A Review of Therapeutic Ultrasound: Effectiveness Studies

Background and Purpose. Therapeutic ultrasound is one of the most widely and frequently used electrophysical agents. Despite over 60 years of clinical use, the effectiveness of ultrasound for treating people with pain, musculoskeletal injuries, and soft tissue lesions remains questionable. This article presents a systematic review of randomized controlled trials (RCTs) in which ultrasound was used to treat people with those conditions. Each trial was designed to investigate the contributions of active and placebo ultrasound to the patient outcomes measured. Depending on the condition, ultrasound (active and placebo) was used alone or in conjunction with other interventions in a manner designed to identify its contribution and distinguish it from those of other interventions. Methods. Thirty-five English-language RCTs were published between 1975 and 1999. Each RCT identified was scrutinized for patient outcomes and methodological adequacy. Results. Ten of the 35 RCTs were judged to have acceptable methods using criteria based on those developed by Sackett et al. Of these RCTs, the results of 2 trials suggest that therapeutic ultrasound is more effective in treating some clinical problems (carpal tunnel syndrome and calcific tendinitis of the shoulder) than placebo ultrasound, and the results of 8 trials suggest that it is not. Discussion and Conclusion. There was little evidence that active therapeutic ultrasound is more effective than placebo ultrasound for treating people with pain or a range of musculoskeletal injuries or for promoting soft tissue healing. The few studies deemed to have adequate methods examined a wide range of patient problems. The dosages used in these studies varied considerably, often for no discernable reason. [Robertson VJ, Baker KG. A review of therapeutic ultrasound: effectiveness studies. Phys Ther. 2001;81:1339–1350.]

Key Words: Clinical trials, Systematic review, Therapeutic ultrasound.

Valma J Robertson
Kerry G Baker
In this article, we present a systematic review of studies of the therapeutic effectiveness of ultrasound. All studies included both active and placebo ultrasound treatment groups. Depending on the problem, ultrasound (active and placebo) was used in conjunction with other interventions but in such a manner that its contribution could be distinguished from any other components of treatment. We focused on studies using ultrasound to treat people with pain and a range of musculoskeletal injuries and to promote soft tissue healing. Nussbaum1 and Robertson and Spurritt2 reported that ultrasound is one of the most frequently used electrophysical agents (EPAs) in physical therapy practice. Ultrasound is widely used in many countries, including Canada,1 Australia,2–4 Denmark,5 Finland,6 New Zealand,7 Switzerland,8 the United Kingdom,9,10 and the United States.11 Physical therapists have given many reasons for using ultrasound such as for the "physiological effects" or because of beliefs in "clinical results"4 or "expected effects."2

We used a systematic review to examine whether there is sufficient evidence to accept the premise that therapeutic ultrasound is effective. That is, does active ultrasound used alone, or with other interventions, produce a different outcome than placebo ultrasound applied under the same conditions? Based on the available evidence, we also examined issues related to dosage and usage, specifically, the total energy and the energy density applied.

Reviews of Ultrasound Research

Nussbaum1 reported that early clinical trials attempting to examine the effectiveness of therapeutic ultrasound were typically flawed. Holmes and Rudland12 reported that, of the 18 trials they evaluated, most had methodological flaws, including a lack of control groups, of standardized treatment and assessment criteria, and of statistical analyses of the results. Gam and Johannsen5 reviewed articles published between 1950 and 1992 on ultrasound used to treat subjects with musculoskeletal problems. They concluded that only 22 of the 293 articles they reviewed were methodologically adequate and that any contribution of ultrasound to the treatment outcomes was not evident on the basis of the findings of controlled studies. Gam and Johannsen also reported that they were unable to investigate any possible dosage-response relationship because of the inadequacy of the treatment details provided.

Based on their meta-analysis of trials of physical therapy treatments for soft tissue lesions of the shoulder, van der Heijden et al13 concluded that ultrasound is not effective and its use should be discouraged. This conclusion was later criticized as being based on methodologically inadequate research studies.14 The criticisms emphasized the need for studies using double blinding, an internally valid method of placebo treatment, and adequate group sizes and including details of the dosage of ultrasound used. That is, the type of study required for credible conclusions in this context are randomized controlled trials (RCTs), which also provide adequate dosage details.

In a recent systematic review, van der Windt et al15 analyzed 38 RCTs and controlled clinical trials (CCTs) (trials in which subjects were not randomly assigned to groups) of the effectiveness of ultrasound used to treat people with musculoskeletal disorders. They concluded that there is little evidence to support the use of active ultrasound therapy for treating people with those disorders. Most methodologically adequate studies (n=13) they reviewed lacked evidence of either meaningful outcomes or statistically significant differences from using ultrasound.

