Title	Effect of skin surface stimulation on acupoints for phonotraumatic injuries
Author(s)	Cheung, Lok-chun, Benjamin; 張樂儁
Citation	Cheung, L. B. [張樂儁]. (2013). Effect of skin surface stimulation on acupoints for phonotraumatic injuries. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
Issued Date	2013
URL	http://hdl.handle.net/10722/238524
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Effect of skin surface stimulation on acupoints for phonotraumatic injuries

Cheung Lok Chun Benjamin

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2013

Effect of skin surface stimulation on acupoints for phonotraumatic injuries

Cheung Lok Chun Benjamin

Abstract

From a recent large-scaled randomized controlled trial, it was found that genuine acupuncture was not significantly different from the placebo, or skin surface stimulation, in improving the vocal functions of dysphonic subjects with benign vocal pathologies. This randomized, double-blind, placebo-controlled study investigated the effects of skin surface stimulation on acupoints for treating phonotraumatic injuries. Four female subjects were randomly assigned to either a treatment group or a placebo group. Subjects in the treatment group received skin surface stimulation on voice-related acupoints *Hegu* (Li4), *Lieque* (Lu7), *Lianquan* (CV23), *Renying* (St9) and *Zhaohai* (Ki6), whereas subjects in the placebo group received skin surface stimulation on non-voice related acupoints *Tianchuang* (SI16), *Neiguan* (PC6), *Zhongzhu* (SJ3) and *Shangqiu* (SP5). All subjects completed 12 sessions within a 6-week timeframe. Outcome measures included aerodynamic measures, the voice range profile, and self-perceived voice-related quality of life by subjects. Limited by a small sample size, significant changes over time were not found in both groups from pre-treatment to post-treatment and 14 days post-treatment. There were no conclusive results regarding whether skin surface stimulation on acupoints is an effective treatment for improving the vocal functions and quality of life in patients with phonotraumatic injuries.

Effect of Skin Surface Stimulation on Acupoints for Phonotraumatic Injuries

Phonotrauma refers to pathological changes in the laryngeal tissues due to hyperfunctional vocal behaviors (Colton, Casper, & Leonard, 2006). The formation of vocal fold nodules, polyps or edema, and thickening of the vocal folds often result in change in voice quality and functions. Voice disorder affects 3% to 9% of the general population (Aronson & Bless, 2009). Conservative treatment such as voice therapy and vocal hygiene programs are often preferred over medications or surgical excision of the laryngeal lesion (Ramig & Verdolini, 1998). However, not all patients are responsive to conventional therapies, as the resolution of the vocal pathology is not always complete (Leonard, 2009; Nakagawa, Miyamoto, Kusuyama, Mori, & Fukuda, 2012; Woo, Casper, Colton, & Brewer, 1994).

Acupuncture, as an alternative treatment approach, has been shown to be effective in treating phonotraumatic voice disorders (Yiu et al., 2006). More recent findings (Yiu et al., in preparation) have also found that stimulation on the skin surface of acupoints produced significant effects as much as needling in genuine acupuncture. The current study served to further investigate the effects of skin surface stimulation on acupoints in treating phonotraumatic voice disorders associated with benign vocal fold pathologies.

Establishing the scientific evidence of skin surface stimulation effect on phonotraumatic vocal fold lesion will promote the adoption of this alternative treatment approach that speech therapists can undertake. Being a non-invasive procedure, risks associated with needling, such as bleeding, trauma to internal organs and infections of blood-borne pathogens (Ernst, 2006), can be eliminated in skin surface stimulation. Findings determined in this study would have clinical and theoretical significances in the field of Traditional Chinese Medicine in addition to voice science.

Traditional and modern theories of acupuncture

Acupuncture, being first practiced in China for over 2000 years ago (Birch & Felt, 1999), is gaining popularity among the general public and researchers in the West. The principles of acupuncture originate from theories in Traditional Chinese Medicine. *Qi* is the vital energy that flows through the meridians, or channels that connect the internal organs (Birch & Felt, 1999). Acupoints or *xue* are sites that *qi* and blood flow on the surface of the body along the meridians. By dredging the obstructed meridians through the stimulation of specific acupoints with needles, *qi* can be regulated to restore the homeostatic balance between *yin* and *yang* (Shi, 2007). Therapeutic effect of acupuncture making use of other modalities of stimulation, including electricity, laser, pressure and temperature, has been documented (World Health Organization, 2002).

Acupuncture has been studied for possible neurological substances using a scientific model. The growth control model described in Shang (2009) is a model that has met the gold standard of science. Phenomena observed in acupuncture therapy, such as therapeutic effects achieved through non-specific stimulation, long-term healing effects, and the functions of acupoints and meridians can be explained using the model. Shang (2009) suggests that acupoints originate from the growth points in the system of embryogenesis called *organizers*. By stimulating these growth points, long term growth effects or regulatory effects on neural, circulatory and immune processes can be triggered. Wound healing is accelerated as a result of an excessive release of endorphins and other neurohumoral factors, such as growth hormones and basic fibroblast growth factor (Shang, 2001). When an acupoint is stimulated, *Qi* is channeled to corresponding internal organs through the meridians according to the Traditional Chinese Medicine model. In the growth control model, meridians are interpreted as boundaries having high electrical conductance where the organizers or acupoints are located. The acupoints and meridians are within a system that activates growth control in an organism.

Outcomes in basic science research in acupuncture converge with the emerging theories for the wound healing and anti-inflammatory actions of acupuncture therapies. Lee, Jeong, Park, Jeon, and Hong (2010) demonstrated faster skin regeneration following deep second degree burns following acupuncture treatment compared with Duoderm dressing in mice. A study by Park et al. (2012) showed that acupuncture improved wound healing using a rat model by demonstrating enhanced cell proliferation, angiogenesis, extracellular matrix remodeling, and a decreased pro-inflammatory response. Vocal fold nodules, polyps and cysts are benign lesions that can be viewed as manifestations of chronic phonotrauma or arrests at the inflammatory phase of the vocal fold wound healing cascade (Branski, Verdolini, Sandulache, Rosen, & Hebda, 2006). Theories on wound healing processes in the vocal folds are still being developed, but it is hypothesized that the processes resemble those that occur in skin, joint and other body tissues, and is mediated by cytokines, growth factors and other biochemical markers (Branski, Rosen, Verdolini & Hebda, 2004). It is hypothesized that for acupuncture to be effective for treating benign laryngeal lesions, the benefits are mediated by facilitating wound healing.

