How Accurate are Therapeutic Ultrasound Machines?

Siobhan Schabrun,[1] B Physio (Hons); **Hayley Walker**,[2] B Physio (Hons); **Lucy Chipchase**,[3] PhD

Abstract: Ultrasound machines are frequently used by physiotherapists. Thus, it is imperative that machines are accurate for safe and effective treatments. This study aimed to examine the accuracy of therapeutic ultrasound machines in terms of power output and timing function, and to determine factors that relate to the degree of machine accuracy. An observational design was used. Sixty-four machines were sampled. Machines were tested at three power settings, two frequencies, and pulsed and continuous modes. Timers were assessed for 5- and 10-minute durations. Ultrasound machines were considered inaccurate if they deviated by more than ±20% and timers if they deviated by ±10%. Fifty-nine percent (291/492) of power tests were inaccurate, with 79% of these producing less power output than depicted on the dial. Thirty-seven percent of timers were inaccurate, with mechanical timers more inaccurate than digital timers ($p<0.05$). Older ultrasound machines were more likely to be inaccurate ($p<0.05$). Length of time since machine calibration was also associated with machine inaccuracy at 3 MHz ($p<0.05$). There is a high level of power and timer inaccuracy in machines that may be related to machine age and frequency of calibration. This study highlights the need for therapists to be aware of the potential for machine inaccuracy and for stricter guidelines on machine calibration to be introduced.

Key words: accuracy, physiotherapy, safety, ultrasound

Introduction

Therapeutic ultrasound (US) is frequently incorporated into treatment regimens used by physiotherapists [1]. In fact, survey data demonstrates that US is now the most frequently used electrophysical agent worldwide, used at least daily for patient treatment by the majority of physiotherapists [2–6]. This high frequency of usage makes the need for equipment accuracy imperative.

Equipment accuracy ensures that patients receive correct therapeutic dosages and underpins patient safety. In cases in which equipment fails to be accurate, two potential scenarios exist. The first is that a higher, harmful dosage may be received by the patient, potentially compromising patient safety [2,7,8]. For example, tissue destruction and blood cell stasis may occur with high doses of US therapy [1]. In the second scenario, the patient may receive a lower dosage than the therapist intended, potentially compromising treatment efficacy [9]. To ensure consistent, safe and efficacious outcomes with US therapy, machine accuracy is of the utmost importance.

The importance of US accuracy was first identified in 1956 when the United States established standards for calibration [10]. The current International Electrotechnical Commission standard for US power output is ±15% [11], with the current Australian/New Zealand standard at ±20% [12,13]. This means that the output produced by an US machine should not deviate by 20% from the value indicated on the meter [7,14]. A similar standard applies to the accuracy of the US timing device, with a ±5% difference considered acceptable [15].

Previous literature has reported startlingly high levels of inaccuracy [10,16–19]. In fact, on average, 65% of US machines have been demonstrated to be inaccurate [8,10,14,16]. However, the majority of available research was conducted more than 20 years ago, and US machines have since become digital in nature and are often

[1]School of Molecular and Biomedical Science (Physiology), University of Adelaide, [2]School of Health Sciences, University of South Australia (City East), Adelaide, and [3]School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Queensland, Australia.
Received: 6 August 2008 Accepted: 25 November 2008
Reprint requests and correspondence to: Siobhan Schabrun, School of Molecular and Biomedical Science (Physiology), University of Adelaide, North Terrace, Adelaide, SA 5000, Australia.
E-mail: siobhan.schabrun@adelaide.edu.au

multifunctional [1]. Thus, there is a paucity of published research on the accuracy of such machines. In addition, US accuracy is dependent upon several machine variables, including the intensity setting, US wave frequency (commonly 1 or 3 MHz), and whether US therapy is delivered in a continuous or pulsed mode [1]. To date, no study has examined US accuracy at the complete range of settings available for patient treatment.

Currently, routine calibration of US equipment is recommended only every 2 years [20]. Thus, there is the potential for machines producing inadequate or unsafe doses to be used in clinical practice. Therefore, the aims of this study were: (1) to examine the accuracy of clinically used US machines with respect to both power output and timing function; and (2) to investigate the features of US machines that might contribute to the degree of accuracy, such as age, brand, common intensities used and frequency of machine use.