We have since identified additional RCTs of ultrasound published in the English language and not reviewed by van der Windt et al.15 Some of these RCTs, which were beyond the scope of van der Windt and colleagues’ study, focused on tissue healing16–22 or pain.23 Other trials focused on using ultrasound for either a combination of pain and soft tissue healing or to change a consequential functional loss.24–33 In all trials, at least active and placebo ultrasound were applied to treat one of these conditions. In some trials, additional, identical interventions were given to both groups but in such a manner that any contribution of active ultrasound to patient outcomes was evident with analysis.

VJ Robertson, PT, PhD, is Associate Professor, School of Physiotherapy, La Trobe University, Bundoora, Victoria 3086, Australia (V.Robertson@latrobe.edu.au). She was Visiting Professor, Division of Physical Therapy, University of Miami, Miami, Fla, when this article was written. Address all correspondence to Dr Robertson.

KG Baker, PT, PhD, is Senior Lecturer, Department of Health Science, Faculty of Health, Science, and Technology, UNITEC, Auckland, New Zealand. He was Lecturer, School of Physiotherapy, Faculty of Health Sciences, University of Sydney, when this article was written.

Both authors provided concept/project design, writing, and data collection and analysis. Dr Robertson provided project management and consultation (including review of manuscript before submission).

This article was submitted February 18, 2000, and was accepted March 5, 2001.
Given the importance of ultrasound in physical therapy, we believe that a new systematic review of ultrasound for pain and soft tissue is necessary. Besides additional studies being available, we will use criteria that are slightly different from those used in the reviews discussed and will examine only RCTs. In addition, we will try to elucidate dosage variables that might affect outcome. These variables include the intensity of ultrasound applied, the size of the area treated, and the duration of treatment for specific problems.

Randomized controlled trials are widely recognized as the best way of comparing the effectiveness of different treatments. We do not believe that RCTs are the only method of obtaining information about an intervention. Many laboratory trials of ultrasound, for example, have demonstrated that it can effect changes consistent with healing. In addition, other research methods might demonstrate differences and suggest important research hypotheses. Without applying the rigorous criteria implicit in an RCT, however, our degree of certainty about the outcome of such research is limited. Also, these approaches, although providing information about ultrasound, cannot produce clinically applicable evidence of effectiveness.

Randomized controlled trials, like other research methods, can be biased if done badly, and they often have deficiencies in what they report. The discrepancy between what aspects of a study should be reported and what aspects of a study are reported has long been known. Readers, however, have to judge a study on the basis of what is reported and available in the public domain.

For this reason, the present review was done in 2 stages. In the first stage, we identified a set of relevant RCTs and investigated their methodological adequacy. The criteria each RCT needed to meet for inclusion in the present study were based on those developed by Sackett et al and are shown in Table 1. In the second stage, we analyzed relevant aspects of their content in order for us to judge the clinical effectiveness of therapeutic ultrasound for treating people with pain and musculoskeletal disorders and for promoting soft tissue healing.

### Method

Our first step was to identify all relevant research articles for this study. The following methods were used: perusal of physical therapy journals from 1975 to 1999, searches of relevant medical and allied health care databases (MEDLINE and CINAHL), reading of recent review articles and reference lists, and consultations with colleagues.

A total of 35 RCTs of therapeutic ultrasound published in the English language were identified and are listed in Table 2. Both authors of this article independently read the articles describing these trials. The initial reading was done in an effort to ensure that all articles described RCTs that investigated the clinical use of ultrasound for treating people with pain or musculoskeletal injuries or for promoting soft tissue healing. Some studies were excluded at this stage. Two studies using subjects without impairments in what we would consider laboratory conditions were not analyzed further. The first study involved treatment of a local skin inflammation caused by applying ultraviolet therapy, and the second study investigated the effects of ultrasound on the mechanical pain threshold in subjects with no known pathology. Neither represents one of the conditions affecting soft tissue or causing pain or functional restrictions for which ultrasound is used clinically. Next, 4 studies using multiple interventions were rejected. The method of using ultrasound in these studies made it difficult for us to distinguish any effects of it from that of other interventions given concurrently. Therefore, these 4 studies were not included in this systematic review. Finally, 2 studies that duplicated other published results were excluded. Both studies presented subjects and results published in other articles that are included in our review. We then applied the methodological filters shown in Table 1 to the remaining 27 RCTs.