Acupuncture for voice disorders

Efficacy studies of acupuncture for the treatment of benign vocal fold lesions or hyperfunctional voice disorders are limited. Many acupuncture studies have focused on studying the analgesic effects of acupuncture and the treatment of musculoskeletal disorders (Du, Xiong, Huang, Li, & Shi, 2009). In China, where conservative behavioral voice therapy is not available, acupuncture on hyperfunctional dysphonia is applied and studied widely. Most studies, however, relied on perceptual evaluation alone to draw conclusions. The majority of the studies reported the use of vague, ill-defined descriptions and qualitative measures. Only a limited number of the studies used trials with objective outcomes measures.

Wang and Tao (2005) compared the effects of acupuncture on treating vocal fold nodules with other therapies. Eighty subjects were randomized to either the acupuncture group (N = 40) to receive acupuncture on the acupoint *Kaiyin No. 1* [located at 1 to 1.5 cm lateral to *Renying* (St9)]; the Chinese herb group (N = 20) to take *Jinsang Sanjie* pills; and the Western medicine group (N= 20) to receive spray inhalation of hormones and antibiotics. A three-point perceptual rating scale (cured, effective, ineffective) was used for assigning different levels of treatment effectiveness upon stroboscopic examination. Acoustic measurements of fundamental frequency, jitter, shimmer and the magnitude of glottal noise energy were also used. Statistical significance was found when the authors compared the total effective rate in the acupuncture group (92.5%) with the Chinese herb group (85.0%) and the Western medicine group (85.0%). The improvement in acoustic parameters for the acupuncture group was found to be statistically significant (p <.01) and superior to the other two groups (p <.01). The authors concluded acupuncture at acupoint *Kaiyin No. 1* showed stronger therapeutic effects for treating vocal nodules than other treatments. Nevertheless, the standards used to measure treatment effects in the perceptual rating scale were qualitative and subjective. Procedures to ensure objectivity in stroboscopic examination of the glottal closure, redness and lesion size were lacking. Moreover, the investigators failed to state the norms and the standardization process for the acoustic parameters.

A large-scale study (N = 160) by Yang, Xie, and Jiang (2006) investigated the effects of acupuncture and Western medicine on treating submucosal hemorrhage of vocal folds. The subjects were randomized into the acupuncture group (N = 80) to receive acupuncture at acupoint *Kaiyin No. 1* and practice deep respiratory exercises, and the Western medicine group (N = 80) to receive medications through spray inhalation. Outcome measures and findings were similar to that reported by Wang and Tao (2005), and similar limitations in the methodological designs also restricted the interpretation of the study.

Another study by Sun et al. (2010) compared the effectiveness of acupuncture and routine medication with routine medication alone on the rehabilitation of vocal folds post-surgery of vocal polyps. Sixty subjects were randomized into the acupuncture group (N = 30), which received needling on *Sheng's* Four Points of Throat and spray inhalation of Western medication, while the control group (N = 30) received medication only. The effectiveness of the acupuncture group with medication (83.3%) was determined to be significantly higher (p <.05) than the control group with medication only (60%). The investigators also employed a rating scale for outcome measures similar to the previous two studies. In addition, they also used more objective standards such as quantitative increase of the maximum sustained phonation. Norms of acoustic parameters such as jitter, shimmer and magnitude of glottal noise were stated explicitly. However, the authors did not report any measures to quantify the size of the vocal fold lesion. There were also no details on how the norms for acoustic measures were obtained.

A study by Yiu et al. (2006) investigated the efficacy of treating benign vocal fold pathologies with acupuncture using a randomized, placebo-controlled design. The double-blind design controlled for experimenter effects and improved the validity of the results. Outcome measures used included the voice range profile, perceptual rating of the voice, and self-perceived quality of life measured by the Voice Activity and Participation Profile (VAPP; Ma & Yiu, 2001). Twenty-four female subjects were randomized into the experimental group (N = 12) and the placebo group (N = 12). Subjects in the experimental group received acupuncture on 3 pairs of voice-related acupoints: *Renying* (St9), *Lieque* (Lu7) and *Zhaohai* (Ki6). Subjects in the placebo group received acupuncture on 2 pairs of non-voice related acupoints: *Houxi* (Si3) and *Kunlun* (Bl60). Significant improvement was found in the maximum fundamental frequency, frequency range, perceptual breathiness and roughness, Total VAPP Score, Self-Perceived Severity Score and Emotion Score in the

experimental group only. This study, being the first published randomized controlled trial for acupuncture on voice disorders, employed a multi-dimensional approach to evaluate objective treatment outcomes. Nevertheless, there was not a no-treatment group to control for spontaneous recovery. In addition, being a preliminary trial, the sample size is considered small for drawing sound conclusions for clinical application.

Paradoxes in clinical trials for acupuncture

Recently, a large scale (N = 167) double-blind randomized controlled trial was carried out by Yiu et al. (in preparation) to evaluate the efficacy of acupuncture to treat benign vocal fold lesions. Subjects in the treatment group (N = 58) received genuine acupuncture on voice-related acupoints: 2 *Hegu* (Li4), 2 *Lieque* (Lu7), 1 *Lianquan* (Cv23), 2 *Renying* (St9) and 2 *Zhaohai* (Ki6). Subjects in the placebo group (N = 59) received sham acupuncture with needle stimulating the skin surface at the same acupoints in the treatment group. Subjects in the control group (N = 50) did not receive any treatment, and only attended the voice assessment sessions. Outcome measures used included the voice range profile, stroboscopic examination, and self-perceived quality of life. The authors reported similar improvements in vocal functions post-treatment in subjects receiving genuine acupuncture and skin surface stimulation on acupoints. No improvement was shown in the no-treatment group.

