institute for Pain Research and Care, McMaster University, Hamilton, ON L8S 4K1, Canada
4. Oslo University Hospital, Forskningsveien 2b, Postboks 1089, Blindern 0317 Oslo, Norway
5. Department of Medicine, University of Toronto, 200 Elizabeth Street, Toronto, Ontario, Canada M5G 2C4
6. Division General Internal Medicine & Division of Clinical Epidemiology, University Hospitals of Geneva, Rue Gabrielle-Perret-Gentil 4, CH-1211, Geneva, Switzerland
7. Department of Medicine, McMaster University, 1280 Main St West, Hamilton, Ontario, Canada, L8S 4L8, Canada
8. Institute of Health and Society, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway
9. Department of Medicine, Innlandet Hospital Trust-division, Gjøvik, Norway
10. Orillia Soldiers' Memorial Hospital; 41 Frederick Street, Orillia, Ontario, Canada L3V 5W6
11. Department of Anesthesia, McMaster University, Hamilton, ON L8S 4K1, Canada

Word count: 4148

Keywords: systematic review, meta-analysis, randomised controlled trials, fracture healing, orthopaedic procedures, ultrasonic waves, ultrasonic therapy, patient-reported outcomes

Correspondence
Stefan Schandelmaier, MD, MSc, PhD student
Department of Clinical Epidemiology & Biostatistics
McMaster University, 1280 Main St W, Hamilton, Ontario, Canada L8S 4L8
schandes@mcmaster.ca

ABSTRACT

Objective: To determine the efficacy of low intensity pulsed ultrasound (LIPUS) for healing of fracture or osteotomy.

Design: Systematic review and meta-analysis**.**

Data sources: MEDLINE, EMBASE, CINAHL, the Cochrane Central Register of Controlled Trials, and trial registries up to November 2016.

Study selection: Randomised controlled trials (RCTs) comparing LIPUS to sham device or no device in patients with any kind of fracture or osteotomy.

Review methods: Two independent reviewers identified studies, extracted data, and assessed risk of bias. A parallel guideline committee (*BMJ* Rapid Recommendation) provided input on the design and interpretation of the systematic review, including selection of patient-important outcomes. We assessed the quality of evidence using GRADE.

Results: We included 26 RCTs with a median sample size of 30 (range 8 to 501). The most trustworthy evidence came from four trials at low risk of bias including patients with tibia or clavicle fractures. Compared with control, LIPUS does not reduce time to return to work (percent difference: 2.7% later with LIPUS, 95% confidence interval [CI] 7.7% earlier to 14.3% later, moderate certainty) or the number of subsequent operations (risk ratio: 0.80, 95% CI 0.55 to 1.16, moderate certainty). For pain, days to weight bearing, and radiographic healing, effects varied substantially between studies. For all three outcomes, trials at low risk of bias failed to demonstrate a benefit with LIPUS, while trials at high risk of bias suggested a benefit (interaction $p<0.001$). Considering only low risk of bias trials, LIPUS does not reduce days to weight bearing (4.8% later, 95% CI 4.0% earlier to 14.4 % later, high certainty), pain at 4 to 6 weeks (mean difference on 0-100 visual analogue scale: 0.94 lower, 95% CI 2.54 lower to 0.65

higher, high certainty), and days to radiographic healing (1.7% earlier, 95% CI 11.2% earlier to 8.8% later, moderate certainty).

Conclusions: Based on moderate to high quality evidence from studies in patients with fresh fracture, LIPUS does not improve patient-important outcomes and probably has no effect on radiographic bone healing.

Systematic review registration: PROSPERO CRD42016050965

What is already known of this topic?

Low intensity pulsed ultrasound (LIPUS) devices are marketed worldwide to accelerate recovery from a fracture or osteotomy.

Previous systematic reviews provided no definite conclusions about the effect of LIPUS on patient-important outcomes and radiographic healing.

What this study adds

A guideline panel including patients and clinical experts informed outcome selection, importance of outcomes, subgroup analyses, and interpretation of results.

Subgroup analyses suggested that beneficial effects of LIPUS are restricted to trials at high risk of bias.

With inclusion of the recently published TRUST trial, sufficient high quality data for patients with fresh fractures has accumulated to conclude that LIPUS fails to improve patient-important outcomes and radiographic healing.

INTRODUCTION

For over 20 years, patients have used low intensity pulsed ultrasound (LIPUS) as an adjunct therapy to improve bone healing. Based on radiographic outcomes, the US Food and Drug Administration and the UK National Institute for Health and Care Excellence NICE have approved LIPUS for fracture healing.(1, 2) Depending on country and device model, LIPUS devices currently cost between £1000-4000. In 2008, 45% of Canadian trauma surgeons prescribed bone stimulators to manage tibia fractures, equally split between LIPUS and electrical stimulation (21% each).(3) Sales from LIPUS amounted to approximately $250 million in 2006 in the US alone.(3, 4)

Within the last seven years, 10 systematic reviews have assessed the effectiveness of LIPUS for bone healing.(5-14) Because existing randomised controlled trials (RCTs) were limited by small sample size, risk of bias, inconsistent results, and failure to address patient-important outcomes, no review offered definitive conclusions. All reviews identified the need for additional RCTs. In addition, recent reviews used suboptimal strategies for outcome selection, data synthesis analysis, and interpretation, leading to potentially misleading conclusions. For instance, the most recent systematic review, published in the top speciality journal for orthopaedic surgeons, considered radiographic union a “critically important outcome” and did not assess the effect of LIPUS on the patient-important outcomes of pain relief or re-operation. Their conclusion that “LIPUS treatment effectively reduces the time to radiographic fracture union” is questionable because it is based on the pooled absolute difference in days to healing, which does not account for the large variation in healing time, showed high unexplained heterogeneity (I^2= 94%), and was driven by studies at high risk of bias. This positive conclusion has the potential to expand the already considerable use of a potentially ineffective therapy.

This systematic review is part of the *BMJ* Rapid Recommendations project, a collaborative effort from the MAGIC research and innovation program (www.magicproject.org) and *The BMJ.* The aim of the project is to respond to new potentially practice-changing evidence and provide a trustworthy practice guideline in a timely manner.(15) In this case, the publication of the TRUST trial,(16) a multicentre trial that randomised 501 patients with tibia fractures and has cast doubt on the effectiveness of LIPUS, initiated the process. This systematic review informed a parallel guideline published in a multi-layered electronic format on *The BMJ*(17) and MAGICapp (https://www.magicapp.org/public/guideline/mL6yYj).

Our objective was to assess whether LIPUS compared to sham device or no device improves patient important outcomes and radiographic healing in patients with any kind of fracture or osteotomy.

METHODS

Guideline panel and patient involvement

According to the *BMJ* Rapid Recommendations process,(15) a guideline panel provided critical oversight to the review and identified populations, subgroups, and outcomes of interest. The panel included six content experts (five orthopaedic or trauma surgeons and one physiotherapist), six methodologists (four of whom are also front-line clinicians), and four patients with personal experience of fractures (one of whom had used LIPUS). All patients received personal training and support to optimise contributions throughout the guideline development process. The patient panel members led the interpretation of the results based on what they expected the typical patient values and preferences to be, as well as the variation between patients.

Information sources

We searched MEDLINE, PubMed, EMBASE, CINAHL, and the Cochrane Central Register of Controlled Trials up to 16 November 2016, using a combination of keywords and MeSH terms for fracture, orthopaedic surgical procedures, and ultrasound. Additional searches included trials registries clinicaltrials.gov and isrctn.com. An experienced research librarian designed the search strategies (appendix 1). Two independent reviewers scanned the references from eligible studies, related systematic reviews, and all studies citing eligible RCTs on Google Scholar.

Study selection

We included RCTs comparing LIPUS to a sham device or no device in patients with any type of fracture regardless of location (long-bone or other bone), type (fresh fracture, delayed union, non-union, or stress fracture), or clinical management (operative or non-operative). We included

any type of osteotomy, including distraction osteogenesis. We excluded trials published only as protocol or abstract if attempts to get the final results from investigators were unsuccessful.

Two reviewers, independently and in duplicate, screened the titles and abstracts of identified articles and acquired the full text of any article that either reviewer judged to be potentially eligible. They independently applied the eligibility criteria to the full texts and, when consensus could not be reached, resolved disagreements through discussion or adjudication by a third reviewer.

Data collection

Two reviewers used standardised forms to independently abstract data; they resolved disagreements by discussion or involved a third reviewer when required. Extracted data included patient characteristics, fracture characteristics, clinical management, risk of bias, intervention details, statements about compliance with treatment, and outcomes.