### Filter 1—Controls

Appropriate controls (methodological filter 1) were believed to be present if subjects were randomly allocated to groups and if there was an active ultrasound treatment group and a placebo ultrasound treatment group, with both groups receiving otherwise apparently identical treatments. According to Hashish et al, this methodological filter is crucial when investigating an intervention that is alleged to have a high placebo effect. Some studies included in this review also had a true control group, which received no ultrasound treatment. Those studies remaining after screening are discussed in the second stage of this review.

### Table 1. Methodological Filters Applied to Randomized Controlled Trials

<table>
<thead>
<tr>
<th>Filter</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Controls</td>
<td>Adequate controls, including placebo treatment and randomized group allocation.</td>
</tr>
<tr>
<td>2. Blinding</td>
<td>Adequate blinding of observers, subjects, and therapists to group allocation.</td>
</tr>
<tr>
<td>3. Description</td>
<td>Adequate description of treatment variables (including checking of machine output).</td>
</tr>
<tr>
<td>4. Outcome</td>
<td>Meaningful outcome measures (ie, valid for patient problem being treated).</td>
</tr>
<tr>
<td>5. Adequacy</td>
<td>Adequacy of sample size for trials showing no treatment effect.</td>
</tr>
</tbody>
</table>

We deemed 26 of the 27 remaining studies to be adequate for further analysis. The study that we excluded as not adequate for further analysis had 2 groups, a group that received ultrasound and the standard treatment and a group that received the standard treatment only. Without including a group that receives both placebo ultrasound and the standard treatment, we believe that it is impossible to distinguish the placebo component in the outcome possibly contributed by using therapeutic ultrasound equipment.

Filter 2—Blinding of Assessors, Subjects, and Users

Three aspects of experimental blinding (methodological filter 2) were considered: (1) blinding of the assessor, (2) blinding of the subjects, and (3) blinding of the users of the ultrasound equipment (ie, the therapists). There was insufficient evidence of adequate blinding on at least one of these grounds for 16 of the remaining 26 studies.16,17,20,21,23,24,26,29,30,32,33,51–55

Blinding of subjects and therapists is a complex and important issue when using ultrasound. Intensity is only one factor affecting the ability of subjects and therapists to identify whether there is or is not an output. Other factors that we believe can contribute to an accurate identification of an output are the speed of applicator movement, the anatomical location insonated, and the temperature of the contact medium used.

The articles by Dyson et al and Downing and Weinstein indicate the extent to which subjects and therapists can correctly identify whether the subjects are receiving ultrasound or a placebo. Dyson et al reported that 6 (66.6%) of the 9 subjects in the experimental treatment group correctly identified the treatment they were receiving.
group detected the skin heating produced by pulsed ultrasound with an estimated space-averaged time-averaged (SATA) intensity of 0.2 W/cm² using a 1-MHz applicator and a dosage of 1 minute per 0.5 cm² surface area. Downing and Weinstein⁵² reported that 50% of the volunteers who were each treated 6 times could correctly identify whether their treatment consisted of active ultrasound or placebo ultrasound. Unfortunately, no information was provided on the intensity used, which we believe is a major factor. Downing and Weinstein⁵² then used warmed gel to increase the likelihood of subjects not knowing whether the ultrasound machine was producing sound waves. The mean level of ultrasound dosage was 1.2 to 1.3 W/cm² for an area of about 150 cm² for 6 minutes, a higher dosage than is used in many studies but also a larger area insonated. Six (40%) of 20 subjects guessed correctly whether they were receiving active ultrasound or placebo ultrasound, and the therapist correctly identified 8 (72.7%) of the 11 subjects who received ultrasound. These findings indicate the difficulties in ensuring that both subjects and therapists are unaware of the output status of an ultrasound machine (double blinding).

McLachlan²⁸ decreased the likelihood of both subject and operator identification of active or placebo ultrasound through local heating by using a resistor to produce some heating in a modified, but otherwise identical, ultrasound applicator. In view of the problems in implementing effective blinding strategies using ultrasound, no studies could be excluded with certainty because of this methodological problem. As a result, 26 studies were left.

There was, however, one exception to our ignoring the blinding method.²⁹ In a study in which the investigators used an applicator with a discernibly different output (45 kHz), the authors recognized this as a problem due to the frequency of ultrasound used.²⁹ They commented that their method was inadequate for double blinding. All therapists and many patients would have been aware that there was no output with the placebo 45-kHz ultrasound. At that frequency, ultrasound has a barely audible output and produces a detectable cutaneous sensation. Excluding that study²⁹ left us with 25 studies as methodologically acceptable.