Langevin et al. (2011) discusses acupuncture treatment by breaking it down into specific needling components, specific non-needling components and non-specific components. Needling components such as the location, insertion depth, modality for stimulating acupoints or skin contact, needle size and number are commonly varied or controlled in placebo-controlled clinical trials. Non-specific components refer to the creditability and expectation of the acupuncturist and the patient, the clinical setting, or other factors that are not unique to acupuncture. It is commonly believed that therapeutic effects of acupuncture

are channeled through needling components. Findings by Yiu et al. (in preparation) were similar to the paradox discussed in Langevin et al. (2011). They showed that sham acupuncture with skin stimulation on acupoints produced efficacious results as genuine acupuncture in a number of well-designed clinical trials studying various disorders and conditions. Moffet (2008) reported 46.4% (26/56) of the efficacy trials published in English in 2006 that are placebo-controlled did not show statistical difference in treatment outcomes. To account for this phenomenon, one explanation would be the therapeutic mechanism of acupuncture is entirely, or to a very large extent, due to non-specific effects. At the same time, sham acupuncture having retained therapeutic specific needling components as in genuine acupuncture is also probable (Langevin et al., 2011).

From the efficacy studies of acupuncture on benign vocal pathologies (Yiu et al., 2006; Yiu et al., in preparation), treatment effects are observed in groups which voice-related acupoints are stimulated, regardless being needling or skin surface stimulation. This is in accordance to the theory of non-specific stimulation of the organizers, or acupoints, stated in the growth control model (Shang, 2009). Thus, we hypothesize that when non-specific stimulation exerts a physiologic wound healing effect directly, skin surface stimulation on voice-related acupoints, presented non-deceptively, will evoke therapeutic changes to the larynx. Given the results in Yiu et al. (2006), it was hypothesized that only skin surface stimulation on voice-related acupoints, but not on non-voice related acupoints, will give positive results. Provided that subjective, self-reported measures are susceptible to enhanced placebo effects in sham acupuncture (Kaptchuk et al., 2006), we also hypothesize that improvement in the voice-related quality of life will not be significantly different for the two groups. This randomized, double-blind, placebo-controlled study aims to investigate the efficacy of skin surface stimulation on voice-related acupoints on improving vocal and aerodynamic functions, and self-perceptions of voice-related quality of life by participants

with benign vocal fold lesions.

Method

Participants

Six Chinese subjects were recruited to participate in the study. The subjects aged between 20 and 55 years to rule out changes in voice quality due to puberty or aging (Titze, 1994; Jasper & Colton, 2000). All subjects obtained a laryngological diagnosis of phonotraumatic injuries with vocal fold mass lesion from an otolaryngologist within three months prior to the study. All subjects self-reported no history of asthma, a diagnosis of or known neurological conditions, hearing problems and having attended voice therapy before. During the time of the study, they were not having an upper respiratory tract infection, or receiving medical or other therapies that may affect voice or laryngeal condition. Four female subjects (mean age = 37.11 years; SD = 12.10 years; range = 21-50 years) had a complete set of data for analysis. They had been randomly assigned into the treatment group (N = 2) to receive skin surface stimulation on voice-related acupoints, or the placebo group (N = 2) to receive skin surface stimulation on non-voice related acupoints. Two (50%) were teachers, one (25%) was a nurse and one (25%) was a student. Two subjects were diagnosed with vocal nodules, one had a vocal polyp, and one had vocal fold thickening.

Outcome measures

1) Aerodynamic measures

Aerodynamic evaluations were carried out using the KayPENTAX *Phonatory Aerodynamic System*, Model 6600. Calibration of the system was performed according to the instructions from the manufacturer. Tasks chosen and recording procedures were adapted from Ma and Yiu (2006). Subjects were asked to put on a facemask tightly over the mouth and the nose during recording to ensure air seal. All subjects completed five trials of recording for each of the four speech tasks listed below. In the maximum sustained

phonation task, each subject was asked to inhale deeply, and then produce the longest possible sustained /a/ at a comfortable pitch and loudness. In the most comfortable phonation task, each subject was asked to sustain an /a/ for five seconds at a comfortable pitch and loudness. To measure the peak intraoral pressure, the remaining two tasks were performed with the subjects maintaining a flexible silicon rubber tube, which was connected to the air pressure transducer, centrally on top of the tongue. In the nonsense vowel-consonant string production task, each subject was asked to produce the vowel /i/, followed by seven consecutive syllables comprised of the bilabial plosive /p/ and the vowel /i/ at a rate of 1.5 seconds per syllable. The subjects were reminded to keep the stress of each syllable equal. In the sentence production task, each subject produced a Cantonese sentence /ba ba da bɔ/ (“Father hits the ball”) for five trials.

2) Voice range profile recording

Voice range profiles were recorded using the Swell *Real-time DSP Phonetograph* Version 2.0 (*Phog* 2.0, AB Nyvalla DSP) with a head-mounted condenser microphone (AKG Acoustics C520). Calibration and recording procedures were adapted from Ma (2011). The *Phog* 2.0 system was first set to convert the intensity level of signals captured at a 5cm distance, the microphone-to mouth distance used in this study, to resemble a source-recording distance of 30 cm. A sound level meter (Rion Co., Ltd., Model NL-20), placed in line with the microphone, was held at 30 cm away from a loudspeaker. The intensity level of a 1-kHz sinusoidal tone generated by the *Phog* 2.0 system and directed through the loudspeaker was registered. Adjustments at the DSP Mixer were made until the intensity level displayed by the system and sound level meter were within 1 dB difference to complete calibration. A sinusoidal 261.6 Hz tone, equivalent to C4 on the piano scale, was presented to the subject at the start of the recording. The subject was instructed to sustain the vowel /a/ at a comfortable intensity level with reference to this tone, and then gradually decrease the

loudness until the lowest intensity level was attained. This procedure was repeated by going down and up the piano scale for every musical note until the subject was unable to sustain phonation at any lower or higher pitch to obtain the lower intensity contour of the voice range profile. To obtain the higher intensity contour, the subject repeated similar steps except to gradually increase the loudness until the highest intensity level was attained.

