Background

Pain is generally assessed by subjective reports, using visual analog scales (Price et al. 1983), questionnaires, which can be converted to numeric scores (McGill, 1975) or discrete numeric scales (Price et al. 1994). Since pain perception is influenced by psychosocial factors, as well as gender, assessment of the severity of an injury from subjective reports of pain alone may be unreliable. In order to quantify pain in a more objective fashion, several studies have been conducted using a pressure algometer to measure the pressure necessary to trigger a sensation of myofascial pain. The results of these studies suggest that a trigger point for myofascial pain has a reliable and reproducible pain-pressure threshold (Reeves et al. 1986; Bendsten et al. 1996; Nussbaum and Downes 1998; Antonaci et al. 1998; Isselee et al 1997). This measure of pain still requires a subjective response on the part of the patient to indicate whether the pressure stimulus is painful so it is still not an objective means of determining the severity of soft tissue injury.

Another approach to locating sites of injury is skin thermography. Skin temperature, assessed by thermography, has been found to be a sensitive test for myofascial pain syndrome. At myofascial trigger points, temperature is higher than that of surrounding locations on the skin (Kruse and Christiansen 1992). It has also been suggested that electrical conductivity of the skin can be used as a diagnostic measure to locate tender areas in soft tissue. Acupuncture points, known to be tender areas, appear to have higher conductivity than that of surrounding tissue (Kwok et al. 1998; Comunetti et al. 1995). Thus, it might be possible to combine these two measures to detect the presence or absence of soft tissue injury in a more objective fashion than pain threshold or pain scores.

Because skin temperature and electrical resistance of healthy tissue may vary among individuals, measurements of a relative nature would be more useful than measurements of an absolute nature. For example, differential values could be used, where the difference is calculated between a measurement made at a reference site known to be uninjured and a test site. A test site would be classified as a site of injury if the differential in skin temperature and electrical conductivity were significantly higher than the differentials between uninjured sites and the reference site. Thus, it would first be necessary to calculate differentials between various sites known to be uninjured and a reference site. A distribution of differentials for uninjured sites would then be established. To determine whether a test site was a site of injury, the differential between the test site and the reference site would have to fall outside the range of differentials for uninjured sites.

A new technology called ASSESSx has been developed to measure skin temperature, skin conductivity, crepitus (sound), and applied force. A review of the literature has indicated that skin temperature and skin conductivity may be reliable measures for locating trigger points for myofascial pain. Therefore, we have conducted a controlled study to determine whether measurement of these quantities with the ASSESSx

technology could be used to accurately locate sites of myofascial pain. In particular, we examined subjects with and without myofascial pain in the neck and shoulder region.

Objective

The primary objective of the study was to establish the validity, accuracy, and reliability of ASSESSx technology in locating sites of myofascial pain.

We tested the hypothesis that differentials in skin temperature and skin conductivity between a reference site and test points on healthy tissue (judged by high pain threshold) were significantly lower than differentials between the reference site and sites of myofascial pain associated with whiplash injury.

Methodology

There were two aspects to validating the technology. The first was to show that it accurately measures temperature, electrical conductivity and applied force. The second was to show that skin temperature and conductivity differentials, were correlated with the intensity of myofascial pain. There were also two aspects to determining the reliability of the technology. The first was to show that measurements are repeatable, i.e., that repeated measurements at the same sites were consistent, whether made by one or more examiners. The second was to show that a range of normal values could be established for differentials in skin temperature and conductivity and that those values outside the normal range were reliable indicators of myofascial trigger points in different individuals.

Instrument Testing

Measurement accuracy was assessed by comparing readings from the ASSESSx instrument with calibrated standards. We first compared force measurements made with the ASSESSx instrument with those of a calibrated force transducer (ATI F/T 3175). Three sets of measurements were made, consisting of 77, 86, and 90 measurements in the range of 0.5 to 10.5 lbs. A regression line was then fit to the data. The average regression coefficient for the three trials was 0.96 and the average value of r2 was 0.99. A regression coefficient of 1.0 and an r2 value of 1.0 would have indicated perfect agreement. Both were sufficiently close to 1.0 to conclude that the ASSESSx instrument accurately measures force. The results from the three trials were very similar, indicating that the force measurements were reliable.