**Filter 3—Treatment Variables**

In 9 of the remaining 25 trials, the researchers did not provide sufficient details about treatment to allow replication of the study. In 8 trials, there was no indication of whether the output of the ultrasound machines used was checked.²⁰,²¹,²³,²⁵,³⁰,³³,⁵⁵–⁵⁷ Several researchers⁵⁸–⁶¹ have measured ultrasound machine output and reported that it frequently varied more than 30% of that indicated. We believe that such results justify concerns about possible discrepancies between the dosage displayed on the machine and that given to the patient. Such discrepancies could affect outcome and have been suggested as a possible reason why ultrasound treatments are believed to be ineffective.⁶² With the exception of those listed, the remaining RCTs documented that the machine’s output was checked, in some cases prior to each session.

We excluded one study²⁴ because we believed there were inadequate details of the intensity of ultrasound applied. Of the remaining 16 studies, a frequency of 3 MHz was used in 7 studies (3.28 MHz in the study by ter Riet et al²²) and a frequency of 1 MHz was used in 9 studies (0.89 MHz in the study by Ebenbichler et al³¹ and 1.1 MHz in the study by McLachlan²⁸). McLachlan²⁸ provided the name of the machine used but not its frequency; the Medtron P300⁴–⁸ used in that study has a frequency of 1 MHz.²⁸ In the study by Hashish et al,²⁵ the SATA intensities for the 3-MHz machines ranged from 0.02 to 0.3 W/cm². The SATA intensities for the 1-MHz machines ranged from 0.2 W/cm² in 3 RCTs⁶,⁵⁴,⁶³ to 2.4 to 2.6 W/cm² in 2 RCTs.²⁸,⁵₃ Researchers who commented on their selection of treatment variables said that their choice reflected practice.²²,²⁸

Details of either the effective radiating area (ERA) or the geometric area of the applicator were accepted for the purpose of our study. Because the piezoelectric element that generates the ultrasound does not vibrate uniformly, the ERA of the applicator is smaller than its geometric area.⁶⁴ As a consequence, calculations of dosage based on geometric area may, in some cases, slightly understate the actual wattage per unit area applied relative to those using ERA. As the applicator needs to be kept moving to avoid hot spots, the area affected by ultrasound energy cannot be precisely determined and the dosage at a particular depth of tissue cannot be known.⁶¹,⁶⁵ The beam nonuniformity ratio (BNR) expresses one variable contributing to this outcome; others include the wattage applied, the depth and types of tissue, and the frequency of ultrasound. The BNR is not relevant to our article, given the omissions of more basic aspects of dosage, but it will have to be considered if appropriate dosages are ever to be determined through research.

The area treated and the duration of application also affect dosage. All research reports that we reviewed contained information on the treatment duration but often not the size of the area treated. Although constant movement of the applicator means underlying tissues receive a variety of ultrasound energy, we believe that details of the size of area treated (ideally volume) should always be provided by researchers. The studies we exam-
ined provided descriptions of the areas insonated, enabling an estimation of the size of area treated.

Filter 4—Outcome Measures
We believed that the different measures used for patient outcomes (methodological filter 4) in the remaining 16 studies were acceptable. In each study we judged, at least one outcome measure with face validity was used. We also believed that these measures have a widely and generally acceptable level of reliability, as evidenced by their use in more than one study. For example, in the studies in which the effect of ultrasound on ulcers was examined, there was use of tracings or a similar method of measuring ulcer area. In some studies, the measures used included the scoring of pain using a visual analog scale. Where appropriate, additional measures such as grip strength were used.

Some authors used what we would consider unacceptable measures; however, because multiple outcome measures were used, these studies were not excluded from our study. For example, Ebenlicher et al used an additive scale for ordinal data, but without any evidence of the appropriate Rasch analysis that we believe is necessary to justify that approach. Although we believed that particular measure not valid, other measures they used, in our opinion, were acceptable.

Filter 5—Sample Size and Power
We believe that sample size is only an issue when there are negative findings (methodological filter 5). A power analysis indicated that 3 of the RCTs that remained following screening had results that indicated that ultrasound is no more effective than placebo treatment, but we believe that these RCTs had too few subjects to identify even a large treatment effect. Following dropouts, Gam et al had 18 subjects in the active ultrasound, massage, and exercise group; 22 subjects in the placebo ultrasound, massage, and exercise group; and 18 untreated control subjects. Downing and Weinstein treated all 20 subjects in their study with active, active assisted, and passive range of movement exercises, followed by active ultrasound given to 11 subjects and placebo ultrasound given to 9 subjects. McDiarmaid et al had 21 subjects in their active ultrasound treatment group and 19 subjects in the placebo ultrasound group. To attain an 80% probability of being able to detect a treatment effect (alpha = .05) requires a minimum of 26 subjects per group for a 2-group study if a large difference in outcome is expected and a parametric statistic is used. If a nonparametric statistic is used, as in these studies, up to 20% more subjects are needed if a large treatment effect is to be identified. This methodological filter of sample size left us with 8 studies in which there were no discernable treatment effects and with 5 studies with positive findings.