Methods

Study design

An observational design was used to measure the accuracy of US machines in terms of power output and timing function, while an observational correlational design was used to establish the presence of any associations between US accuracy and machine age, time since calibration, frequency of use, and common intensities used for treatment. Ethical approval was granted by the University of South Australia's Divisional Ethics Committee (Health Sciences).

Machines

A power calculation (beta error, 0.8) revealed that a sample size of 64 machines was required to produce an adequate level of power. US machines from a range of physiotherapy practice environments, including major metropolitan hospitals (public and private) and private practices were sampled.

US machines were included if they had been used clinically in the last 6 months. Thus, those machines which were unlikely to have been involved with the routine calibration procedures of the sampled practice were removed. Conversely, machines were excluded if they had not been used in the previous 6 months or if they had been identified as defective with a faulty tag attached.

Materials

The power output of all US machines was tested using a portable, battery operated, digital wattmeter (Model UW-11; BioTek Instruments, Winooski, VT, USA) [21], with a resolution of 0.1 W and an accuracy of ±5% (Figure 1). The wattmeter was calibrated by Domo Technica (New South Wales), in accordance with the ISO 90001 Quality Management Standard. The device was certified accurate for a period of 1 year, under normal use. A digital stopwatch was used to test the accuracy of the US machine timers.

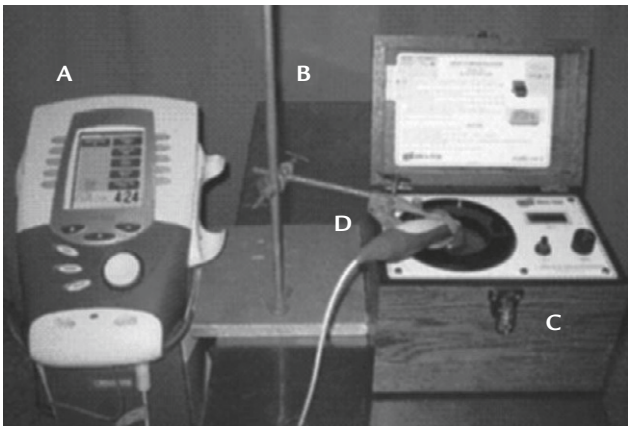

Figure 1. The testing set-up for ultrasound machine power output. A = ultrasound machine; B = retort stand; C = wattmeter; D = ultrasound head in degassed water.

Reliability testing

Prior to the commencement of data collection, the principal investigator received training in the use of the wattmeter from a biomedical engineer. Both inter- and intrareliability testing was performed. Interreliability involved both the principal investigator and the biomedical engineer to ensure that the principle investigator was well trained and competent in the use of the equipment. The study method was found to have excellent inter- and intrareliability with intraclass correlation coefficients (ICCs) of 0.99 and 0.96, respectively. Thus, only a single measurement of power output was taken for each setting throughout data collection.

Procedure

Data collection took place at the individual physiotherapy practices. Initially, a short survey was administered verbally. The survey included information about machine age, brand and model, date of last calibration, and frequency of use. During this time, the wattmeter was "rested" for 15 minutes prior to use, to stabilize it following transportation.

Following this, testing commenced. Fifty-five millilitres of degassed deionized water was added to the transducer well in the wattmeter. The water acts as a coupling agent to allow ultrasonic energy to pass from the US head to the target (cone) of the wattmeter unimpeded. To minimize effects associated with oxygenation and temperature increase, degassed water was used for no more than 60 minutes after opening, and water in the transducer well was replaced every 15 minutes [22,23].

A retort stand was used to secure the US head in position, to prevent movement and ensure repeatable

results [21]. Particular care was taken as incorrect placement can reduce accuracy [10]. The US head's surface was submerged in the water and, for consistency, positioned 5 mm from the cone of the wattmeter. The distance from the cone was kept small in order to minimize absorption of US energy as it passed through the water [15]. The wattmeter was then turned on, and the LCD display was adjusted to read 0 W. A 10-minute warm-up period was given prior to testing to ensure that readings obtained from the wattmeter were stable.

Power output testing

The power output of US machines was tested on a combination of 12 different settings, but as machines vary in their possible dosage options, not all machines were able to be tested on all 12 combinations. No warm-up stabilization period was given for the US units as the intent was to replicate how US units are used in clinical practice. A range of clinically used power outputs were tested (2, 5 and 8 W) as well as two frequencies (1 or 3 MHz) and continuous or pulsed options [1]. Testing order was randomly allocated for each machine.