Risk of bias assessment

Two reviewers independently assessed risk of bias using a modified Cochrane risk of bias instrument that includes response options of "definitely or probably yes" (assigned a low risk of bias) or "definitely or probably no" (assigned a high risk of bias), an approach we have previously validated.(18) On the study level, we assessed generation of randomisation sequence, concealment of allocation, blinding of patients, caregivers, and outcome reporting (by comparing each publication with their corresponding published protocol, when available). For each outcome within studies, we assessed blinding of outcome assessors, loss to follow-up, and additional limitations. We considered ≥20% loss-to follow-up to represent a high risk of bias unless the investigators performed appropriate sensitivity analyses demonstrating the robustness of the

results. As a sensitivity analysis, we alternatively considered a more conservative threshold of ≥10% loss to follow-up. We categorised a trial as being at low risk of bias for a particular outcome if we identified no limitation for any risk of bias item.

Outcomes

Patients identified functional recovery (time to return to work and time to full weigh bearing), pain reduction, and number of subsequent fracture or osteotomy related operations (re-operation for operatively managed fracture and osteotomy) as the most important outcomes for patients considering LIPUS for bone healing. Because many clinicians currently base their management on time to radiographic healing, a surrogate outcome important only insofar as it influences patient experience, the panel requested its inclusion in our review. We extracted all outcomes that fell into these categories as well as ultrasound device-related adverse effects.

Synthesis of results

We pooled treatment effects of LIPUS on similar outcomes across eligible trials, regardless of clinical subgroups, focusing on complete case analysis. We calculated pooled estimates and associated 95% confidence intervals (CI) using random effects models for meta-analysis with three or more studies, and fixed-effects models for meta-analysis with two studies. We examined heterogeneity associated with all pooled analyses using both the X^2 test and I^2 statistic. SAS version 9.4, R version 3.1, and Review Manager 5.3 provided software for the statistical analysis.

For time-to-event outcomes, we pooled hazard ratios. For studies that did not apply methods of survival analysis, we considered time to event reported as a continuous variable (e.g. days to return to work) at the longest follow-up time. We used the relative effect measure ratio of means (mean LIPUS/mean control) in order to account for the baseline difference in fracture healing

depending on type of bone and (e.g. scaphoid, clavicle, tibia) and fracture or procedure (e.g. stress fracture or distraction osteogenesis). We pooled the natural logarithm of the ratio of means and presented the results as percentage difference (relative change). For studies that reported the proportion of patients who achieved the event at a specific time point, we calculated risk ratios.

When studies used different instruments to measure the same construct on a continuous scale, we converted all instruments to the most commonly used instrument among studies and then pooled results using the weighted mean difference.(19)

For the outcomes number of subsequent operations and device related adverse events, we calculated both risk ratios, which are preferable in case of varying baseline risks, and risk differences, which allow inclusion of studies with zero events in both groups.

In consultation with the expert and patient guideline panel, we pre-specified three subgroup hypotheses to explain heterogeneity of effects between studies: (1) LIPUS will show larger effects in high risk of bias studies, (2) LIPUS effects will differ based on clinical subgroups, and (3) LIPUS will show larger effects with greater patient compliance. In consultation with the six clinical experts on the parallel guideline panel, we classified eligible RCTs according to the following five clinical subgroups: (1) operatively managed fresh fractures, (2) non-operatively managed fresh fractures, (3) stress fractures, (4) non-union, and (5) osteotomy (including distraction osteogenesis). Because compliance was reported inconsistently, two reviewers independently categorised trials using response options of "definitely or probably high compliance" or "definitely or probably moderate compliance" using as a guide a definition of high compliance as at least 80% of patients applied LIPUS for at least 80% of the total time prescribed. We conducted univariable tests of interaction to establish if the effect size from the

subgroups differed significantly from each other, and, in order to test independence of subgroup effects, performed multivariable meta-regression in which we included risk of bias (high versus low), compliance with LIPUS treatment (high versus moderate), and clinical subgroups (as above) as independent variables in a single model.

Only one outcome, days to radiographic healing, included enough studies to perform all planned subgroup analysis. As a rule of thumb we had pre-specified in our protocol at least three studies per group. We assessed the credibility of significant subgroup effects using the criteria suggested by Sun et al.(20) Based on the finding that risk of bias appeared to independently explain the high heterogeneity in the outcome days to radiographic healing, we performed subgroup analysis by risk of bias for all outcomes.

The authors and the guideline panel achieved consensus in categorising the quality of evidence for all reported outcomes as high, moderate, low, or very low using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. In the GRADE approach, RCTs begin as high quality evidence but can be rated down due to: (1) risk of bias; (2) inconsistency; (3) indirectness; (4) imprecision; or, (5) publication bias.(21) We considered rating down for inconsistency if the magnitude and direction of effects were dissimilar, the confidence intervals had minimal overlap, the test of heterogeneity was significant, or the I^2 was high.(22) For outcomes with ten or more studies, we inspected symmetry of funnel plots and performed Egger's statistical test for publication bias.(23)

To calculate absolute effects, we applied the effect estimate from the meta-analysis to the control arm of the TRUST trial, which enrolled patients with tibia fractures and had the largest sample

size of any eligible study that was at low risk of bias. The approach to rating certainty of individual outcomes was fully contextualised: that is, in rating quality about any individual outcome, we took into account the findings on the other outcomes.

RESULTS

Search results

We identified 3489 potentially eligible abstracts, retrieved 42 studies in full text, and found 26 eligible RCTs (fig 1).(16, 24-50) Two RCTs, Handolin et al.(30, 31) and Emami et al.,(27, 28) provided two publications reporting on the same group of patients. There were no shared patients between the TRUST pilot (24) and the definitive trial.(16) Our registry search yielded four protocols of potentially eligible RCTs; one was discontinued due to slow recruitment (ISRCTN90844675, personal communication, outcome data not available yet), one manuscript is under peer-review (NCT00744861, personal communication: “no difference between the control group and the ultrasound group”), one is completed but unpublished (JPRN-UMIN000002005, no response from investigators), and the last is still ongoing (NCT02383160). Attempts to acquire the full text of another potentially eligible RCT,(51) reported in a recent systematic review,(11) were unsuccessful.

Study characteristics

Eligible trials enrolled patients with operatively managed fresh fractures (n=7); non-operatively managed fresh fractures (n=6); stress fractures (n=2); non-unions (n=3); and osteotomies (n=8), of which five were distraction osteogenesis (table 1). Most trials enrolled patients with tibia fractures or osteotomies (n=14). All but two trials applied LIPUS for 20 minutes every day either for a fixed period or until radiographic healing. Otherwise, one trial applied LIPUS for 15

minutes per day,(36) and another trial for 5 minutes every second day.(39) Fifteen RCTs (60%) provided their control group with an inactive device that was indistinguishable from the active LIPUS. Only three trials (12%) were explicitly free of industry funding.(26, 42, 48)

Risk of bias

We contacted authors to resolve areas of uncertainty and successfully clarified details in five RCTs.(30, 31, 35, 37, 40) We considered six trials to be at low risk of bias,(16, 24, 27, 37, 46, 47) and the remaining 20 studies to be at high risk of bias (table 2). The main limitations were failure to report a method for allocation concealment (15 RCTs), unblinded patients (10 RCTs), caregivers or outcome assessors (10 RCTs), and high or unclear numbers of patients excluded from the analysis (13 RCTs; table 2).

Outcomes

Table 3 summarises findings of all outcomes. Interactive summary of findings tables are available online at https://www.magicapp.org/public/guideline/mL6yYj.

Functional recovery

Only the TRUST trial assessed time to return to work using a time-to-event analysis, and found no significant effect (hazard ratio 1.11 favouring control, 95% CI 0.82 to 1.50; 343 patients).(16) Three trials assessed the number of days to return to work; the pooled effect was not significant (2.7% later return with LIPUS, 95% CI 7.7% earlier to 14.3% later; I^2=0%; 392 patients) (fig 2). We found no significant interaction with risk of bias (p=0.86). Considering an alternative threshold of ≥10% loss-to follow-up for assessing risk of attrition bias, all three studies would fall into the category of high risk of bias. However, given the consist absence of effects this would not lower our confidence in the result. A fourth trial in patients with delayed union of tibia fracture provided insufficient data for inclusion in meta-analysis (table 2), but reported no significant difference in days to return to work.(50)

Only the TRUST trial assessed time to full weight bearing using a time-to event analysis, and found no significant effect (hazard ratio 0.87 in favour of LIPUS, 95% CI 0.70 to 1.08; 451 patients). Three trials assessed the number of days to full weight bearing. Overall results suggested no significant effect on full weight bearing with LIPUS but high heterogeneity (I^2=95%). The effect of the one trial at high risk of bias (40.0% earlier, 95% CI 48.4% to 30.3 earlier) differed significantly from the consistent results from the two trials at low risk of bias (4.8% later, 95% CI 4.0% earlier to 14.4% later; 483 patients; interaction p<0.001, subgroup effect not effected by alternative threshold for missing data) (fig 3).