Filter 6—Data Analysis
We identified 3 of the 13 remaining studies as having problems with aspects of their data analysis. The analysis of results presented by Dyson et al is confusing because the number of subjects (ulcers or patients) in each group in the first part of their study is unclear. There is another major problem of ulcer variability in that study. The authors noted that the chronic varicose ulcers tested ranged in size from 1.5 to 12.75 cm² and had existed from 6 to 360 months in their subjects and that the response to intervention was very variable. Although robust, a Student t test, like other parametric tests, is dependent on a level of homogeneity of variance. We believe that this is especially important when small numbers of subjects are used and, in this context, suggests that a nonparametric analysis rather than a parametric analysis should have been used. Either problem, we believed, was sufficient to exclude this study from further consideration.

Roche and West and Binder et al examined the effects of ultrasound, but they did not demonstrate an equivalence of the experimental and control subject groups before the study began. In the study by Roche and West, the ulcer size was markedly less in the placebo ultrasound group (23.62 versus 32.51 cm²) and of longer duration (12.35 versus 5.37 years) than for the active ultrasound group. Binder et al provided evidence of equivalence of patient details such as age but not of the initial measures of grip strength or of pain used to evaluate the effect of using ultrasound. Consequently, any apparent differences between the groups after treatment may have been due to differences between subjects in each group. For example, more subjects in the ultrasound group may have had lesions that were likely to recover more quickly, perhaps because of their relative recency or lesser severity or because of other factors such as the age and activity level of a subject.

Initial equivalence of groups cannot be assumed; therefore, we believe the results are not compelling. Dyson identified another problem in the study by Binder et al. Subjects with poor results from treatment typically did not rest adequately from the precipitating cause. The etiology of overuse injury suggests that rest is likely to be effective in assisting the resolution of this condition. We found, however, that the study by Binder et al provided too few details to examine this claim and to differentiate the contributions of active ultrasound from those of resting in patients with lateral epicondylitis.

Summary
From the original 35 studies identified, Table 2 shows that 10 studies remained. In 2 of those
10 studies, there were differences in outcomes as those subjects treated with active ultrasound improved but those subjects treated with placebo ultrasound did not improve. \(^{31,54}\) In the other 8 studies, no differences between groups treated with active ultrasound or placebo ultrasound were found.

**Analysis of Methodologically Acceptable Studies**

In this section, we analyze the 10 studies that we believed to be methodologically acceptable. The focus, however, will be on identifying common factors that might distinguish the 2 studies in which ultrasound affected the outcomes from those studies in which ultrasound did not affect the outcomes.

**Dosage**

Table 3 shows the conditions that the subjects had and the dosages used in the 10 studies. Some details had to be estimated, specifically, the size of the applicator for articles published prior to 1992\(^ {25–28,51}\) and the size of area treated. \(^ {6,25–28,31,54}\) For the applicator, this is likely to be correct as machines used prior to 1992 typically had 5-cm\(^2\) applicators with a smaller ERA. In the absence of information on the geometric size of the applicator, we assumed it to equal the ERA. This produces, in some instances, a small overestimation of the wattage applied. With the size of area treated, this was estimated from descriptions of the area treated and is clearly a potential source of error when used in subsequent calculations of energy density. \(^ {6,25–28,31,54}\)

Total energy (in joules) was calculated as watts per square centimeter × applicator size (in square centimeters) × time (in seconds). This calculation was done to enable comparisons of dosage between trials. Table 3 shows considerable variation among studies, probably compounded by the necessity for us to use estimates of applicator size and size of the area treated with ultrasound in many calculations. Studies in which 3-MHz frequency ultrasound was used had outputs ranging from 30 J\(^ {25}\) to 180 J.\(^ {22}\) For ultrasound with a frequency of 0.89 to 1 MHz, the range was from 600 J\(^ {6}\) to 11,600 J.\(^ {28}\) As might be expected, subjects who used the 3-MHz frequency ultrasound used less total energy. If multiplied by 3 to give some measure of comparability with 1-MHz frequency energy levels that are available at a superficial depth, the dosage estimates fall within a range of 90 to 1,450 J in the lower range outputs used with a 1-MHz frequency. The ultrasound dosages used in the 2 studies in which differences were found between placebo and active ultrasound\(^ {31,54}\) (estimated total energy applied as 2,250 and 900 J, respectively) were within the range of those used in studies in which no differences were found. This finding suggests to us that there was no obvious source of differences between the 2 categories of studies in the dosages of ultrasound applied.