3) Voice Activity and Participation Profile (VAPP)

VAPP is a 28-item self-report questionnaire that measures voice-related quality of life. It evaluates self- perception of voice problem, and the activity limitations and participation restrictions on job, daily communication, social communication and emotion. A Chinese version which employs an 11-point equal- appearing interval scaling procedure (Ma & Yiu, 2007) was used.

Procedures

All subjects received 12 sessions of skin surface stimulation therapy, and underwent three assessment sessions throughout the study: a pre-treatment assessment before the first therapy session (PRE), a post-treatment assessment conducted immediately after the 12th therapy session (POST), and a follow-up assessment conducted 14 days after the twelfth therapy session (FU). The outcome measures obtained were aerodynamic measures, voice range profile recording, and VAPP. The assessments were conducted in a sound-treated laboratory. The sequence of the tasks was randomized within subjects to eliminate possible order effects on the outcome measures. The raters were blind to the group allocation of the subjects.

Skin surface stimulation on acupoints

The subjects were randomly assigned to the treatment or placebo group to complete 12 skin surface stimulation sessions within six weeks. They were blind to the group allocation. A qualified acupuncturist registered in Hong Kong was responsible for conducting all the

skin surface stimulation sessions. The acupuncture needles used were disposable stainless steel needles manufactured by Tai Chi. They were 0.3 mm in diameter, 25 mm in length, and had blunt ends.

In each session, the subjects were asked to relax and lie supine on a couch in a quiet room. Swabs saturated with 75% alcohol were used for disinfecting the skin area over the acupoints to be stimulated. Foam stands with double-sided adhesives were taped onto the designated skin surface sites, and plastic guide tubes of 30 mm were then inserted into the holes in the middle of the stands. After the blunt needles were positioned in the guide tubes and stands, the needles were struck towards the subject's skin surface and exerted pressure on it. In order to maintain the sensation of pressure and fixate the needles in position, needle-holding tags were inserted into the guide tubes at the upper end. Extra pressure was exerted once every 5 minutes on their respective locations by gently striking the needle ends so as to stimulate the skin surface. Each skin surface stimulation session lasted for 30 minutes.

Subjects in the treatment group received stimulation on voice-related acupoints: two *Hegu* (Li4) and two *Lieque* (Lu7) points on the wrist, one *Lianquan* (CV23) and two *Renying* (St9) points on the neck and two *Zhaohai* (Ki6) points on the ankle. Subjects in the placebo group received stimulation on non voice-related acupoints: two *Tianchuang* (SI16) points on the neck, two *Neiguan* (PC6) points near the wrist, two *Zhongzhu* (SJ3) points on the hand, and two *Shangqiu* (SP5) points on the mellelous. All the acupoints were chosen according to their indications listed in Shi (2007), which is a Chinese national planned textbook for the study of acupuncture. The non-voice related acupoints used in the placebo group were ensured not to belong to the same meridians as the acupoints used in the treatment group. This was to prevent any unwanted stimulation of the voice-related acupoints.

Data Analysis

Aerodynamic analysis

Five aerodynamic measurements were obtained for each subject from four speech tasks:

- Maximum phonation task: the trial with the longest time of sustained phonation and the mean airflow rate corresponding to that trial were selected out of the five trials each subject attempted.
- Most comfortable phonation task: the five mean airflow rates were averaged and collected for each subject.
- Vowel-consonant string production task: the mean peak intraoral pressure was obtained by averaging the peak intraoral pressure of the middle five /pi/ syllables in five trials.
- Sentence production task: the mean peak intraoral pressure was obtained by averaging the peak intraoral pressure of the second /ba/ in five trials.

A Mann-Whitney *U*-test was carried out on each of the five aerodynamic measures to compare the between-subject variable (group effect; treatment versus placebo) in the pre-treatment assessment. Two Friedman tests were used independently to compare the within-subject variable (session effect; PRE versus POST versus FU) for the treatment group and the placebo group. Post-hoc Wilcoxon signed ranks test were used for pair-wise comparisons if there were significant differences between the measures obtained in the three assessments. To determine the intra-rater reliability in the selection of the segment of the recording for analysis, 10% of the sample (two sets of data) were analyzed the second time on a separate occasion. To determine the inter-rater reliability, another rater was given the same two sets of data for segment selection.

Voice range profile analysis

Seven measurements from the voice range profile were obtained for each subject, including the highest, lowest frequency and the frequency range; the highest, lowest intensity and the intensity range; and the voice range profile area (in semitone dBA). The frequency

and intensity measures required manual calculation, whereas the profile area was calculated by the *Phog* 2.0 system automatically.

A Mann-Whitney *U*-test was carried out on each of the seven measurements to compare between the two groups (treatment versus placebo) in the pre-treatment assessment session. Two Friedman tests and post-hoc Wilcoxon signed ranks test were used independently to compare the effects of time (PRE versus POST versus FU) for the treatment group and the placebo group. Intra-rater reliability in the calculation of the frequency and intensity measures was determined by having 10% of the samples (two sets of data) analyzed the second time on a separate occasion. To determine the inter-rater reliability, another rater was given the same 2 sets of data for calculation.

VAPP analysis

The Self-perceived Severity Score, Activity Limitation Score, Participation Restriction Score, Emotion Score and the Total VAPP Score were obtained for each subject. A Mann-Whitney *U*-test was carried out on each of the five scores to compare between the two groups (treatment versus placebo) in the pre-treatment assessment session. Two Friedman tests and post-hoc Wilcoxon signed ranks test were independently used to compare the effects of time (PRE versus POST versus FU) for the treatment group and the placebo group.

Results

Reliability

Inter-rater and intra-rater reliability were calculated for the aerodynamic measures and the voice range profile. Out of the 20 sets aerodynamic measures used for determining inter- and intra-rater reliability, 13 showed exact agreement. The remaining seven sets had less than 1% difference with one another and demonstrated good correlation using Spearman's rank correlation coefficient ($\rho = 1.000$, $p < .001$). All 28 sets of inter- and intra-rater measures obtained from the voice range profile showed exact agreement. The

results were proven to be reliable.