Appendix 2 presents results of other functional outcomes including return to leisure activities, return to household activities, return to pre-injury level of function, and physical function measured with a multidimensional questionnaire. None of these were significantly affected by use of LIPUS, nor did they show substantial inconsistency.

Pain reduction

Four trials assessed pain, two using a 100mm visual analogue scale(37, 49) and two using the subdomain "bodily pain" of the SF-36 instrument.(16, 24) After transforming all results to a 100mm visual analogue scale, findings at 3 to 6 weeks follow-up showed no significant effect of LIPUS on pain reduction but high heterogeneity (I^2=97%). The effect of the one trial at high risk of bias (28.12 mm lower, 95% CI, 37.05 to 19.19 lower) differed significantly from the consistent results from the three trials at low risk of bias (0.93 mm lower, 95% CI 2.51 lower to 0.64 higher; 626 patients; I^2=0%; interaction p<0.001; fig 4). The subgroup effect was no longer significant when we used a threshold of ≥10% missing data to designate a trial at high risk of attrition bias (p=0.35, fig 4 in Appendix 4). Two other small studies assessed pain intensity at 5

months but could not be included in the meta-analysis. One reported pain outcomes only narratively (no effect),(41) another used a modified instrument with unclear scale and variance (no effect).(49)

Other outcomes for pain included pain intensity assessed at multiple time-points and number of painful days (appendix 3). None showed a significant effect of LIPUS, nor substantial inconsistency.

Number of subsequent operations

Ten trials reported the number of subsequent operations including three trials reporting zero events in both arms. Neither the pooled risk ratio (0.8 in favour of LIPUS, 95% CI 0.55 to 1.16; I^2=0%; 7 trials, 693 patients; fig 5) nor the pooled risk difference (3% reduction with LIPUS, 95% CI 7% reduction to 2% increase; I^2=0%; 10 trials, 740 patients) showed a significant effect. There was no significant interaction with risk of bias on either scale (risk ratio: p=0.75; risk difference: p=0.64. The results did not depend on the threshold for missing data).

Time to radiographic healing

Two trials used time-to-event analysis methods to assess time to radiographic healing,(16, 24) and showed no significant effect of LIPUS (hazard ratio 1.06 in favour of control, 95% CI 0.86 to 1.32; I^2=0%; 532 patients). Fifteen trials reported the number of days to radiographic healing. Overall results suggested accelerated radiographic healing with LIPUS (26% earlier, 95% CI 33.6% to 17.8% earlier; I^2=84.7%). The effect differed significantly between the 12 trials at high risk of bias (31.8% earlier; 95%CI 38.6% to 24.3% days earlier; I^2=77.8%; 446 patients) and the three trials at low risk of bias (1.7% earlier, 95% CI 11.2% earlier to 8.8% later, I^2=9.8%; 483 patients; interaction p<0.001; fig 6). This subgroup effect fulfilled 8 of 9 credibility criteria

relevant to risk of bias as an explanation of heterogeneity (table 4). In addition, the subgroup effect was robust to our sensitivity analysis using a more conservative threshold for defining risk of attrition bias (interaction p=0.004, fig 5 in appendix 4). The effect of LIPUS on days to radiographic healing did not differ significantly across clinical subgroups (p=0.13, fig 1 in appendix 4) or between high and moderate compliance with treatment (p=0.79, fig 2 in appendix 4). In our multivariable meta-regression, which included risk of bias, clinical subgroups, and compliance with treatment, the only significant effect modifier was the risk of bias (p=0.005).

Another RCT in patients with delayed union of tibia fracture reported only the proportion of healed fractures at 16 weeks and did not find a significant difference (65% in the LIPUS and 46% in the control arm, p=0.07; high risk of bias towards LIPUS due to serious imbalance in age of fracture at baseline).(44)

The funnel plot based on time to radiographic healing was not clearly asymmetrical and Egger's test for publication bias was not significant (p=0.25, fig 3 in appendix 4).

Device related adverse effects

Seven studies reported explicitly the absence of any device-related adverse effects; two other studies reported mild transient skin irritations in 6 of patients. The pooled risk ratio based on these two studies (2.65 in favour of control, 95% CI 0.32 to 22.21; 129 patients) was not significant, nor was the pooled risk difference based on all nine trials (0%, 95% CI 1% reduction to 1% increase; I^2=0%; 839 patients; fig 7). We found no significant interaction with risk of bias on the risk difference scale (p=0.75).

DISCUSSION

Main findings

Our systematic review demonstrated moderate quality evidence that LIPUS applied to patients with fractures or osteotomies has no effect on time to return to work or the number of subsequent operations (table 3). Overall results suggested a possible reduction of days to full weight bearing, pain, and days to radiographic healing, but with large variability between studies strongly associated with risk of bias as an effect modifier: only trials with high risk of bias demonstrated benefit. Based on RCTs at low risk of bias, we found high quality evidence that LIPUS has no effect on pain reduction, days to full weight bearing, or device-related adverse effects, and moderate quality evidence that LIPUS has no effect on days to radiographic healing (table 3).

Comparison with other systematic reviews

Our results are consistent with other systematic reviews in concluding that most RCTs addressing LIPUS therapy are poorly reported, lack patient important outcomes, and are at high risk of bias.(5-14) Our systematic review, however, differs from previous systematic reviews in several important aspects. First, we include the recently published TRUST trial,(16) by far the largest trial addressing LIPUS therapy for bone healing, which reported a number of patient-important outcomes. Second, our choice of outcomes and interpretation of findings was informed by a guideline panel including patients with personal experience of fractures in the context of BMJ Rapid Recommendations. Patients considered functional recovery, pain reduction and operations as critical outcomes, while expressing little interest in the commonly reported surrogate outcome of radiographic healing. Third, we used optimal statistical approaches, and in particular the ratio of means to combine days to radiographic healing, return to work, or full weight bearing across studies. This relative effect measure is most appropriate in the context of

LIPUS where the average time to recovery differs substantially between clinical subgroups. For instance, a lower grade stress fracture is likely to heal much faster than a complicated tibia fracture. It is not surprising, therefore, that previous meta-analyses found high heterogeneity when they used absolute mean differences to pool across studies.(8, 11, 12)

Finally, we used the GRADE approach to assess the quality of evidence, taking into account the results of subgroup analysis based on risk of bias: when effects differed significantly between high and low quality trials, we based our conclusions on trials at low risk of bias. Our approach of limiting conclusions to low risk of bias trials depends on our judgement of risk of bias; however, our ratings of risk of bias were consistent with those of a previous Cochrane systematic review.(5) Further, most trials judged to be at high risk of bias had limitations in more than one domain, and some had additional sources of bias including baseline imbalance or unclear clustering when patients had more than one fracture or surgery. Applying our risk of bias judgments as an effect modifier met 8 of 9 relevant criteria for a credible subgroup analysis (table 4). A post-hoc sensitivity analysis exploring a more conservative threshold for attrition bias (≥10% loss to follow-up) yielded, for all outcomes, results essentially consistent with the primary analyses

Limitations

The primary limitation of our review is the failure of most trials to measure or report patient-important outcomes. Of the 26 eligible trials, 11 reported, in sufficient detail for inclusion in meta-analysis, outcomes that patient consider critical for decision making.(16, 24, 25, 27, 29-31, 35, 37, 39, 46, 47) Of these, the only four trials that contributed substantial data were either conducted in patients with operatively managed fresh tibia fracture(16, 24, 27) or conservatively

managed clavicle fracture.(37) One could question the extent to which our results apply to patients not included at all (such as children) or underrepresented (stress fractures, non-union, and osteotomies) in the eligible trials. Qualitative subgroup effects (e.g. no benefit in one subgroup and important benefit in another) are, however, unusual. In the absence of evidence to the contrary, it is therefore reasonable to apply our results to these populations. Our subgroup analysis and meta-regression for radiographic healing found no effect modification based on clinical subgroups. Certainly, the burden of proof regarding the effect of LIPUS in children and underrepresented populations rests with those who might postulate a benefit.

LIPUS compared with electrical stimulation

Our findings are similar to a 2016 systematic review of 15 small trials that explored electrical stimulation vs. sham therapy for fracture healing; only 4 of which were at low risk of bias.(52) This review found moderate quality evidence for a 35% reduction (95% CI 19% to 47%; I^2=46%) in the rate of radiographic nonunion. The authors found no evidence of a subgroup effect based on clinical presentation (i.e. fresh fractures, delayed union or nonunion, spinal fusion, or surgical osteotomy; interaction p = 0.41) – they did not explore whether risk of bias explained heterogeneity, but all 4 trials at low risk of bias showed no significant effect on radiographic union.(53-56) This review found a small reduction in pain (mean difference of −7.7 mm on a 100mm visual analogue scale for pain, 95% CI −13.92 to −1.43), and low quality evidence for no difference in functional outcome (mean difference of −0.88 points on the 100 point Short Form 36 Physical Component Summary score, 95% CI −6.63 to 4.87).