To test the accuracy of temperature measurement, the ASSESSx instrument was placed in a closed chamber that was gradually heated from room temperature to approximately 40º C. This procedure was repeated three times. The readings were compared with readings from a calibrated thermistor (BoxcarPro) placed next to the ASSESSx instrument. The ASSESSx reading was an average of 2.5° C higher at room temperature than the reading of the thermistor, but increased more slowly during heating. The final temperature recorded by the ASSESSx instrument was 40.4° C, while the thermistor reading ranged from 46-48° C. We found that the relation between the ASSESSx reading and the thermistor reading was best represented by an exponential equation. The average r2 value

for the three trials was 0.98. Although the temperature reading was not accurate, the fact that the r2 value was so close to 1.0 indicated that it was predictable and could be corrected. The difference between the temperature reading of the ASSESSx instrument and the thermistor is probably partly due to a slow response time, i.e., the ASSESSx instrument takes longer to reach a steady-state temperature than the thermistor. The results from the three trials were very similar, indicating that the temperature measurements were reliable.

The ability of the ASSESSx instrument to measure electrical conductivity was tested by measuring the conductivity of a set of 18 calibrated resistors from 0.01 to 22 MΩ. We found that the ASSESSx readings matched the conductivity of the resistors almost exactly for values between 5 x 10-7 and 10-5 Ω-1. The regression coefficient was 1.0 and the r2 value was 1.0. The ASSESSx instrument was somewhat less accurate for lower conductivities. Using a range of 4.5 x 10-8 to 10-5 Ω-1, the regression coefficient was 1.3 and the r2 value was 0.99. The higher value for the regression coefficient was due to inaccurate readings below 5 x 10-7 Ω-1, i.e., the ASSESSx conductivity readings were consistently higher than the true values and the errors became larger as the conductivity became very low.

Subjects

Twenty subjects, ranging in age from 18 to 49 years old, were examined. Twelve of the subjects were female and eight were male. Six subjects (three male and three female) reported that they had experienced a whiplash injury within the past six months and were still experiencing pain in the neck and shoulder region. We were unable to recruit more subjects with whiplash injury despite considerable effort. We advertised extensively on the campus of Simon Fraser University, sent requests to about 40 local chiropractors and to 10 local physiotherapy clinics. The remaining 14 subjects were recruited as healthy subjects from among students and staff at Simon Fraser University. None of them reported that they had never experienced musculoskeletal injury to the shoulder or neck prior to the examinations. The first subject tested was female with whiplash injury. She served only as a test subject for training of the examiners since she asked that her data not be included in the analysis.

The mean age of the subjects with whiplash injury, who were included in the analysis, was 25.6 (S.D. 1.7) years. The mean age of the subjects, reporting no neck or shoulder injury, was 24.1 (S.D. 7.8) years. The difference in mean age was not statistically significant.

Protocol

Subjects were asked to read an information sheet and to sign a consent form. Copies of both are provided in the Appendix. The subject was introduced to the first examiner and the research assistant and was informed that a second examiner would conduct a second examination later. Examinations were conducted in a corner of the Biomechanics Laboratory at Simon Fraser University, partitioned with dividers and curtains for privacy. During the examination, the subject lay face down on a bed with the head supported on a

pillow if desired. After the two examinations, the subject completed a modified version of the McGill Pain Questionnaire, which is attached as part of the Appendix.

The examiner applied a constant force of approximately two pounds with the tip of the ASSESSx instrument at five locations on the neck and shoulder on each side of the body (Fig. 1). The order in which these sites were tested was randomized for each examination. The research assistant instructed the examiner as to the next site in the sequence, during the course of the examination. The examiner applied force for 10 seconds, after which the force was reduced to near zero, without lifting the instrument from the skin. The correct force was indicated to the examiner by illumination of a green light on the instrument. Red lights on either side of the green light indicated if the force was too high or too low. The research assistant continuously monitored force, skin temperature and skin conductivity, which were printed on a computer display, facing away from the examiner. The examiner could not see the computer display and was not informed that the instrument measured skin temperature or skin conductivity. At the end of the 10-second period, the research assistant recorded the final values of force, skin temperature and skin conductivity on the computer and on a data sheet.

The examination began by recording a reference baseline measurement at a neutral test location. The neutral location was chosen as a point on the center of the spine in the middle of the back, just below the trapezius muscle (Fig. 1). Subjects did not report experiencing pain when force was applied at this location. Following each measurement, the subject was asked to indicate the level of pain experienced during the time that the force had been applied. The pain response was given by pressing a handheld keypad. The keypad had four positions, labelled one to four. If no pain was experienced, the subject was instructed not to press the keypad. A rating of four would have indicated extreme pain. No subject ever rated the pain level as four. The highest rating was three. The examiner turned away from the subject during the response period so that only the research assistant had knowledge of the subject's response, which appeared on the computer display. The pain response was also recorded on the data sheet.

The first examiner conducted two examinations at each of the test locations, after which the second examiner conducted two examinations. The examiner, who began first varied, such that the number of subjects for which each examiner began first was approximately equal.