Aerodynamic measures

The mean aerodynamic values for the treatment and placebo groups across three assessment time points are shown on Table 1. A Mann-Whitney *U*-test (Table 2) comparing the values at pre-treatment between the two groups showed no significance for all measures ($p > .05$), indicating that the two groups did not differ statistically before the therapy. No significant within-group effects over time ($p > .05$), as revealed by the Friedman tests, were found in both groups (Table 3). Despite not supported by inferential statistics, the mean peak intraoral pressure detected in /pi/ string production in both groups appear to increase during treatment. While the values in the treatment group continued to increase after treatment ended, the level was maintained in the placebo group.

Voice range profile measures

The mean values of measurements for both groups obtained from the voice range profile are listed in Table 4. No significant differences were found between the groups at pre-treatment ($p > .05$) by a Mann-Whitney *U*-test (Table 5). The effect of time on both the treatment and placebo groups were found to be not significant as well ($p > .05$) as revealed in two Friedman tests (Table 6). Even though there is no statistical difference, a consistent trend indicating a decrease in highest fundamental frequency, frequency range and area of the voice range profile can be seen in the placebo group from pre-treatment to post-treatment. The values also failed to return to pre-treatment level when assessed in the follow-up session. Most measures obtained from subjects in the treatment group, on the other hand, remained steady throughout the course of treatment.

Table 1
Mean and standard deviation of aerodynamic measures across time

	Pre-treatment		Post-treatment		Follow-up	
Measurements/ Groups	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Maximum sustained phonation (s)						
Treatment	17.26	(9.68)	18.57	(2.70)	17.96	(2.85)
Placebo	14.64	(6.45)	16.10	(4.48)	16.32	(6.33)
Mean airflow rate— maximum sustained phonation (l/s)						
Treatment	0.20	(0.18)	0.11	(0.10)	0.15	(0.08)
Placebo	0.17	(0.01)	0.12	(0.09)	0.10	(0.10)
Mean airflow rate— most comfortable phonation (l/s)						
Treatment	0.22	(0.18)	0.18	(0.16)	0.29	(0.19)
Placebo	0.15	(0.07)	0.16	(0.09)	0.12	(0.06)
Mean peak intraoral pressure— consonant-vowel string (cm H_2O)						
Treatment	8.78	(3.79)	10.17	(1.72)	11.64	(0.55)
Placebo	13.35	(1.88)	14.64	(5.94)	14.88	(6.99)
Mean peak intraoral pressure— sentence production (cm H_2O)						
Treatment	10.08	(1.88)	10.16	(1.23)	10.49	(0.93)
Placebo	13.47	(2.01)	11.96	(4.15)	13.76	(6.95)

VAPP

The mean scores of the VAPP for both the treatment and placebo groups across time are listed in Table 7. No significant differences in mean scores ($p > .05$) were found between the groups at pre-treatment (Table 8); and the within-group effects of time were also not significant ($p > .05$) (Table 9). Despite not gaining statistical support, it was clear that the numerical values of all the VAPP scores have reduced unanimously in both groups after treatment. Nevertheless, most of them failed to maintain until follow-up.

Table 2
Mann-Whitney U between-group analysis of aerodynamic measures in pretreatment

Measurements	***U***	***Z***	***p***
Maximum phonation time (s)	1.000	-0.775	.439
Mean airflow rate— maximum sustained phonation (ml/s)	2.000	0.000	1.000
Mean airflow rate— most comfortable phonation (ml/s)	2.000	0.000	1.000
Mean peak intraoral pressure— consonant-vowel string (cm H_2O)	0.000	-1.549	.121
Mean peak intraoral pressure— sentence production (cm H_2O)	0.000	-1.549	.121

Table 3
Friedman within-group analysis of aerodynamic measures across time

Measurements/ Groups	χ^2	***df***	***p***
Maximum phonation time (s)			
Treatment	1.000	2	.607
Placebo	3.000	2	.223
Mean airflow rate— maximum sustained phonation (ml/s)			
Treatment	3.000	2	.223
Placebo	3.000	2	.223
Mean airflow rate— most comfortable phonation (ml/s)			
Treatment	4.000	2	.135
Placebo	3.000	2	.223
Mean peak intraoral pressure— consonant-vowel string (cm H_2O)			
Treatment	0.000	2	1.000
Placebo	0.000	2	1.000
Mean peak intraoral pressure— sentence production (cm H_2O)			
Treatment	0.000	2	1.000
Placebo	0.268	2	.867

Table 4
Mean and standard deviation of voice range profile measures across time

	Pre-treatment		**Post-treatment**		**Follow-up**	
Measurements/ Groups	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Highest frequency (Hz)						
Treatment	893.25	(216.73)	863.90	(175.22)	893.25	(216.73)
Placebo	647.65	(328.59)	446.65	(108.40)	466.20	(0.00)
Lowest frequency (Hz)						
Treatment	110.15	(8.98)	110.15	(8.98)	106.90	(4.38)
Placebo	155.80	(12.73)	134.70	(5.52)	147.80	(24.04)
Frequency range (Hz)						
Treatment	783.10	(207.75)	753.75	(166.24)	786.35	(212.34)
Placebo	490.95	(342.59)	311.95	(102.88)	318.40	(24.04)
Highest Intensity (dB)						
Treatment	97.50	(6.36)	94.50	(3.54)	97.00	(5.66)
Placebo	92.00	(7.07)	92.50	(4.95)	95.00	(1.41)
Lowest Intensity (dB)						
Treatment	56.00	(0.00)	57.50	(3.54)	55.50	(3.54)
Placebo	58.00	(5.66)	58.00	(2.83)	60.50	(6.36)
Intensity Range (dB)						
Treatment	41.50	(6.36)	37.00	(7.07)	41.50	(2.12)
Placebo	34.00	(1.41)	34.50	(7.78)	34.50	(4.95)
Area (in semitone x dB)						
Treatment	897.50	(3.54)	836.50	(44.55)	904.00	(60.81)
Placebo	472.50	(228.40)	410.00	(130.11)	416.00	(52.33)

Table 5
Mann-Whitney U between-group analysis of voice range profile measures in pretreatment