Conclusions

In conclusion, moderate to high quality evidence demonstrates that LIPUS fails to accelerate return to work, return to full weight bearing, pain, or the need for subsequent operation. If one

gives highest credibility to combined effects from all available RCTs, low quality evidence would suggest a large reduction in time to radiographic healing. If, however, one gives higher credence to low risk of bias trials, moderate to high quality evidence suggests that LIPUS not only has no effect on patient-important outcomes, but also fails to accelerate radiographic healing. The evidence applies directly to patients with fresh fractures and indirectly to children and other underrepresented populations, particularly those with non-union, for which no trustworthy direct evidence exists.

Acknowledgements

We thank members of the Rapid Recommendations panel for critical feedback on outcome and subgroup selection, GRADE judgments, and manuscript feedback, including Rudolf Poolman (chair and orthopaedic surgeon), Ian Harris (orthopaedic and trauma surgeon), Inger Schipper (trauma surgeon), Maureen Smith (patient representative), Alexandra Albin (patient representative), Sally Nador (patient representative), William Sasges (patient representative), Ton Kuijpers (methodologist), Loes van Beers (physiotherapist), and Michael Verhofstad (trauma surgeon).

Competing interest

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work. We acknowledge that JWB, DHA, and GHG were co-authors of the TRUST trial, which was supported in part by an industry grant from Smith & Nephew, a manufacturer of LIPUS devices. No other relationships or activities that could appear to have influenced the submitted work.

Funding

This study was unfunded.

Contributors

JWB, RAS, GHG and POV conceived the study idea. SS and JWB coordinated the systematic review. SS wrote the first draft of the manuscript. RC designed the search strategy. LL, AK, RC, and SS screened abstracts and full texts. LL, AK, RAS, TA, and SS acquired the data and judged risk of bias in the studies. SS, JWB and DHA performed the data analysis. DHA and GHG

provided statistical advice. SS, RAS, POV, JWB and GHG interpreted the data analysis. All authors critically revised the manuscript. SS had full access to all of the data in the study, and takes responsibility for the integrity of the data and the accuracy of the data analysis. SS is the guarantor.

Ethical approval

Not required.

Data sharing

All data informing the study is freely available in the appendices.

Data access

All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Patient involvement

Four patient representatives were full members of the guideline, and contributed to the selection and prioritisation of outcomes, values and preferences assessments, and critical feedback to the protocol for the systematic review and the BMJ Rapid Recommendations manuscript.

Transparency declaration

SS is guarantor and affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

REFERENCES

1. National Institute for Helacare Excellence. Low-intensity pulsed ultrasound to promote fracture healing. 2010. URL: https://www.nice.org.uk/guidance/ipg374/history.
2. Food and Drug Administration. Approval order for Exogen device. 2000, URL: http://www.fda.gov/ohrms/dockets/dailys/00/mar00/031300/aav0001.pdf.
3. Busse JW, Morton E, Lacchetti C, Guyatt GH, Bhandari M. Current management of tibial shaft fractures: a survey of 450 Canadian orthopedic trauma surgeons. Acta Orthop. 2008;79:689-94.
4. Wachovia Capital Markets. Equity research: bone growth stimulation 2008 outlook. 2007.
5. Griffin XL, Parsons N, Costa ML, Metcalfe D. Ultrasound and shockwave therapy for acute fractures in adults. Cochrane Database Syst Rev. 2014:CD008579.
6. Ebrahim S, Mollon B, Bance S, Busse JW, Bhandari M. Low-intensity pulsed ultrasonography versus electrical stimulation for fracture healing: a systematic review and network meta-analysis. Can J Surg. 2014;57:E105-18.
7. Bashardoust Tajali S, Houghton P, MacDermid JC, Grewal R. Effects of low-intensity pulsed ultrasound therapy on fracture healing: a systematic review and meta-analysis. Am J Phys Med Rehabil. 2012;91:349-67.
8. Hannemann PF, Mommers EH, Schots JP, Brink PR, Poeze M. The effects of low-intensity pulsed ultrasound and pulsed electromagnetic fields bone growth stimulation in acute fractures: a systematic review and meta-analysis of randomized controlled trials. Arch Orthop Trauma Surg. 2014;134:1093-106.
9. Martinez de Albornoz P, Khanna A, Longo UG, Forriol F, Maffulli N. The evidence of low-intensity pulsed ultrasound for in vitro, animal and human fracture healing. Br Med Bull. 2011;100:39-57.
10. Raza H, Saltaji H, Kaur H, Flores-Mir C, El-Bialy T. Effect of Low-Intensity Pulsed Ultrasound on Distraction Osteogenesis Treatment Time: A Meta-analysis of Randomized Clinical Trials. J Ultrasound Med. 2016;35:349-58.
11. Rutten S, van den Bekerom MP, Sierevelt IN, Nolte PA. Enhancement of Bone-Healing by Low-Intensity Pulsed Ultrasound: A Systematic Review. JBJS Rev. 2016;4.
12. Snyder BM, Conley J, Koval KJ. Does low-intensity pulsed ultrasound reduce time to fracture healing? A meta-analysis. Am J Orthop (Belle Mead NJ). 2012;41:E12-9.
13. Watanabe Y, Matsushita T, Bhandari M, Zdero R, Schemitsch EH. Ultrasound for fracture healing: current evidence. J Orthop Trauma. 2010;24 Suppl 1:S56-61.
14. Busse JW, Kaur J, Mollon B, et al. Low intensity pulsed ultrasonography for fractures: systematic review of randomised controlled trials. BMJ. 2009;338:b351.
15. Siemieniuk RA, Agoritsas T, Macdonald H, Guyatt GH, Brandt L, Vandvik PO. Introduction to BMJ Rapid Recommendations. BMJ. 2016;354:i5191.
16. TRUST Investigators writing group, Busse JW, Bhandari M, Einhorn TA, et al. Re-evaluation of low intensity pulsed ultrasound in treatment of tibial fractures (TRUST): randomized clinical trial. BMJ. 2016;355:i5351.
17. Poolman RW, Agoritsas T, Siemieniuk RAC, et al. Low intensity pulsed ultrasound (LIPUS) for bone healing: a clinical practice guideline. Submitted to The BMJ 2016.
18. Akl EA, Sun X, Busse JW, et al. Specific instructions for estimating unclearly reported blinding status in randomized trials were reliable and valid. J Clin Epidemiol. 2012;65:262-7.

19. Thorlund K, Walter SD, Johnston BC, Furukawa TA, Guyatt GH. Pooling health-related quality of life outcomes in meta-analysis-a tutorial and review of methods for enhancing interpretability. Res Synth Methods. 2011;2:188-203.
20. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. BMJ. 2010;340:c117.
21. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336:924-6.
22. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. J Clin Epidemiol. 2011;64:1294-302.
23. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629-34.
24. Busse JW, Bhandari M, Einhorn TA, et al. Trial to re-evaluate ultrasound in the treatment of tibial fractures (TRUST): a multicenter randomized pilot study. Trials. 2014;15:206.
25. Dudda M, Hauser J, Muhr G, Esenwein SA. Low-intensity pulsed ultrasound as a useful adjuvant during distraction osteogenesis: a prospective, randomized controlled trial. Journal of Trauma-Injury Infection & Critical Care. 2011;71:1376-80.
26. El-Mowafi H, Mohsen M. The effect of low-intensity pulsed ultrasound on callus maturation in tibial distraction osteogenesis. Int Orthop. 2005;29:121-4.
27. Emami A, Petren-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.
28. Emami A, Larsson A, Petren-Mallmin M, Larsson S. Serum bone markers after intramedullary fixed tibial fractures. Clin Orthop Relat Res. 1999:220-9.
29. Gan TY, Kuah DE, Graham KS, Markson G. Low-intensity pulsed ultrasound in lower limb bone stress injuries: a randomized controlled trial. Clin J Sport Med. 2014;24:457-60.
30. Handolin L, Kiljunen V, Arnala I, et al. Effect of ultrasound therapy on bone healing of lateral malleolar fractures of the ankle joint fixed with bioabsorbable screws. J Orthop Sci. 2005;10:391-5.
31. Handolin L, Kiljunen V, Arnala I, Pajarinen J, Partio EK, Rokkanen P. The effect of low intensity ultrasound and bioabsorbable self-reinforced poly-L-lactide screw fixation on bone in lateral malleolar fractures. Arch Orthop Trauma Surg. 2005;125:317-21.
32. Handolin L, Kiljunen V, Arnala I, et al. No long-term effects of ultrasound therapy on bioabsorbable screw-fixed lateral malleolar fracture. Scand J Surg. 2005;94:239-42.
33. Heckman JD, Ryaby JP, McCabe J, Frey JJ, Kilcoyne RF. Acceleration of tibial fracture-healing by non-invasive, low-intensity pulsed ultrasound. J Bone Joint Surg Am. 1994;76:26-34.
34. 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.
35. Leung KS, Lee WS, Tsui HF, Liu PP, Cheung WH. Complex tibial fracture outcomes following treatment with low-intensity pulsed ultrasound. Ultrasound Med Biol. 2004;30:389-95.
36. Liu Y, Wei X, Kuang Y, et al. Ultrasound treatment for accelerating fracture healing of the distal radius. A control study. Acta Cir Bras. 2014;29:765-70.
37. Lubbert PH, van der Rijt RH, Hoorntje LE, van der Werken C. Low-intensity pulsed ultrasound (LIPUS) in fresh clavicle fractures: a multi-centre double blind randomised controlled trial. Injury. 2008;39:1444-52.
38. Mayr E, Rudzki MM, Rudzki M, Borchardt B, Hausser H, Ruter A. Does pulsed low-intensity ultrasound accelerate healing of scaphoid fractures?. [German]. Handchirurgie Mikrochirurgie Plastische Chirurgie. 2000;32:115-22.