The energy density (total energy [in joules] per unit area [in square centimeters]) for the remaining studies was from a low of 2 J/cm\(^2\) with a frequency of 3 MHz in the study by Hashish et al\(^ {25}\) to a high of 150 J/cm\(^2\) with a frequency of 1 MHz in the study by Ebenbichler et al\(^ {31}\) \((X=55.79\ J/cm^2, \ 95\% \ confidence \ interval=19.8–84.3)\). This is a large range. For example, in the 2 studies in which shoulders were treated, \(^ {6,31}\) a 1-MHz frequency and energy densities of 40 and 150 J/cm\(^2\), respectively, were used. Carpal tunnel syndrome and epicondylalgia were treated using a frequency of 1 MHz and energy densities of 60 and 120 J/cm\(^2\) in the studies by Ebenbichler et al\(^ {34}\) and Lundberg et al\(^ {31}\), respectively. These differences in the energy density used suggest that comparable areas are not treated with similar dosages. There also was no apparent relationship between the year of the study and energy density applied. In the first study\(^ {51}\) and the most recent study\(^ {31}\) using a 1-MHz frequency, researchers applied the highest energy densities. Table 3 shows that a similar phenomenon occurred among studies using a 3-MHz frequency.

The limitation of this analysis was the uncertainty added by the use of some estimates in the calculations. We based estimates we made of the geometric size of the applicator\(^ {25–27,51}\) and the size of the area treated\(^ {6,25–28,31,54}\) on the details provided in the respective articles. We made these estimates to compare the effect of different levels of dosage (energy density) on patient outcomes. The data, however, suggest that there was considerable variation in energy density that is not accounted for by the type of patient problem, by the size or depth of the area treated, or by the year of the study.

**Problems Treated**

The diversity of problems treated with ultrasound limits comparisons between studies and possible conclusions on effective dosages. Each study reviewed compared the effects of active and placebo ultrasound. Depending on the condition treated, both groups also had either identical concurrent treatment or no additional concurrent treatment. This allowed the contribution of ultrasound to be distinguished from other components of multiple interventions, the aim of an RCT. Between-study comparisons were difficult even for studies that seemed similar because of differences in the inclusion criteria used. For example, Ebenbichler et al\(^ {31}\) used ultrasound for people with calcific tendinitis of the shoulder, whereas Nykanen\(^ {6}\) simply said that the subjects had shoulder pain. The possibility of a meaningful comparison is limited by one study\(^ {6}\) having included a wider range of shoulder problems than the other study\(^ {31}\).  

---

Table 3.
Details of Dosage in Studies Deemed to Be Methodologically Acceptable

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency (MHz)</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1.1</td>
<td>1</td>
<td>1</td>
<td>3.28</td>
<td>1</td>
</tr>
<tr>
<td>Output</td>
<td></td>
<td>Pulsed</td>
<td>Continuous</td>
<td>Pulsed 1:4</td>
<td>Continuous</td>
<td>Continuous</td>
<td>Continuous</td>
<td>Continuous</td>
<td>Pulsed 1:4</td>
<td>Pulsed 1:4</td>
</tr>
<tr>
<td>SATA (W/cm²)²</td>
<td>0.02, 0.1, 0.3</td>
<td>1</td>
<td>0.1</td>
<td>0.12</td>
<td>2.4–2.6</td>
<td>2.5</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Applicator size (cm²)b</td>
<td>5 estc</td>
<td>5</td>
<td>5 est</td>
<td>5 est</td>
<td>10 (8.5 ERA)</td>
<td>5</td>
<td>4 or 1</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Time [min]d</td>
<td></td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>8–15</td>
<td>3</td>
<td>10</td>
<td>7.5</td>
<td>15</td>
</tr>
<tr>
<td>Total energy [J]²</td>
<td>30–450</td>
<td>3,000</td>
<td>60</td>
<td>72</td>
<td>5,760–11,600</td>
<td>4,500</td>
<td>600</td>
<td>45 or 180</td>
<td>900</td>
<td>2,250</td>
</tr>
<tr>
<td>Area [cm²]</td>
<td>15 est</td>
<td>25</td>
<td>15 est</td>
<td>10 est</td>
<td>50–125 est</td>
<td>100</td>
<td>15 est</td>
<td>10</td>
<td>15 est</td>
<td>15 est</td>
</tr>
<tr>
<td>Energy density [J/cm²]f</td>
<td>2–30</td>
<td>120</td>
<td>4</td>
<td>7.2</td>
<td>93.6–115.2</td>
<td>45</td>
<td>40</td>
<td>4.5 or 18</td>
<td>60</td>
<td>150</td>
</tr>
</tbody>
</table>