During testing, it was important to minimize the accumulated amount of US energy entering the wattmeter. US energy (and heat) are known to accumulate on the target, resulting in a greater power reading and an inaccurate result [21,23]. If testing took longer than an hour in duration, the wattmeter was allowed 20 minutes to "rest" before testing continued, enabling the accumulated US energy and heat to be dispersed from the wattmeter [21].

Timer testing

The accuracy of the US timing device was determined for 5 and 10 minutes using a digital stopwatch (Technos Sports Timer, Technos, Brazil).

Calculating percentage error

The accuracy of machine output was assessed in terms of power output (watts). A small number of machines (four) displayed machine output only in terms of intensity (W/cm^2). Therefore, the intensities were converted to power output prior to analysis, using the formula: intensity (W/cm^2) = power output (W)/effective radiating area, where the effective radiating area is the manufacturer's reported value for the area of the US beam.

Percentage output error

The difference between the power output measured by the wattmeter (true power output) and the power output displayed on the US dial (indicated output) was calculated and expressed as a percentage, known as the percentage output error, using the formula: [(true power output − indicated power output)/indicated power output] × 100.

If the percentage error was greater than ± 20%, then for that particular set of dosage parameters, the power output was classified as inaccurate [13]. A similar calculation was performed for the timer accuracy, with a ± 5% standard used [13].

Statistical analysis

Descriptive statistics were used to determine the number and percentage of machine settings that produced power outputs outside the ± 20% standard error range (SPSS version 14; SPSS Inc., Chicago, IL, USA). The number and percentage of timers outside the standard error range (± 5%) at both 5- and 10-minute durations were also determined [13]. Descriptive statistics (e.g. mean, range and frequency) were obtained for each survey question. A regression analysis was performed using general linear modelling with a cluster log-binomial version to assess the effect of a number of variables on US accuracy. Significance was set at $p < 0.05$.

Results

A total of 64 machines were sampled from 31 physiotherapy practices. The majority of machines were used in private musculoskeletal physiotherapy practices (47 machines) with 11 machines drawn from the public hospital setting. The mean age of the machines tested was 10 years (standard deviation, 7.8; range, 0–30 years). The majority (90.6%) of US machines sampled were calibrated annually. The average time since calibration was 11 months (standard deviation, 11; range, 3–60 months).

Power accuracy

As multiple settings were tested on each machine, a total of 492 power tests were performed. Fifty-nine percent (291/492) of all tests were outside the recommended ± 20% standard (confidence interval, 54.76–63.43) and were thus deemed inaccurate. A total of 13 US machines (20.3%) were found to produce inaccurate power outputs on all settings tested. In contrast, three machines (4.7%) produced accurate power outputs on all tested settings.

The percentage error in machine power output ranged from −100% to +210%. Of those tests found to be inaccurate, 79% (230/291) were within the standard, delivering an accurate US output, and 11% of tests (32/291) had a percentage error that was exactly on the ± 20% standard.

Potentially, each machine could have 12 different power output tests performed if the full range of machine settings were available. The Table presents the results of machine inaccuracy for the 12 different settings tested. For tests 1–3, 68 tests were performed, as one US machine had two sound heads.

Table. Summary of findings from the 12 individual tests of power output

Test no.	Frequency (MHz)	Waveform	Power (W)	Total no. of tests	Inaccurate machines (*n*)	Inaccurate machines (%)	95% CI
1	1	CUS	2	68	40	58.8	46.9–70.0
2			5	68	33	48.5	36.8–60.3
3			8	68	33	48.5	36.8–60.3
4		PUS	2	60	37	61.7	48.9–73.3
5			5	60	30	50.0	37.5–62.5
6			8	59	28	47.5	35.0–60.2
7	3	CUS	2	15	12	80.0	54.7–94.6
8			5	15	11	73.3	47.5–90.9
9			8	15	11	73.3	47.5–90.9
10		PUS	2	21	18	85.0	65.9–96.2
11			5	21	19	90.5	72.0–98.4
12			8	21	19	90.5	72.0–98.4

CI = confidence interval; CUS = continuous US; PUS = pulsed US.