39. Patel K, Kumar S, Kathiriya N, et al. An Evaluation of the Effect of Therapeutic Ultrasound on Healing of Mandibular Fracture. Craniomaxillofac Trauma Reconstr. 2015;8:299-306.
40. Ricardo M. The effect of ultrasound on the healing of muscle-pediculated bone graft in scaphoid non-union. Int Orthop. 2006;30:123-7.
41. Rutten S, Klein-Nulend J, Guit GL, et al. Use of low-intensity pulsed ultrasound stimulation of delayed unions of the osteotomized fibula: a prospective randomized double-blind trial (Thesis). Low-intensity pulsed ultrasound treatment in delayed bone healing [thesis]. Amsterdam: Vrije Universiteit Amsterdam; 2012.
42. Rue JP, Armstrong DW, 3rd, Frassica FJ, Deafenbaugh M, Wilckens JH. The effect of pulsed ultrasound in the treatment of tibial stress fractures. Orthopedics. 2004;27:1192-5.
43. Salem KH, Schmelz A. Low-intensity pulsed ultrasound shortens the treatment time in tibial distraction osteogenesis. International Orthopaedics. 2014;38:1477-82.
44. Schofer MD, Block JE, Aigner J, Schmelz A. Improved healing response in delayed unions of the tibia with low-intensity pulsed ultrasound: results of a randomized sham-controlled trial. BMC Musculoskelet Disord. 2010;11:229.
45. Kamath JB, Jayasheelan N, Reddy B, Muhammed S, Savur A. The effect of low-intensity pulsed ultrasound therapy on fracture healing. Muller Journal of Medical Sciences and Research. 2015;6:49-53.
46. Schortinghuis J, Bronckers AL, Stegenga B, Raghoebar GM, de Bont LG. Ultrasound to stimulate early bone formation in a distraction gap: a double blind randomised clinical pilot trial in the edentulous mandible. Arch Oral Biol. 2005;50:411-20.
47. Schortinghuis J, Bronckers AL, Gravendeel J, Stegenga B, Raghoebar GM. The effect of ultrasound on osteogenesis in the vertically distracted edentulous mandible: a double-blind trial. Int J Oral Maxillofac Surg. 2008;37:1014-21.
48. Tsumaki N, Kakiuchi M, Sasaki J, Ochi T, Yoshikawa H. Low-intensity pulsed ultrasound accelerates maturation of callus in patients treated with opening-wedge high tibial osteotomy by hemicallotasis. J Bone Joint Surg Am. 2004;86-a:2399-405.
49. Urita A, Iwasaki N, Kondo M, Nishio Y, Kamishima T, Minami A. Effect of low-intensity pulsed ultrasound on bone healing at osteotomy sites after forearm bone shortening. Journal of Hand Surgery - American Volume. 2013;38:498-503.
50. Zacherl M, Gruber G, Radl R, Rehak PH, Windhager R. No midterm benefit from low intensity pulsed ultrasound after chevron osteotomy for hallux valgus. Ultrasound Med Biol. 2009;35:1290-7.
51. Nolte PA, Maas M, Roolker L, Marti RK, Albers GHR. Effect of low-intensity ultrasound on bone healing in osteotomies of the lower extremity: a randomised trial. In: Nolte PA, editor. Nonunions – surgery and low-intensity ultrasound treatment [thesis]. Amsterdam: Universiteit van Amsterdam; 2002. p. 96-106.
52. Aleem IS, Aleem I, Evaniew N, et al. Efficacy of Electrical Stimulators for Bone Healing: A Meta-Analysis of Randomized Sham-Controlled Trials. Sci Rep. 2016;6:31724.
53. Hannemann PF, Gottgens KW, van Wely BJ, et al. The clinical and radiological outcome of pulsed electromagnetic field treatment for acute scaphoid fractures: a randomised double-blind placebo-controlled multicentre trial. J Bone Joint Surg Br. 2012;94:1403-8.
54. Hannemann PF, van Wezenbeek MR, Kolkman KA, et al. CT scan-evaluated outcome of pulsed electromagnetic fields in the treatment of acute scaphoid fractures: a randomised, multicentre, double-blind, placebo-controlled trial. Bone Joint J. 2014;96-B:1070-6.

55. Mammi GI, Rocchi R, Cadossi R, Massari L, Traina GC. The electrical stimulation of tibial osteotomies. Double-blind study. Clin Orthop Relat Res. 1993:246-53.
56. Martinez-Rondanelli A, Martinez JP, Moncada ME, Manzi E, Pinedo CR, Cadavid H. Electromagnetic stimulation as coadjuvant in the healing of diaphyseal femoral fractures: a randomized controlled trial. Colomb Med (Cali). 2014;45:67-71.

TABLES

Table 1, Study characteristics

Author Year	Bone	Type of fracture / surgery	% open fracture	Manage-ment	% women	Mean age	N randomised		Sham device	Dose and duration of LIPUS therapy	Max follow-up	Explicit free of industry funding
							LIPUS	No ultrasound				
Busse 2014(24)	Tibia	Fresh fracture	27%	Operative	24%	40	23	28	Yes	20 min/day to healing*	1 year	No
Busse 2016(16)	Tibia	Fresh fracture	23%	Operative	31%	40	250	251	Yes	20 min/day to healing*	1 year	No
Dudda 2011(25)	Tibia	Distraction osteogenesis	NA	Operative	11%	39	16	20	No	20 min/day to healing*	35 weeks	No
El-Mowafi 2005(26)	Tibia	Distraction osteogenesis	NA	Operative	0%	35	10	10	No	20 min/day to healing*	12 months	Yes
Emami 1999(27, 28)	Tibia	Fresh fracture	13%	Operative	25%	37	15	17	Yes	20 min/day to healing*	20 weeks	No
Gan 2014(29)	Tibia, fibula, metatarsal	Stress fracture	0%	Non-operative	83%	30	15	15	Yes	20 min/day for 28 days	12 weeks	No
Handolin 2005a(30, 31)	Lateral malleolus	Fresh fracture	0%	Operative	47%	42	11	11	Yes	20 min/day for 42 days	12 weeks	No
Handolin 2005b (32)	Lateral malleolus	Fresh fracture	0%	Operative	56%	40	15	15	Yes	20 min/day for 42 days	18 months	No
Heckman 1994(33)	Tibia	Fresh fracture	4%	Non-operative	19%	33	48	49	Yes	20 min/day to healing*	140 days	No
Kamath 2015(45)	Tibia and femur	Fresh fracture	0%	Operative	NR	36	33	27	No	20 min/day for 1 month	16 weeks	No
Kristiansen 1997(34)	Distal radius	Fresh fracture	0%	Non-operative	84%	56	40	45	Yes	20 min/day for 70 days	140 days	No
Leung 2004(35)	Tibia	Fresh fracture	47%	Operative	11%	35	16	14	Yes	20 min/day for 4 months	5 months	No
Liu 2014(36)	Distal	Fresh	NR	Non-	36%	67	41	40	No	15 min/day	At least 12	No