Measurements	***U***	***Z***	***p***
Highest frequency (Hz)	1.000	-0.775	.439
Lowest frequency (Hz)	0.000	-1.549	.121
Frequency range (Hz)	1.000	-0.775	.439
Highest Intensity (dB)	1.000	-0.775	.439
Lowest Intensity (dB)	2.000	0.000	1.000
Intensity Range (dB)	0.000	-1.549	.121
Area (in semitone x dB)	0.000	-1.549	.121

Table 6
Friedman within-group analysis of voice range profile measures across time

Measurements/ Groups	χ^2	***df***	***p***
Highest frequency (Hz)			
Treatment	2.000	2	.368
Placebo	1.000	2	.607
Lowest frequency (Hz)			
Treatment	2.000	2	.368
Placebo	2.000	2	.368
Frequency range (Hz)			
Treatment	2.000	2	.368
Placebo	1.000	2	.607
Highest Intensity (dB)			
Treatment	3.714	2	.156
Placebo	0.000	2	1.000
Lowest Intensity (dB)			
Treatment	0.000	2	1.000
Placebo	0.268	2	.867
Intensity Range (dB)			
Treatment	3.000	2	.223
Placebo	0.000	2	1.000
Area (in semitone x dB)			
Treatment	3.000	2	.223
Placebo	0.000	2	1.000

Table 7
Mean and standard deviation of VAPP scores across time

	Pre-treatment		**Post-treatment**		**Follow-up**	
Scores/ Groups	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Self-Perceived Severity Score						
Treatment	7.00	(2.83)	5.50	(0.71)	5.50	(2.12)
Placebo	7.50	(0.71)	6.50	(0.71)	6.50	(0.71)
Activity Limitation Score						
Treatment	61.50	(37.48)	53.00	(25.46)	55.00	(28.28)
Placebo	63.00	(28.28)	48.00	(25.46)	46.00	(22.62)
Participation Restriction Score						
Treatment	78.00	(8.49)	61.50	(6.36)	66.50	(17.68)
Placebo	47.00	(42.43)	33.00	(25.46)	35.00	(29.70)
Emotion Score						
Treatment	51.00	(16.97)	39.50	(12.02)	46.00	(15.56)
Placebo	42.50	(19.10)	37.00	(22.63)	36.00	(19.80)
Total VAPP Score						
Treatment	197.50	(65.76)	159.50	(44.55)	173.00	(63.64)
Placebo	160.00	(90.51)	124.50	(74.25)	123.50	(72.83)

Table 8
Mann-Whitney U between-group analysis of VAPP scores in pretreatment

Measurements	***U***	***Z***	***p***
Self-Perceived Severity Score	2.000	0.000	1.000
Activity Limitation Score	2.000	0.000	1.000
Participation Restriction Score	1.000	-0.775	.439
Emotion Score	1.000	-0.775	.439
Total VAPP Score	1.000	-0.775	.439

Table 9
Friedman within-group analysis of VAPP scores across time

Measurements/ Groups	χ^2	***df***	***p***
Self-Perceived Severity Score			
Treatment	2.000	2	.368
Placebo	4.000	2	.135
Activity Limitation Score			
Treatment	2.000	2	.368
Placebo	3.714	2	.156
Participation Restriction Score			
Treatment	3.000	2	.223
Placebo	3.000	2	.223
Emotion Score			
Treatment	4.000	2	.135
Placebo	3.000	2	.223
Total VAPP Score			
Treatment	3.714	2	.156
Placebo	3.714	2	.156

Discussion

The study set out to investigate the effects of skin surface stimulation therapy on dysphonic subjects having phonotraumatic lesions. Using a protocol that employed the same set of voice-related acupoints and the same course of treatment, it was hypothesized that skin surface stimulation, presented non-deceptively, could bring about positive changes to the vocal functions and voice-related quality of life just as presented deceptively to subjects in the placebo group in the study by Yiu et al. (in preparation). Aerodynamic measures and the voice range profile were used to evaluate changes in vocal functions. The subjects' self-perceived voice-related quality of life was measured using the VAPP.

Aerodynamic measures

Aerodynamic measures were chosen as outcome measures for being sensitive to treatment outcomes and reflective of the impact of a lesion on the vocal functions (Holmberg,

Doyle, Perkell, Hammarberg, & Hillman, 2002). The presence of a greater mean airflow rate in speech tasks is often associated with an inadequate glottal closure due to obstruction by the protruding vocal fold lesion. According to Smitheran and Hixon (1981), the peak intraoral pressure was found to be reliable and valid in estimating the sub-glottal pressure. In order to overcome the difficulty of achieving adequate closure of the vocal folds during phonation, the phonatory threshold pressure or the sub-glottal pressure has to be increased (Colton, Casper, & Leonard, 2006). Maximum phonation time is considered an indicator of vocal functions, and also commonly used in the evaluation of therapy outcomes (Hirano, Koike, & Von Leden, 1968). A reduction in mean airflow rate and mean peak intraoral pressure, and an increase in maximum phonation time will thus support a reduction of vocal fold lesion size and an improvement of vocal functions.

The treatment group and the placebo group were comparable before therapy as no significant differences between them were found at pre-treatment level. Thus, further within-subject comparisons conducted independently for the two groups would be valid and meaningful. However, no statistically significant changes were found across time for the treatment and placebo groups in all the measures. The peak intraoral pressure during consonant-vowel string production seemed to increase from pre- to post-treatment in both groups. A possible interpretation is that skin surface stimulation or the voice-related acupoints selected were not effective in reducing the vocal lesion size and improving glottal closure during phonation. Since voicing behaviors were not controlled in this study, greater phonotraumatic voice use after therapy ended for the treatment group may account for the further increase of mean peak intraoral pressure at 14 days post-treatment.