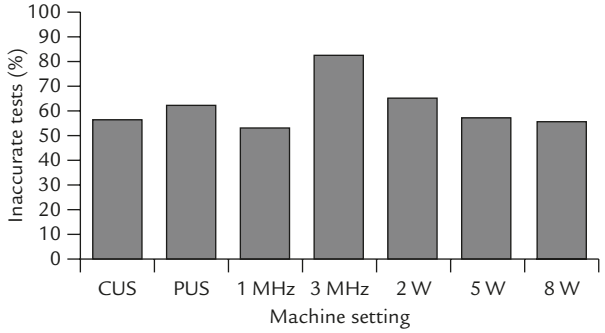

Figure 2. Percentage of tests found to be inaccurate at the various subgroups.

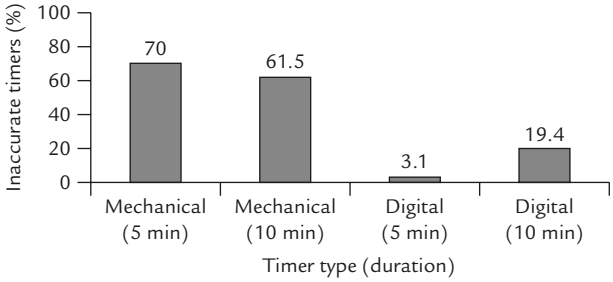

Figure 3. Percentage of timers found to be inaccurate.

Data were divided into subgroups: continuous or pulsed US, frequencies (1 or 3 MHz), and the three power settings (2, 5 or 8 W). A χ^2 test demonstrated a statistically significant difference between the two frequency settings, demonstrating that machines were more accurate at 1 MHz than at 3 MHz ($\chi^2 = 3.3677$, $p < 0.001$). No statistical difference was found between continuous or pulsed modes or the three power settings. These findings are depicted in Figure 2.

Timer accuracy

Of the 64 machines, 62 timers were assessed for 5 minutes and 57 machines for 10 minutes. The remaining machines were unable to be tested because of inadequate dial markings or because they were unable to be set for 10 minutes. Overall, 37% of timers were found to be inaccurate (range, 26.6% undertime to 93.4% overtime). Sixty-six percent (37/56) of mechanical timers were inaccurate, compared with 11% of digital timers (6/62). A χ^2 test revealed a statistically significant difference in accuracy between the two types of timers, with digital timers being significantly more accurate than analogue timers ($\chi^2 = 38.42$, $p < 0.05$) (Figure 3).

Factors including machine age, time since calibration, frequency of use and common intensities used were assessed to determine their effect on machine accuracy. A statistically significant association was found between frequency (1 MHz and 3 MHz) and machine age ($p = 0.026$). For a frequency of 1 MHz, the association was found to have an incident rate ratio of 1.025, indicating that each year added to a machine age increases the risk of inaccuracy by 1.025 times. Furthermore, at a frequency of 3 MHz, a statistically significant association was found between machine accuracy and the time since calibration ($p = 0.045$). Hence, the longer the time since calibration, the greater the risk of US machine inaccuracy.

Discussion

For patients to receive an effective and safe US treatment, it is critical that the total amount of power delivered to the tissues and the overall time of exposure are accurate [8,10]. Yet, the results of the present study demonstrate high and widespread levels of machine inaccuracy. In fact, 59% of all power output tests and 37% of timers were inaccurate. This suggests that approximately one in every two patients will receive an inaccurate dose than the one intended by the physiotherapist.

The majority of US machines in clinical use appear to be emitting a power output that does not accurately correspond to the value depicted on the dial. Interestingly, the majority of inaccurate US machines (79%) in this study were found to provide a lower power output than that indicated on the machine's dial. This finding has a number of ramifications for the treatment received by patients. Physiotherapists prescribe US dosage parameters based on a variety of considerations, including the depth of the tissue, the nature of the presenting complaint and the stage of tissue healing. An US machine that produces a power output lower than that intended by the therapist may produce an ineffective treatment, which may fail to penetrate to the target tissue or has insufficient acoustic energy for the healing process. Furthermore, 21% of the inaccurate machines were found to be above the standard, providing the tissues with more US energy than intended. This suggests that some patients were receiving an US dose which could at best counteract the beneficial healing effects of US therapy and at worst cause harmful tissue effects such as blood cell stasis [7,8,24].

The findings of this study were surprisingly similar to previous research in which, on average, 65% of machines were found to produce inaccurate power outputs. Furthermore, previous research reported that the majority of US machines produced less power output than intended and machines set at 3 MHz were less accurate than those set at 1 MHz [7,10,18,25–28]. This suggests that regardless of the change in machine technology over the last 20 years or potential differences in US usage and calibration practices internationally, two out of every three US machines produce an inaccurate dose.