	radius	fracture		operative						for ≥12 weeks	weeks	
Lubbert 2008(37)	Clavicle	Fresh fracture	0%	Non-operative	16%	38	61	59	Yes	20 min/day for 28 days	8 weeks	No
Mayr 2000(38)	Scaphoid	Fresh fracture	0%	Non-operative	17%	37	15	15	No	20 min/day to healing*	120 days	No
Patel 2014(39)	Mandible	Fresh fracture	NR	Non-operative	25%	15-35	14	14	No	5 min q.a.d. for 24 days	5 weeks	No
Ricardo 2006(40)	Scaphoid	Non-union	NA	Operative	0%	27	10	11	Yes	20 min/day to healing*	4 years	No
Rue 2004(42)	Tibia	Stress fracture	0%	Non-operative	50%	19	Probably 20	Probably 20	Yes	20 min/day to healing*	NR	Yes
Rutten 2012(41)	Tibia	Non-union	0%	Operative	70%	41-63	10	10	Yes	20 min/day for 5 months	5 years	No
Salem 2014(43)	Tibia	Distraction osteogenesis	NA	Operative	14%	30	12	9	No	20 min/day to healing*	NR	No
Schofer 2010(44)	Tibia	Non-union	NA	Operative	24%	44	51	50	Yes	20 min/day for 16 weeks	16 weeks	No
Schortinghuis 2005(46)	Mandible	Distraction osteogenesis	NA	Operative	75%	65	4	4	Yes	20 min/day for 4 weeks	30 months	No
Schortinghuis 2008(47)	Mandible	Distraction osteogenesis	NA	Operative	NR	56	5	4	Yes	20 min/day for 6 weeks	44 months	No
Tsumaki 2004(48)	Tibia	Distraction osteogenesis	NA	Operative	81%	68	21 knees	21 knees	No	20 min/day to healing*	NR	Yes
Urita 2013(49)	Ulna and radius	Osteotomy (shortening)	NA	Operative	63%	48	14	13	No	20 min/day to healing* or 12 weeks	24 weeks	No
Zacherl 2009(50)	Hallux valgus	Osteotomy (deformity correction)	NA	Operative	85%	53	26 toes	26 toes	Yes	20 min/day for 42 days	1 year	No

*Until radiographic healing. q.a.d . = every other day

Table 2, Risk of bias

Author Year	Sequence generation adequate	Concealment of treatment allocation	Patients blinded	Caregivers blinded	Outcome assessors blinded	Outcomes reported as planned (link to protocol)	No other bias detected	Loss to follow-up (%) for outcome radiographic healing unless specified otherwise
Busse 2014(24)	Yes	Yes	Yes	Yes	Yes	Yes[a]	Yes	2%
Busse 2016(16)	Yes	Yes	Yes	Yes	Yes	Yes[a]	Yes	19% for radiographic healing, 11% for return to work, 9% for weight bearing
Dudda 2011(25)	Yes	No	No	No	No	Unclear[b]	Yes	Unclear, assumed to be 0
El-Mowafi 2005(26)	Yes	No	No	No	No	Unclear[b]	Yes	5%
Emami 1999(27, 28)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	3%
Gan 2014(29)	Yes	No	Yes	Yes	Yes	Unclear[b]	Yes	23% (pain)
Handolin 2005a(30, 31)	Yes	No	Yes	Yes	Yes	Unclear[b]	Yes	5%
Handolin 2005b (32)	Yes	No	Yes	Yes	Yes	Unclear[b]	Yes	No eligible outcome reported
Heckman 1994(33)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	31%
Kamath 2015(45)	Yes	No	No	No	Yes	Unclear[b]	Yes	No eligible outcome reported
Kristiansen 1997(34)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	28%
Leung 2004(35)	No[c]	No[c,d]	No[d]	No[d]	No[d]	Unclear[b]	No[e]	Unclear, assumed to be 0
Liu 2014(36)	Yes	No	No	No	yes	Unclear[b]	No[f]	Unclear, assumed to be 0
Lubbert 2008(37)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	16%
Mayr 2000(38)	Yes	No	No	No	Yes	Unclear[b]	Yes	0
Patel 2014(39)	Yes	No	No	No	No	Unclear[b]	Yes	Unclear, assumed to be 0
Ricardo 2006(40)	Yes	No	Yes	Yes	Yes	Unclear[b]	Yes	Unclear, assumed to be 0

Rue 2004(42)	Yes	No	Yes	Yes	Yes	Unclear[b]	Yes	Unclear, probably 35%
Rutten 2012(41)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	45%
Salem 2014(43)	Yes	No	No	No	No	Unclear[b]	Yes	Unclear, assumed to be 0
Schofer 2010(44)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	No[g]	Unclear, assumed to be 0
Schortinghuis 2005(46)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	0 for subsequent operation
Schortinghuis 2008(47)	Yes	Yes	Yes	Yes	Yes	Unclear[b]	Yes	0 for subsequent operation
Tsumaki 2004(48)	Yes	Yes	No	No	No	Unclear[b]	No[h]	Unclear, assumed to be 0
Urita 2013(49)	No[i]	No	No	No	Yes	Unclear[b]	Yes	Unclear, assumed to be 0
Zacherl 2009(50)	Yes	No	yes	yes	yes	Unclear[b]	No[k]	Not included in meta-analysis due to insufficient reporting[k]

a Protocol: NCT00667849
b No protocol published and trial not registered
c Quasi-randomised based on sequence of admission
d Inactive device was distinguishable from active device
e Unadjusted clustering, 30 fractures of 28 patients were randomized
f Implausibly narrow confidence intervals
g Prognostic imbalance: non-union fractures in LIPUS arm were considerably older
h Bilateral surgery – one tibia was randomised to LIPUS and one to no treatment. We assumed a correlation of 0.5 in our analysis of days to radiographic healing
i Used an odd-even system for treatment allocation
k Randomised 44 patients but analysed 52 toes, clustering unclear, standard deviations not reported

Table 3, GRADE Summary of Findings table

Outcome	Study results and measurements	Absolute effect estimates		Quality of evidence	Narrative Summary
		No ultrasound	LIPUS		
Days to return to work	% Difference: 2.7% (95% CI, -7.7% to 14.3%) in days, lower better Based on data from 392 patients in 3 studies	**200 days** (Mean)	**205 days** (Mean)	**Moderate** Due to serious imprecision	LIPUS probably has little or no impact on time to return to work
		Difference: 5 days later (95% CI, 15 earlier to 20 later)			
Days to full weight bearing	% Difference: 4.8% (95% CI, -4.0% to 14.4%) in days, lower better Based on data from 483 patients in 2 trials at low risk of bias	**70 days** (Mean)	**73 days** (Mean)	**High**	LIPUS has no impact on time to full weight bearing
		Difference: 3 days earlier (95% CI, 3 earlier to 10 later)			
Pain reduction Follow up 4 to 6 weeks	Mean difference: -0.93 (95% CI -2.51 to 0.64) 0 to 100 visual analogue scale, lower better, minimal important difference: 10-15 Based on data from 626 patients in 3 trials at low risk of bias	**40** (Mean)	**39** (Mean)	**High**	LIPUS has no impact on pain reduction
		Difference: 1 lower (95% CI 3, lower to 1 higher)			
Subsequent operations Follow up 8 weeks to 44 months	Risk ratio: 0.80 (95% CI 0.55 to 1.16) Based on data from 740 patients in 7 studies	**160 per 1000**	**128 per 1000**	**Moderate** Due to serious imprecision	LIPUS probably has little or no impact on subsequent operation
		Difference: 32 fewer (95% CI, 72 fewer to 26 more)			
Days to radiographic healing	% Difference: -1.7% (95% CI, -11.2% to 8.8%) in days, lower better Based on data from 483 patients in 3 trials at low risk of bias	**150 days (Mean)**	**147 days (Mean)**	**Moderate** Due to serious imprecision	LIPUS probably has little or no impact on time to radiographic healing
		Difference: 3 days earlier (95% CI, 17 earlier to 13			

		later)			
Device-related adverse effects Follow up 5 to 52 weeks	Risk difference: 0% (CI 95% -1% to 1%) Based on data from 839 patients in 9 studies	**0 per 1000**	**0 per 1000**	**High**	LIPUS has no impact on device-related adverse effects
		Difference: 0 fewer (95% CI, 10 fewer to 10 more)			

Table 4, Credibility of subgroup effects for risk of bias for the outcome days to radiographic healing