Analysis

Skin temperature and skin conductivity differentials were calculated by subtracting the baseline reference value from each recorded value. The differentials and pain responses were tested for intra-rater reliability by comparing the first and second measurement of each examiner for each anatomical location. Inter-rater reliability was assessed by comparing the mean of the two measurements, obtained by the first examiner at each anatomical location, to the mean of the two measurements, obtained by the second examiner. The Spearman rank correlation and the Kendall rank correlation were used as measures of reliability. Where appropriate, linear regression was also used to examine the strength of trends in the data. The differentials were also classified according to pain

response. The resulting distributions were examined for evidence of bimodality to determine whether they could be used to discriminate painful responses from non-painful responses.

Results

The force measurements indicated that the examiners were able to apply the required force accurately and consistently. The skin temperature measurements stabilized within the 10 s period of force application. However, the skin conductivity measurements were generally not stable, continuously increasing throughout the period of force application.

Intra-rater reliability was tested by comparing the measurements obtained by each examiner for each subject for the two examinations at each anatomical location. The skin temperature differentials from measurements performed by Examiner 1 and Examiner 2 are shown in Figs. 2 and 3, respectively. The agreement for Examiner 1 was very good when the temperature differential was greater than -2° C. A regression line fit to the data for temperature differentials greater than -2° C had a slope of 1.04 with an r2 value of 0.89. However, there was a systematic offset of 0.49° C, indicating that the temperature recorded during the second examination was slightly higher than the first. The agreement was also good for Examiner 2 when the temperature differential was greater than -1° C, although the slope was not as close to 1 (0.91) and the r2 value was lower (0.73). There was also a systematic offset, although it was somewhat lower (0.38° C). The most likely reason for the offset was that the reference temperature for computing the differential was always recorded at the beginning of the first examination. Consequently, any warming of the ASSESSx instrument over time would result in a higher differential during the second examination. The Spearman rank correlation for the combined data obtained by the two examiners was 0.89 and the Kendall rank correlation was 0.74, indicating very good agreement between the first and second examination. Rank correlations of one would indicate perfect agreement.

There was much less agreement between the two examinations for skin conductivity differential. From Figs. 4 and 5 it is evident that there is much more scatter in the data than for the temperature differential and that the scatter is similar for the two examiners. The Spearman rank correlation was only 0.066 and the Kendall rank correlation was only 0.045. This indicates that there was little consistency between the first and second examinations. Such inconsistency is in agreement with the observation that the skin conductivity reading did not stabilize during the recording period.

Inter-rater reliability was tested by comparing the mean of the two measurements obtained by the two examiners for each subject at corresponding anatomical locations. The data for skin temperature differential is shown in Fig. 6. It is evident that the agreement is poor. The Spearman rank correlation was only 0.30, while the Kendall rank correlation was 0.19. The data for skin conductivity differential is shown in Fig. 7. The agreement is only slightly better, with a Spearman rank correlation of 0.39 and a Kendall rank correlation of 0.27. The poor inter-rater reliability is not due to differences in the applied force since this was kept in the same range by the two examiners. It is more likely due to differences in technique or inconsistencies in locating the anatomical test sites.

For example, the research assistant noted that Examiner 1 occasionally lifted the tip of the ASSESSx instrument from the skin when moving between test sites, while the Examiner 2 kept the tip of the instrument in contact with the skin throughout the entire examination.

We checked the consistency of the subject's pain response for the two examiners and for the two measurements made by each examiner at each anatomical location. There was greater consistency of responses in examinations conducted by Examiner 2 (Spearman rank correlation 0.71; Kendall rank correlation 0.67) than in those conducted by Examiner 1 (Spearman rank correlation 0.54; Kendall rank correlation 0.50). This may be related to greater care taken by Examiner 2 in conducting the examination. We also compared the consistency in the pain response elicited by one examiner compared to the other. This was done by calculating the mean pain response of the two examinations conducted by each examiner at each anatomical location. There was moderately good agreement (Spearman rank correlation 0.74; Kendall rank correlation 0.67). The difference between examiners may have been related to differences in the exact location where force was applied and trial-to-trial variability in the subject's judgment of pain.

When the distribution of skin temperature differentials and skin conductivity differentials for healthy subjects and subjects with whiplash injury were compared, we did not find two clearly distinct distributions (Figs. 8 and 9). Instead, there was considerable overlap in the distributions, although for one or two whiplash subjects the differentials were clearly higher than for the healthy subjects. This appeared to be more so for higher pain responses. As a result, it possible to distinguish one subject with whiplash injury based on skin temperature differential and one subject with whiplash injury based on skin conductivity differential. However, the differentials for the other three subjects with whiplash injury could not be distinguished from those of healthy subjects.