Voice range profile measures

The voice range profile is a powerful tool widely used for documenting changes in vocal functions as to measure voice therapy outcomes (Heylen et al., 1998). The treatment group

and the placebo group were determined not to be significantly different before treatment. Significant changes over time were not found for the two groups either. Nevertheless, similar trends have been observed for the measures of highest frequency, frequency range and the profile area. The presence of vocal pathologies increases the mass of the vocal folds, and thus reduces their rate of vibration. The increased stiffness also restricts the degree of vocal fold stretching and thus reduces the highest possible frequency that can be attained during phonation (Colton, Casper, & Leonard, 2006). The frequency range is the difference between the highest and lowest frequency. It is a quantifiable measure of the flexibility to increase or decrease the mass per unit length of the vocal folds to vary the vibration frequency. The voice range profile area is a two-dimensional measure that demonstrates the interaction between vocal frequency and intensity ranges (Heylen et al., 1998). Therefore, a larger area is often associated with better vocal functions. While these measures remained relatively steady over the assessments in the treatment group, a drop seemed to appear after treatment for the placebo group, and there was little change in the values afterwards during the follow-up assessment.

Based on the trends observed, there are currently two possible hypotheses that might account for the phenomena, the first one being skin surface stimulation on non-voice acupoints were detrimental to vocal functions, whereas therapy using voice-related acupoints did not yield any changes. The second hypothesis was that the placebo group demonstrated the natural course of development for the vocal pathologies; whereas the treatment group, in fact, benefitted from therapy so that that the vocal functions maintained from pre-treatment to 14 days post-treatment. A no-treatment control group that shows the natural development of the vocal functions would be necessary in order to draw a conclusion.

VAPP

According to Traditional Chinese Medicine theories, acupuncture or skin surface

stimulation is an integrative care that attempt to induce holistic improvements to patients (Paterson & Schnyer, 2007). Quality of life has often been emphasized as an important aspect to patient-centred outcome measures (Enderby, 1992). A clinically significant change, or the effect on the patient's functional abilities, is considered to have equal status as statistically significant improvements for dysphonic patients in evaluating treatment efficacy (Carding, 2000).

In this study, all the VAPP Scores for the treatment and placebo groups did not differ significantly at pre-treatment; and the scores in both groups did not differ significantly across time as well. A negative change of all scores from pre- to post-treatment was apparent in both groups, showing positive changes in the voice-related quality of life during that period. Interestingly, an increase in Activity Limitation, Participation Restriction and the Emotion Scores after post-treatment was noticed. This indicated a reduction of well-being in quality of life when the therapy stopped. Even though a negative trend in some parameters of vocal functions was observed, the self-reported voice-related quality of life bettered during the treatment in the placebo group. These observations were in line with our hypothesis that the improvements in voice-related quality of life will be similar for both groups, and independent of changes in vocal functions.

Limitations of the current study

Firstly, the limitation in sample size is suspected to be the reason that statistically significant results were not obtained in any of the measures used in the study. Discussion was confined to the interpretation of consistent trends observed. The large standard deviations, coming from a small sample, might have masked possible statistical significance as well. Moreover, individual differences such as voicing behaviors were not controlled and unlikely to be balanced after random assignment of subjects in a small sample. Therefore, the study should be repeated with an increased sample size. The second limitation is the small

number of outcome measures being used. Improvement in voice captured in one outcome measure may not necessarily be reflected in the other. Whilst the current study focuses on vocal functions and quality of life measures, laryngoscopic examination and perceptual voice evaluation should be incorporated in future studies so that therapy outcomes can be evaluated in a multi-dimensional manner (Speyer, Wieneke, & Dejonckere, 2004).

Conclusion

This study represents a pilot trial to investigate the effects of skin surface stimulation on acupoints on improving vocal functions and voice-related quality of life. Outcome measures employed including aerodynamic measures, the voice range profile, and the Voice Activity and Participation Profile (VAPP). Due to the limitation of sample size, no statistically significant findings were obtained in any of the measures. Despite that, trends in mean peak intraoral pressure; highest frequency, frequency range and voice range profile area; as well as VAPP scores across the course of treatment were observed. In conclusion, current results remain inconclusive as in whether skin surface stimulation on acupoints is an effective treatment for phonotraumatic lesions.

References

Aronson, A. E., & Bless, D. M. (2009). *Clinical voice disorders* (4th ed.). New York: Thieme.

Birch, S., & Felt, R. L. (1999). *Understanding acupuncture*. Edinburgh: Churchill Livingstone.

Branski, R. C., Rosen, C. A., Verdolini, K., & Hebda, P. A. (2004). Markers of wound healing in vocal fold secretions from patients with laryngeal pathology. *The Annals of Otology, Rhinology & Laryngology, 113*(1), 23-29.

Branski, R. C., Verdolini, K., Sandulache, V., Rosen, C. A., & Hebda, P. A. (2006). Vocal fold wound healing: A review for clinicians. *Journal of Voice, 20*(3), 432-442.

Carding, P. (2000). *Evaluating voice therapy: Measuring the effectiveness of treatment*. London: Whurr.

Colton, R. H., Casper, J. K., & Leonard, R. (2006). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (3rd ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Du, Y., Xiong, J., Huang, W., Li, B., & Shi, L. (2009). Medical conditions treated by acupuncture: A preliminary review of randomized control trials. *Medical Acupuncture, 21*(3), 207-213.

Enderby, P. (*1992*) *Outcome* measures in speech *therapy*: Impairment, disability, handicap, and distress. *Health Trends, 24*(2), 61–64.

Ernst, E. (2006). Acupuncture—a critical analysis. *Journal of Internal Medicine, 259*, 125-137.

Heylen, L., Wuyts, F. L., Mertens, F., De Bodt, M., Pattyn, J., Croux, C., & Van de Heyning, P.H. (1998). Evaluation of the vocal performance of children using a voice range profile index. *Journal of Speech, Language and Hearing Research, 41*, 232-238.

Hirano, M., Koike, Y., & Von Leden, H. (1968). Maximum phonation time and air usage during phonation. *Folia Phoniatrica, 20,* 185-201.

Holmberg, E. B., Doyle, P., Perkell, J. S., Hammarberg, B., & Hillman, R. E. (2002). Aerodynamic and acoustic voice measurements of patients with vocal nodules: variation in baseline and changes across voice therapy. *Journal of Voice, 17*(3), 269-282.