If two-thirds of US machines produce inaccurate power outputs, the results of US clinical trials must also be questioned. The current evidence from randomized clinical trials for US therapy is inconclusive, with studies finding it difficult to establish evidence of a treatment effect and an appropriate dose-response relationship [1,29]. If an US machine is emitting less energy than anticipated, the treatment is likely to be ineffective. This finding may explain the lack of conclusive evidence found in clinical trials, which often report no greater treatment effect than placebo [28,30,31]. Thus, there is a need for US machines to be thoroughly calibrated prior to clinical trials and to be reported upon if the results of clinical trials are to be of value.

This study found a significant association between the time since calibration and machine accuracy [8,10]. This suggests that the longer the duration since the US machine has been calibrated, the greater the chance that the US machine will produce an inaccurate power output. However, this association was only found at a frequency of 3 MHz, and it must be noted that a smaller number of machines were tested at the 3 MHz setting (19 machines). Therefore, it is unclear if this association can be generalized to the whole US population. However, US machine age was also found to be significantly associated with machine inaccuracy. The older the machine, the more likely it was to be inaccurate. It is, therefore, possible that US units have a shelf-life and may need replacing if dosage parameters remain inaccurate despite regular calibration.

The accuracy of the US timer function was slightly better than that obtained for power output. However, over one-third of timers still returned an inaccurate result with 35% and 38.6% of timers inaccurate at 5 and 10 minutes, respectively. The majority of timers ran over the set time (62%). The use of inaccurate timers in clinical practice impacts directly on the amount of US energy delivered to the tissues. As a result, greater amounts of energy may be delivered during treatment, which could potentially cause tissue harm, compounding the effects of inaccurate power outputs [7,8,24]. Conversely, an inaccurate timer that provides less time may reduce the US dose and jeopardize the efficacy of the treatment. It is noteworthy that inaccuracy among timers was found to be significantly reduced when a digital rather than an analogue timer was used. This finding suggests that machines with inbuilt analogue timers may require more regular calibration of their timing function than those with inbuilt digital timers. This finding also supports a shift to digital timers in all new US machines available for purchase, which will ameliorate this aspect of dose inaccuracy.

Implications

This study found that US machines currently being used for physiotherapy treatment display high levels of inaccuracy and should be cause for concern. Based on these results, patients are more likely to receive a treatment dose lower than the one intended by the therapist. Failure to ensure regular calibrations and machine accuracy could have serious consequences, potentially jeopardizing patient safety and treatment efficacy.

In light of the current study's findings, it is imperative that therapists are aware of the potential for machine inaccuracy. It is recommended that therapists perform daily power output checks to ensure that US energy is being emitted from the head and also that US machines are calibrated more regularly than current guidelines suggest. The present study also found age to be associated with machine inaccuracy, suggesting that US machines may have a shelf-life and should be replaced when they no longer respond to calibration practices and repeatedly produce inaccurate power outputs.

The exact reasons for such high levels of machine inaccuracy remain unclear. Further research to examine the relationship between US machine design and the high levels of inaccuracy is necessary. In addition, future research should evaluate how often US machines require calibration to order to ensure machine accuracy.

The provision of guidelines regarding the time period after which US machines no longer respond to calibration and require replacement is also warranted.