We then plotted the skin conductivity differential versus the skin temperature differential (Fig. 10) to determine whether this might better distinguish subjects with whiplash injury from healthy subjects. Figure 10 illustrates that two clusters of data points lay outside the range of differentials for healthy subjects. One of these clusters is along the conductivity axis, with values less than –4 and the other is along the temperature axis with values greater than 5 and negative conductivity values. The cluster along the conductivity axis identifies one subject, while the other cluster identifies two subjects. Thus, three of the five subjects with whiplash injury could be distinguished from healthy subjects, based on combined skin temperature differentials and skin conductivity differentials. However, these distinguishing clusters of data points all correspond to examinations made by Examiner 1 only. From the measurements made by Examiner 2, it would not have been possible to distinguish subjects with whiplash injury from healthy subjects. Furthermore, about a third of the data points in these clusters come from measurements in which the subjects reported no pain.

Although none of the healthy subjects had experienced musculoskeletal injury to the neck or shoulder, seven of them gave pain responses of two or three. Six of these subjects also indicated that they had experienced pain during the examinations, based on their responses to the McGill Pain Questionnaire. We, therefore, decided that it would be

informative to combine the data from all subjects and to compare the skin temperature and skin conductivity distributions for measurements where pain was reported (pain scores greater than 0) and for measurements where pain was not reported (pain score of 0). The result is shown in Fig. 11. The majority of data points from the two distributions overlap. It is, therefore, clear that in the majority of cases the skin temperature and skin conductivity differentials could not be used to distinguish a painful test site from a non-painful test site.

Conclusions

Although there is some hint that sites of whiplash injury might be distinguished by higher (positive) skin temperature differentials and lower (negative) skin conductivity differentials than uninjured sites, our results demonstrate that there is a great deal of inconsistency between measurements by the two examiners with the ASSESSx instrument. Furthermore, while skin temperature differentials obtained from repeated measurements by the same examiner are consistent, there is a lack of consistency in skin conductivity differentials. Since we determined that, the ASSESSx instrument reliably and accurately measures the electrical conductivity of a fixed resistance, either skin conductivity is dynamic or our protocol was inappropriate. Before any further studies are conducted to examine skin conductivity, a protocol must be developed which allows stable measurements to be obtained.

Our study was purposely designed to use a simple protocol that would be easy for inexperienced examiners to repeat reliably. Consequently, we did not attempt to locate sites where force application would trigger the maximum myofascial pain. It may be that larger differentials could be obtained by searching for such trigger points, which would distinguish between healthy subjects and subjects with myofascial injury. However, only an experienced physiotherapist or massage therapist could do this.

In its current state, the ASSESSx instrument cannot be used to reliably distinguish subjects with whiplash injury from healthy subjects. However, if the problem with inconsistency in the skin conductivity measurements could be resolved, it would be worthwhile repeating the study with a larger group of subjects with whiplash injury.

If the study is repeated we would also recommend that the examiners be experienced physiotherapists or massage therapists so that a protocol could be used in which a myofascial trigger point, eliciting the greatest pain response, is first identified and measurements are then made at this specific anatomical location and compared with measurements made at the same anatomical location in a healthy subject.

References

Antonaci, F., Sand, T. and Lucas, G.A. (1998) Pressure algometry in healthy subjects: inter-examiner variability. Scand. J. Rehabil. Med. 30: 3-8

Bendsten, L., Jensen, R. and Olesen, J. (1996) Qualitatively altered nociception in chronic myofascial pain. Pain 65: 259-264

Comunetti, A., Laage, S., Schiessl, N. and Kistler, A. (1995) Characterisation of human skin conductance at acupuncture points. Pharma Clin. Res. 51: 328-331

Isselee, H. De Laat, A., Lesaffre, E. and Lysens, R. (1997) Short-term reproducibility of pressure pain thresholds in masseter and temporalis muscles of symptom-free subjects. Eur. J. Oral Sci. 105: 583-587

Kruse, R.A. and Christiansen, J.A. (1992) Thermographic imaging of myofascial trigger points: a follow-up study. Arch. Phys. Med Rehabil. 73: 819-823

Kwok, G., Cohen, M. and Cosic, I. (1998) Mapping acupuncture points using multi channel device. Australas. Phys. Eng. Sci. Med. 21: 68-72

Melzack, R. (1975) The McGill Pain Questionnaire: major properties and scoring methods. Pain 1: 277-299.

Nussbaum, E.L. and Downes, L. (1998) Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. Phys. Ther. 78: 160-169

Price, D.D., McGrath, P.A., Rafii, A. and Buckingham, B. (1983) The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 17: 45-56

Price, D.D., Bush, F.M., Long, S. and Harkins, S.W. (1994) A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. Pain 56: 217-226

Reeves, J.L., Jaeger, B. and Graff-Radford, S.B. (1986) Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. Pain 24: 313-321