Criteria(20)	Rating (yes means higher credibility)
Is the subgroup variable a characteristic measured at baseline or after randomization?	Not applicable for risk of bias
Is the effect suggested by comparisons within rather than between studies?	No, between studies
Was the subgroup effect specified a priori?	Yes, specified in our protocol
Was the direction of the subgroup effect specified a priori?	Yes, we expected a larger effect for studies at high risk of bias
Is there indirect evidence that supports the hypothesized interaction (biological rationale)?	Not applicable for risk of bias
Was the subgroup effect one of a small number of hypothesized effects tested?	Yes, one of three
Does the interaction test suggest a low likelihood that chance explains the apparent subgroup effect?	Yes, significant in univariable subgroup analysis ($p<0.001$)
Is the significant subgroup effect independent?	Yes, significant in multivariable meta-regression ($p<0.01$)
Is the size of the subgroup effect large?	Yes, 31.8% acceleration in high risk of bias trials versus 1.7% acceleration in low risk of bias trials
Is the interaction consistent across closely related outcomes within the study?	Yes, risk of bias explained heterogeneity in outcomes weight bearing and pain
Is the interaction consistent across studies?	Yes, high risk of bias studies consistently showed large effects, low risk of bias studies small effects

FIGURES

Figure 1, Flow diagram of studies included in review of low intensity pulsed ultrasound compared with control (sham device or no device) for patients with fracture or osteotomy

Figure 2, Forest plot for percent difference of days to return to work for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device)

Figure 3, Forest plot for percent difference of days to full weight bearing for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device), by risk of bias. Interaction p<0.001

Figure 4, Forest plot for mean difference of pain reduction, all instruments transformed to 0-100 visual analogue scale, by risk of bias. Interaction p<0.001

Figure 5, Forest plot for risk ratio for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of number of subsequent fracture-related operations

Figure 6, Forest plot for percent difference for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of days to radiographic healing, by risk of bias. Interaction p<0.001

Figure 7, Forest plot for risk difference for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of ultrasound device related adverse effects

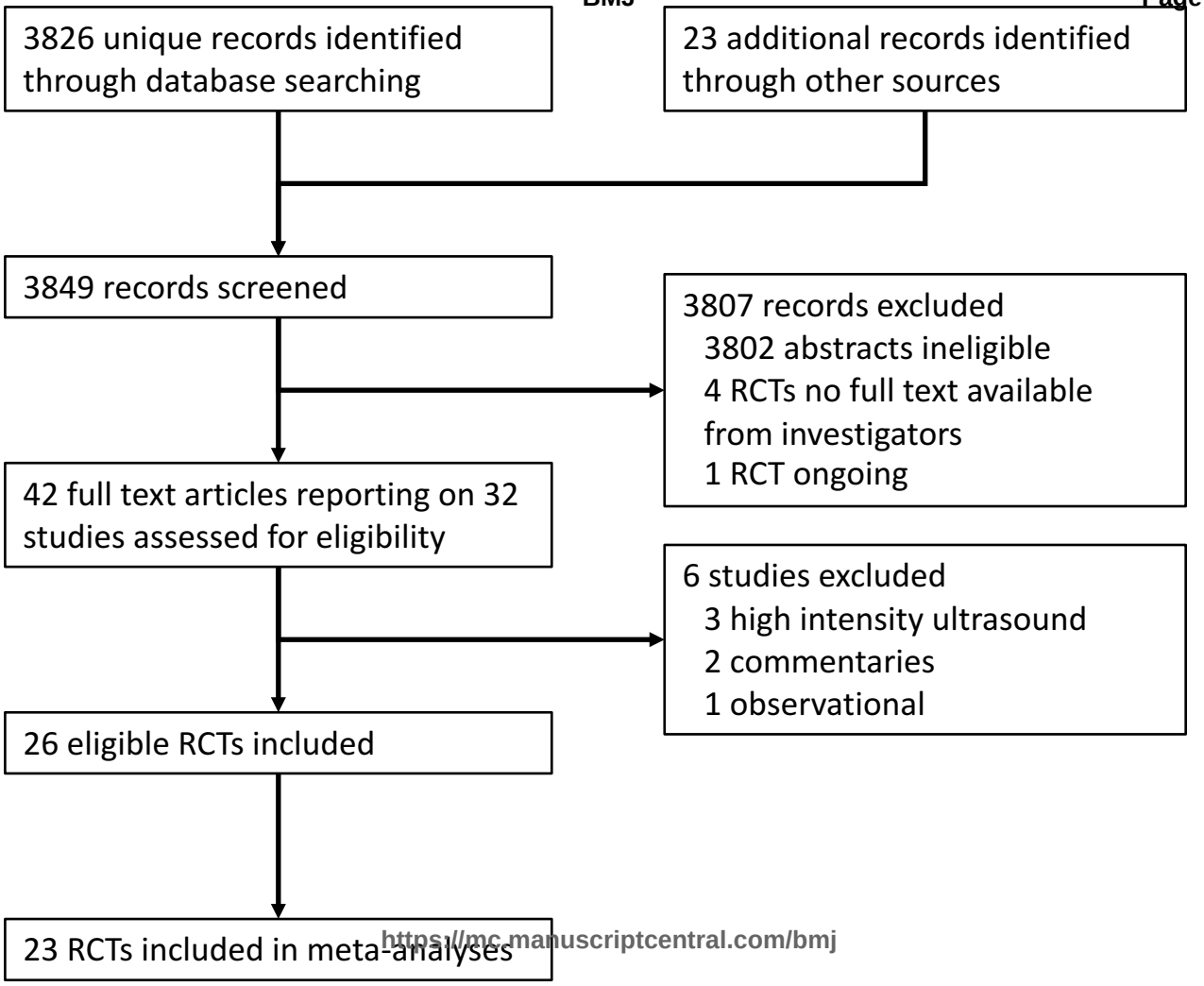

Figure 2, forest plot for percent difference of days to return to work for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device)

	LIPUS			Control					
Study	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**% Difference (95% CI)**	**Weight**	**% Difference (95% CI)**
Rue 2004	56.2	19.6	14	55.8	15.5	12		19.6%	0.7 (-20.9 to 28.2)
Lubbert 2008	17.0	11	50	15.05	11	47		15.0%	13.0 (-14.2 to 48.8)
Busse 2016	202.9	108.6	139	200.7	113.5	130		65.4%	1.1 (-11.4 to 14.3)
Total (heterogeneity: P=0.76, I^2 = 0%)			203			189		100%	2.7 (-7.7 to 14.3)

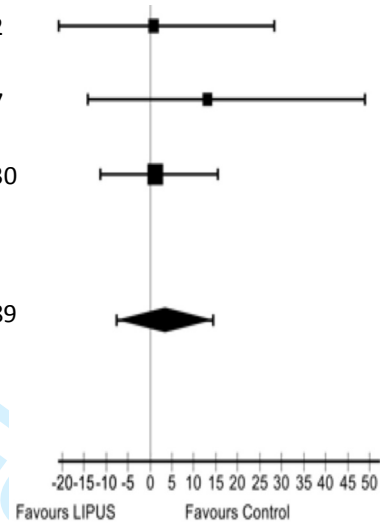

Figure 3, Forest plot for percent difference of days to full weight bearing for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device), by risk of bias. Interaction p<0.001

	LIPUS			Control					
Study and subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**% Difference (95% CI)**	**Weight**	**% Difference (95% CI)**
High risk of bias									
Leung 2004	65.1	14.7	16	108.5	21	14		34.4%	-40.0 (-48.4 to -30.3)
Low risk of bias									
Emami 1999	45.5	18.9	15	49.7	23.1	17		30.1%	-8.5 (-32.6 to 24.3)
Busse 2016	76.9	38.5	228	72.5	35.5	223		35.4%	6.1 (-3.2 to 16.2)
Subtotal (Heterogeneity: P=0.37, I^2 = 0%)			243			240		35.5%	4.8 (-4.0 to 14.4)
Total (Heterogeneity: P<0.001, I^2 =95.0%)			259			254		100%	-16.6 (-44.9 to 26.1)
Test for subgroup differences: P<0.001									

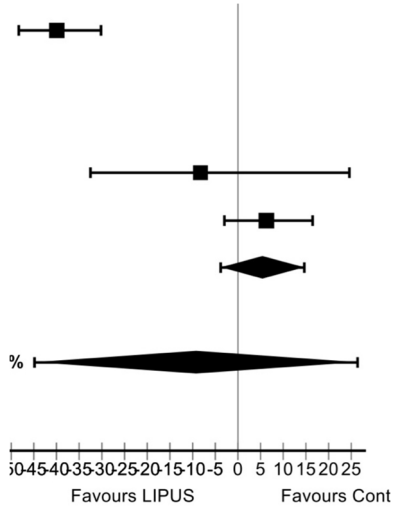

Figure 4, Forest plot for mean difference of pain reduction, all instruments transformed to 0-100 visual analogue scale, by risk of bias. Interaction p<0.001