References

1. Robertson V, Ward A, Low J, et al. *Electrotherapy Explained: Principles and Practice*, 4th edition. Sydney: Butterworth-Heinemann Elsevier, 2006.
2. Goh AC, Chok B, Wong WP, et al. Therapeutic ultrasound: rate of usage, knowledge of use, and opinions on dosimetry. *Physiother Singapore* 1999;2:69–83.
3. Lindsay D, Dearness J, Richardson C, et al. A survey of electromodality usage in private physiotherapy practices. *Aust J Physiother* 1990;36:249–56.
4. Lindsay DM, Dearness J, McGinley CC. Electrotherapy usage trends in private physiotherapy practice in Alberta. *Physiother Can* 1995;47:30–4.
5. Robertson VJ, Spurritt D. Electrophysical agents: implications of their availability and use in undergraduate clinical placements. *Physiotherapy* 1998;84:335–44.
6. Chipchase LS, Trinkle D. Therapeutic ultrasound: clinician usage and perception of efficacy. *Hong Kong Physiother J* 2003; 21:5–13.
7. Pye S, Milford C. The performance of ultrasound physiotherapy machines in Lothian Region, Scotland, 1992. *Ultrasound Med Biol* 1994;20:347–59.
8. Artho PA, Thyne JG, Warring BP, et al. A calibration study of therapeutic ultrasound units. *Phys Ther* 2002;82:257–63.
9. Pye S. Ultrasound therapy equipment—does it perform? *Physiotherapy* 1996;82:39–44.
10. Rivest M, Quirion-de Girardi C, Seaborne D, et al. Evaluation of therapeutic ultrasound devices: performance stability over 44 weeks of clinical use. *Physiother Can* 1987;39:77–86.
11. International Electrotechnical Committee. *Ultrasonics-physiotherapy Systems: Performance Requirements and Methods of Measurements in the Frequency Range 0.5 to 5 MHz*. Geneva, Switzerland: International Electrotechnical Committee, 1996.
12. Australian and New Zealand Standards. *Ultrasonics-physiotherapy Systems: Performance Requirements and Methods of Measurements in the Frequency Range 0.5 to 5 MHz (AS/NZS 4713)*. Sydney: Standards Australia and Standards New Zealand, 2002.
13. Australian and New Zealand Standards. *Guide to Testing and Calibration of Ultrasound Therapy Machines (AS/NZS Draft)*. Sydney: Standards Australia and Standards New Zealand, 2005.
14. Kollmann C, Vacariu G, Schuhfried O, et al. Variations in the output power and surface heating effects of transducers in therapeutic ultrasound. *Arch Phys Med Rehabil* 2005;86:1318–24.
15. Pye S, Zeqiri B. *Guidelines for Testing and Calibration of Physiotherapy Machines: Report No. 84*. England: Institution of Physics and Engineering in Medicine, 2001.
16. Stewart HF, Harris GR, Herman BA, et al. Survey of use and performance of ultrasonic therapy equipment in Pinellas County, Florida. *Phys Ther* 1974;54:707–15.
17. Allen KG, Battye CK. Performance of ultrasonic therapy instruments. *Physiotherapy* 1978;64:174–9.
18. Snow CJ. Ultrasound therapy units in Manitoba and Northwestern Ontario: performance evaluation. *Physiother Can* 1982;34:185–9.
19. Lloyd JJ, Evans JA. A calibration survey of physiotherapy ultrasound equipment in North Wales. *Physiotherapy* 1988; 74:56–61.
20. Robertson V, Chipchase LS, Laakso EL, et al. *Guidelines for the Clinical Use of Electrophysical Agents*. Australia: Australia Australian Physiotherapy Association, 2001.
21. Bio-Tek Instruments Inc. *Digital Ultrasound Wattmeter Model UW-II: Operators Manual*. Winooski, VT: Bio-Tek Instruments Inc., 1989.
22. Beissner K. Primary measurement of ultrasonic power and dissemination of ultrasonic power reference values by means of standard transducers. *Metrologia* 1999;36:313–20.
23. Sutton Y, Shaw A, Zeqiri B. Measurement of ultrasonic power using an acoustically absorbing well. *Ultrasound Med Biol* 2003;29:1507–13.
24. Kitchen S, ed. *Electrotherapy: Evidence Based Practice*, 11th edition. Oxford: Harcourt Publishers Ltd., 2002.
25. Ross RN, Sourkes AM, Sandeman JM. Survey of ultrasound therapy devices in Manitoba. *Health Phys* 1984;47:595–601.
26. Chapman RH. A survey of output calibrations of ultrasound therapy equipment used in physiotherapy practice. *NZ J Physiother* 1986;14:13–4.
27. Hekkenberg RT, Oosterbaan WA, van Beekum WT. Evaluation of ultrasound therapy devices. *Physiotherapy* 1986;72:390–5.
28. Guirro R, Britshcy Dos Santos SC. Evaluation of the acoustic intensity of new ultrasound therapy equipment. *Ultrasonics* 2002;39:553–7.
29. Robertson VJ. Dosage and treatment response in randomized clinical trials of therapeutic ultrasound. *Phys Ther Sport* 2002; 3:124–33.
30. Robertson VJ, Baker KG. A review of therapeutic ultrasound: effectiveness studies. *Phys Ther* 2001;81:1339–50.
31. Cunningham SJ. Calibration chaos makes research irrelevant. *Physiotherapy* 1997;83:440–1.