¹ Space-averaged time-averaged intensity (SATA) calculated from peak output and pulse frequency.
² Effective radiating area (ERA) typically not provided.
³ Estimated value.
⁴ Time per typical initial session(s).
⁵ Total energy calculated as: SATA (W/cm²) × Applicator size (cm²) × Time (sec) = Joules.
⁶ Energy density calculated as total energy per unit area (J/cm²).
Outcome Measures Used
A variety of outcome measures also were used. At least one study had a problematic outcome measure, discussed earlier, in which ordinal data were added. In that study, however, the authors also examined the results of electroneurological testing, which was used as a diagnostic tool and as an outcome measure. Nerve conduction studies can provide an accurate assessment of conditions involving compromise of peripheral nerves, including carpal tunnel syndrome. Because of this, we judged the study as acceptable, as were the other remaining studies. We contend, therefore, that differences in findings were not attributable to problems with outcome measures.

Outcomes for True Control Groups
In all 10 of the studies, active ultrasound and placebo ultrasound were used. In 3 studies there was what we would consider a true control group, which received neither active ultrasound nor placebo ultrasound. In each of these studies, the group that received active ultrasound had better outcomes than the true control group (outcome measures: pain, facial swelling, trismus, serum C-reactive proteins, plasma cortisol, weight test, pain and power with wrist dorsiflexion, and grip strength test). In the 2 studies in which the group that received the placebo treatment and true control group were compared, the authors reported better outcomes for the group that received the placebo treatment than for the true control group. This finding provides some support for the use of ultrasound equipment for treating some conditions. Such a finding is also consistent with known responses to pain, an outcome measure in many of the studies. We contend, therefore, that differences in findings were not attributable to problems with outcome measures.

Differences Between 2 Categories of Studies
We identified 2 categories of methodologically acceptable studies: studies in which ultrasound was found to produce desirable outcomes and studies in which that was not found. The same first authors published the only 2 studies showing active ultrasound as more effective than placebo ultrasound that were, we believe, methodologically adequate. This raises obvious questions as to how those studies differ from the other 8 studies.

Not all methodological filters were strictly applied. In particular, whether experimental blinding (filter 2) was effectively implemented during ultrasound use was not certain in 16 of the 26 studies remaining at that stage in the screening process. This issue is relevant to both studies that showed active ultrasound had desirable outcomes. Blinding of subjects and therapists was not addressed in the earlier study by Ebenbichler and colleagues. In the later study by Ebenbichler and colleagues, subjects and therapists were said to be unaware of which treatment was the placebo ultrasound treatment and which treatment was the active ultrasound treatment, but no details were provided. Had filter 2 been strictly applied, both studies would have been excluded early in the screening process.

In both studies by Ebenbichler and colleagues, the investigators applied ultrasound when it was possibly not the ideal treatment. In response to a critical letter about one of the studies, Ebenbichler described the purpose in their study as “investigating the efficacy of a promising entity—ultrasound treatment.” He agreed with the letter writer that there possibly was a relatively high relapse rate in patients with carpal tunnel syndrome treated with ultrasound. In a letter commenting on a different study investigating the use of ultrasound in people with carpal tunnel syndrome, a physiatrist contended that there were risks of aggravating the condition by heating directly over the nerve rather than over associated structures. This, however, does not detract from the fact that Ebenbichler et al reported a difference in outcome for those treated with ultrasound (outcome measures provided: subjective symptom score for main complaint and sensory loss; electroneurographic measurements of median motor nerve distal latency, antidromic sensory nerve conduction rate, and peak-to-peak amplitude; and physical function levels, including strength of handgrip and of finger pinch).

In the later study by Ebenbichler et al, the researchers treated individuals with calcific tendinitis of the shoulder. A feature of this condition can be the spontaneous resorption of the calcium deposits over time. We excluded one RCT from our analysis because it compared no treatment with a treatment of ultrasound and acetic acid iontophoresis. The study by Perron and Malouin had 2 groups—a treatment group and a true control group. The authors found no differences between the 2 groups over the 3 weeks of the study. During this time, both groups had positive changes, with a decreased size and density of calcium deposits and an increased range of passive shoulder abduction with decreased pain during it. In no instance was the change in any outcome measure statistically significantly different between the groups. That is, the natural course of calcific tendinitis was unaltered by ultrasound.