Study and subgroup	LIPUS Mean	SD	Total	Control Mean	SD	Total	Mean difference (95% CI)	Weight	Mean difference (95% CI)
High risk of bias									
Patel 2014	-72.41	8.93	14	-44.29	14.53	14		21.6%	-28.12 (-37.05 to -19.19)
Low risk of bias									
Lubbert 2008	27.895	15.6	52	28.33	13.7	49		25.2%	-0.43 (-6.15 to 5.28)
Busse 2014	-41.8	10.3	23	-39.9	11.2	27		25.0%	-1.90 (-7.86 to 4.06)
Busse 2016	-39	9.8	237	-38.1	9.2	238		28.2%	-0.90 (-2.61 to 0.81)
Subtotal (heterogeneity: P=0.94, I^2 = 0%)			312			314		78.4%	-0.93 (-2.51 to 0.64)
Total (heterogeneity: P < 0.001, I^2 = 91.0%)			326			328		100%	-6.92 (-15.39 to 1.55)

Test for subgroup differences: P < 0.001

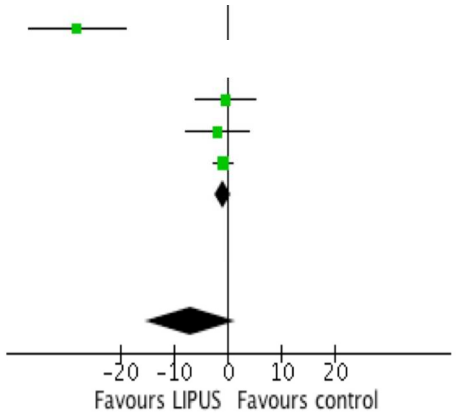

Figure 5, Forest plot for risk ratio for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of number of subsequent fracture-related operations

Study	LIPUS No of events	LIPUS Total	Control No of events	Control Total	Risk ratio (95% CI)	Weight	Risk ratio (95% CI)
Emami 1999	1	15	5	15		3.4%	0.20 (0.03 to 1.51)
Leung 2004	0	15	2	13		1.6%	0.17 (0.01 to 3.34)
Handolin 2005a	0	15	0	15			Not estimable
Schortinghuis 2005	0	4	0	4			Not estimable
Lubbert 2008	5	52	6	49		11.0%	0.79 (0.26 to 2.41)
Schortinghuis 2008	0	5	0	4			Not estimable
Dudda 2011	2	16	4	20		5.6%	0.63 (0.13 to 2.99)
Busse 2014	6	23	6	25		14.4%	1.09 (0.41 to 2.90)
Patel 2014	1	14	0	14		1.4%	3.00 (0.13 to 67.91)
Busse 2016	28	217	32	205		62.6%	0.83 (0.52 to 1.32)
Total (Heterogeneity: P=0.67, I^2 = 0%)	43	376	55	364		100.0%	0.80 (0.55 to 1.16)

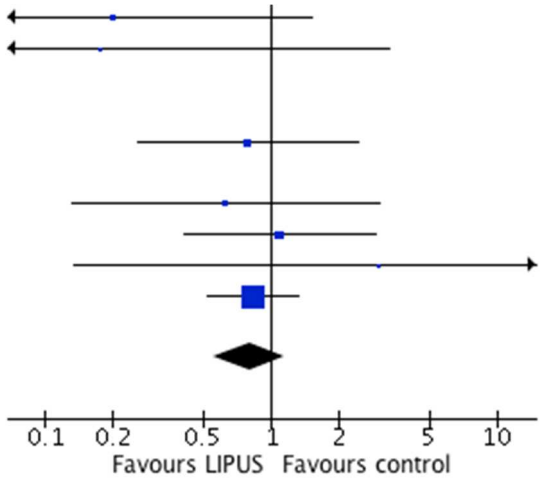

Figure 6, Forest plot for percent difference for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of days to radiographic healing, by risk of bias. Interaction p<0.001

	LIPUS			Control					
Study and subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**% Difference (95% CI)**	**Weight**	**% Difference (95% CI)**
High risk of bias									
Heckman 1994	102	27.6	33	190	106.7	34		7.0%	-46.3 (-56.5 to -33.8)
Kristiansen 1997	51	21.9	30	77	27.8	31		7.2%	-33.8 (-45.7 to -19.1)
Mayr 2000	43.2	10.9	15	62	19.2	15		7.1%	-30.3 (-43.1 to -14.7)
Leung 2004	80.5	21	16	140	30.8	14		7.7%	-42.5 (-51.6 to -31.7)
Tsumaki 2004	49.7	18.2	21	55.3	16.8	21		8.1%	-10.1 (-22.3 to 3.9)
El-Mowafi 2005	30	3	10	48	9.8	9		8.1%	-37.5 (-46.0 to -27.7)
Ricardo 2006	56	10.1	10	94	15.9	11		8.1%	-40.4 (-48.7 to -30.8)
Dudda 2011	32.8	13.1	16	44.6	26.8	20		5.1%	-26.5 (-47.0 to 2.1)
Rutten 2012	80	28	4	187	100.5	7		2.9%	-57.2 (-74.7 to -27.6)
Urita 2013	57	10	14	77	26	13		7.1%	-26.0 (-39.7 to -9.1)
Liu 2014	32	2.6	41	40.8	5.1	40		9.4%	-21.4 (-24.9 to -17.7)
Salem 2014	33	16	12	45	34	9		2.6%	-26.7 (-58.3 to 29.0)
Subtotal (Heterogeneity: P<0.001, I^2 = 77.8)			222			224		80.4%	-32.8 (-39.5 to -25.3)
Low risk of bias									
Emami 1999	155	85.2	15	125	45.4	17		5.1%	24.0 (-10.6 to 72.0)
Busse 2014	151.7	59.9	23	161.6	101.2	24		5.5%	-6.1 (-30.3 to 26.5)
Busse 2016	143.2	62.6	209	148.7	63.9	195		0.9%	-3.7 (-11.5 to 4.8)
Subtotal (Heterogeneity: P= 0.33, I^2 = 9.8%)			247			236		19.6%	-1.7 (-11.2 to 8.8)
Total (Heterogeneity: P<0.001, I^2 =84.7%)			469			460		100%	-27.3 (-34.7 to -19.0)
Test for subgroup differences: P<0.001									

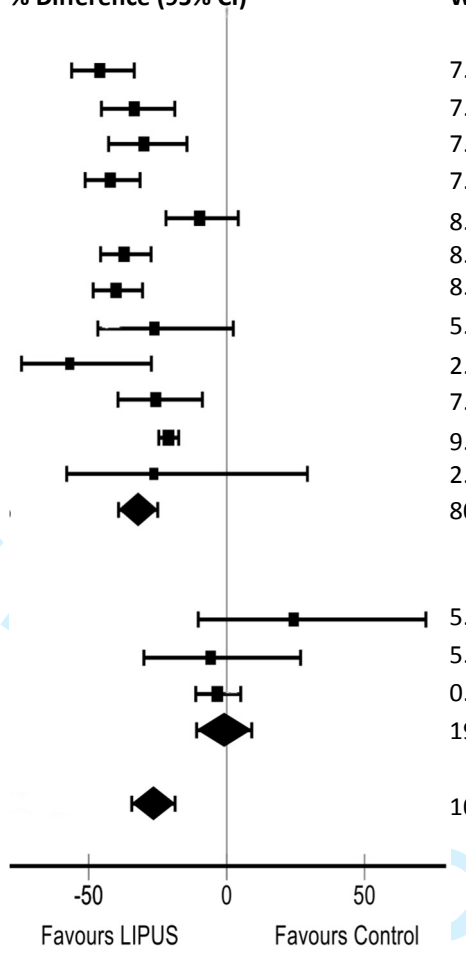

Figure 7, Forest plot for risk difference for low intensity pulsed ultrasound (LIPUS) compared with control (sham device or no device) of ultrasound device related adverse effects

Study	LIPUS No of events	Total	Control No of events	Total	Risk difference (95% CI)	Weight	Risk difference (95% CI)
Kristiansen 1997	0	30	0	31		7.3%	0.00 (-0.06 to 0.06)
Leung 2004	4	15	0	13		3.3%	0.27 (0.03 to 0.51)
Lubbert 2008	1	52	1	49		1.2%	-0.00 (-0.06 to 0.05)
Schofer 2010	0	51	0	50		12.1%	0.00 (-0.04 to 0.04)
Urita 2013	0	14	0	13		3.2%	0.00 (-0.13 to 0.13)
Gan 2014	0	10	0	13		2.7%	0.00 (-0.16 to 0.16)
Patel 2014	0	14	0	14		3.3%	0.00 (-0.13 to 0.13)
Busse 2014	0	23	0	25		5.7%	0.00 (-0.08 to 0.08)
Busse 2016	0	217	0	205		50.3%	0.00 (-0.01 to 0.01)
Total (Heterogeneity: P = 0.40, I^2 = 4.0%)	5	426	1	413		100.0%	0.01 (-0.01 to 0.03)