Discussion
In this review, we found few RCTs that investigated the contribution of therapeutic ultrasound to patient outcomes that met the minimal standards of methodological adequacy. Of the 10 RCTs that did meet our standards, 2 studies demonstrated improvements in outcome measures in subjects treated with ultrasound. In Ebenbichler and colleagues’ study of subjects with calcific tendinitis of the shoulder, the outcome measures were
change in calcium deposits in the shoulder and subjective symptoms and pain. In Ebenbichler and colleagues’ study of subjects with carpal tunnel syndrome, the outcome measures were subjective symptoms, electroneurographic test results, and physical functioning. In the remaining 8 studies, no statistically significant differences in outcome between subjects treated with ultrasound and subjects treated with placebo ultrasound were found.

Ultrasound has been used therapeutically for over 6 decades in the ways reported in the trials examined in this study. Any clinically significant effects should, by now, have been identified in a number of rigorous studies that showed which patient outcomes are improved by using therapeutic ultrasound. In our review, we found that is not the case. Furthermore, in the few methodologically adequate studies that exist, treatment was provided for a wide range of problems; thus, few conclusions can be drawn. Similarly, no replications exist of studies with significant findings. Having different researchers in a different facility using the same procedure and obtaining a similar finding would considerably affect the strength of our certainty about ultrasound.

We found that the dosages of ultrasound used in the studies we reviewed varied considerably and for reasons that were not always clear. No underlying patterns were evident except possibly that the studies with significant outcomes were among those using a higher total energy output. Furthermore, without adequate data, there is little scientific basis for dosage selection in clinical practice. This leaves a question of the extent to which the diversity of dosages used helps explain the limited evidence of effectiveness of therapeutic ultrasound.

Limitations

One possible limitation of our review is its exclusive focus on RCTs. There are a number of other methods of obtaining relevant information about therapeutic ultrasound. Many laboratory studies indicate that ultrasound has an in vitro effect (see “A Review of Therapeutic Ultrasound: Biophysical Effects” by Baker et al in this issue). Unless these effects are not only consistent with healing but also sufficient to alter a relevant patient outcome positively, they do not justify the clinical use of ultrasound. Based on their clinical experiences, many therapists believe that ultrasound benefits healing. Until methodologically adequate studies can demonstrate that people treated with active ultrasound consistently have a better outcome than those treated with placebo ultrasound, we believe that doubts must remain.

Another possible limitation is the particular criteria used as filters on the set of studies identified. Those criteria, however, are entirely consistent with other sets used in previous systematic reviews and meta-analyses. In addition, they are consistent with requirements given for reporting RCTs described in major medical journals published in the United States and the United Kingdom.

One possibly valid criticism of our review is that we did not apply the different filters with sufficient rigor. For example, had studies that provided dubious or imprecise details of blinding of subjects, assessors, and therapists been excluded, few would have passed filter 2. Similarly, had the criterion regarding the establishment of controls (filter 1) been rigorously applied, few studies would have passed, as many authors did not provide adequate details of how they randomly assigned subjects to groups. Consistent with the possibility of a differential application of filters, there are some differences between this review and the review by van der Windt et al. They accepted 2 studies that we rejected as methodologically inadequate. Had we accepted those 2 studies, however, it would have made no difference to the outcome of this review, because both studies demonstrated no differences when using active ultrasound rather than placebo ultrasound.

Conclusions

When methodologically flawed trials were excluded, there were few RCTs that investigated ultrasound and those RCTs provided little clinical evidence for the efficacy of therapeutic ultrasound. The application of the exclusion criteria and methodological filters resulted in the elimination of all except 10 clinical ultrasound trials from the present review. Eight studies showed that active ultrasound is no more beneficial than placebo ultrasound for the treatment of people with pain or soft tissue injury. Few generalizations can be drawn from the 2 trials in which active ultrasound was found to be superior to placebo ultrasound, given their heterogeneity and omission of important details. Consequently, there is still little evidence of the clinical effectiveness of therapeutic ultrasound as currently used by physical therapists to treat people with pain and musculoskeletal injuries and to promote soft tissue healing. There are, however, apparently considerably different beliefs as to what is an acceptable dosage.

Future Directions

The findings of the present review indicate the importance of systematically investigating the clinical effectiveness of therapeutic ultrasound and establishing whether there is a dose-response relationship. The first stage is to identify clinical problems for which ultrasound is undoubtedly effective. The next stage should be to establish experimental and treatment protocols and standardized methods for ensuring the output of all ultrasound.
equipment used. With sufficient such studies, meta-reviews should be possible and able to indicate more convincingly than systematic reviews the extent to which ultrasound affects clinical outcomes and under which conditions.

References