Jasper, J, K., & Colton, R. (2000). Current understanding and treatment of phonatory disorders in geriatric populations. *Current Opinion in Otolaryngology & Head and Neck Surgery, 8*, 158-164.

Kaptchuk, T. J., Stason, W. B., Davis, R.B., Legedza, A. T. R., Schnyer, R. N., Kerr, C. E., . . . Goldman, R. H. (2006). Sham device *v* inert pill: randomized controlled trial of two placebo treatments. *BMJ, 332*(7538), 391–397. doi: 10.1136/bmj.38726.603310.55

Langevin, H. M., Wayne, P.M., MacPherson, H., Schnyer, R., Milley, R. M., Napadow, V., . . . Hammerschlag, R. (2011). Paradoxes in acupuncture research: strategies for moving forward. *Evidence-based Complementary and Alternative Medicine, 2011*, 11 pages. doi:10.1155/2011/180805

Lee, J., Jeong, H. J., Park, H., Jeon, S., & Hong, S. (2010). Acupuncture accelerates wound healing in burn-injured mice. *Burns, 37,* 117-125.

Leonard, R. (2009). Voice therapy and vocal nodules in adults. *Current Opinion in Otolaryngology & Head and Neck Surgery, 17*, 453-457.

Ma, E. P.-M. (2011). Voice range profile: Phog. In E. P.-M. Ma & E. M.-L. Yiu (Eds.), *Handbook of voice assessments*. San Diego: Plural Publishing.

Ma, E. P.-M., & Yiu, E. M.-L. (2001). Voice activity and participation profile: Assessing the impact of voice disorders on daily activities. *Journal of Speech Language and Hearing Research, 44,* 511-524.

Ma. E. P.-M., & Yiu, E. M.-L. (2006). Multiparametric evaluation of dysphonic severity. *Journal of Voice, 20*(3), 380-390.

Ma. E. P.-M., & Yiu, E. M.-L. (2007). Scaling voice activity limitation and participation restriction in dysphonic individuals. *Folia Phoniatrica et Logopaedica, 59*(2), 74-82.

Moffet, H. H. (2008). Traditional acupuncture theories yield null outcomes: a systematic review of clinical trials. *Journal of Clinical Epidemiology, 61*, 741-747.

Nakagawa, H., Miyamoto, M., Kusuyama, T., Mori, Y., & Fukuda, H. (2012). Resolution of vocal fold polyps with conservative treatment. *Journal of Voice, 26*(3), e107-e110.

Park, S. I., Sunwoo, Y., Jung, Y. J., Chang, W C., Park, M, Chung, Y., . . . Lee, S. (2012). Therapeutic effects of acupuncture through enhancement of functional angiogenesis and granulogenesis in rat wound healing. *Evidence-based* C*omplementary and Alternative Medicine, 2012,* 10 pages. doi:10.1155/2012/464586

Paterson, C., & Schnyer, R.N. (2007). Measuring patient-centred outcomes. In H. MacPherson (Eds.), *Acupuncture research: strategies for establishing an evidence base*. Edinburgh: Churchill Livingstone Elsevier.

Ramig, L. O., & Verdolini, K. (1998). Treatment efficacy: voice disorders. *Journal of Speech, Language and Hearing Research, 41*, S101-S116.

Shang, C. (2001). Electrophysiology of growth control and acupuncture. *Life Sciences, 68*, 1333-1352.

Shang, C. (2009). Has basic acupuncture research met the gold standard of science? *Medical Acupuncture, 21*(3), 191-200.

Shi, X. (2007). *Zhen jiu xue*. Beijing: Gao deng jiao yu chu ban she.

Smitheran, J. R., & Hixon, T. J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders, 46*, 138-146.

Speyer, R., Wieneke, G. H., & Dejonckere, P. H. (2004). Documentation of progress in voice therapy: perceptual, acoustic, and laryngostroboscopic findings pretherapy and posttherapy. *Journal of Voice, 18*(3), 325-340.

Sun, J.-H., Zhang, Y., Yan, D.-N., Ma, H.-A., Jiang, Z.-Q., & Geng, X.-W., (2010). Effects on function rehabilitation of vocal cord after vocal cord polyps surgery treated with acupuncture at Sheng's Four Points of Throat. *Chinese Acupuncture & Moxibustion, 30*(10), 822-825.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N.J: Prentice Hall.

Wang, Q., & Tao, B. (2005). Acupuncture treatment of vocal nodule and evaluation of the therapeutic effect by computerized acoustics assay. *Chinese Acupuncture & Moxibustion, 25*(6), 404-406.

Woo, P., Casper, J., Colton, R. & Brewer, D. (1994). Diagnosis and treatment of persistent dysphonia after laryngeal surgery: a retrospective analysis of 62 patients. *Laryngoscope, 104*, 1084-1091.

World Health Organization. (2002). *Acupuncture: review and analysis of reports on controlled clinical trials.* Geneva, Switzerland: WHO.

Yang, S.-R., Xie, Q., & Jiang, J.-J. (2006). Clinical study and acoustical analysis for submucosa hemorrhage of vocal cords treated mainly with acupuncture at "Kaiyin point No. 1". *Chinese Acupuncture & Moxibustion, 26*(4), 243-246.

Yiu, E. M. L., Kwong, E., Chan, K., Tsang, R., Li, N., Verdolini, K., . . . Lin, J. (in preparation). Laryngeal wound healing in acupuncture and skin surface stimulation.

Yiu, E., Xu, J. J., Murry, T., Wei, W. I., Yu, M., Ma, E., . . . Kwong, E. Y.-L. (2006). A randomized treatment-placebo study of the effectiveness of acupuncture for benign vocal pathologies. *Journal of Voice, 20*(1), 144-156.

Acknowledgement

I would like to express my deepest gratitude to Professor Edwin Yiu, my dissertation supervisor, for his guidance, encouragement and useful critiques to this work. I would like to thank Ms. Yung Tsang for conducting the skin surface stimulation sessions and her assistance in administrative duties; Ms. Miki Shek for her insights in Traditional Chinese Medicine and her help with recording voice range profiles; and Miss Yann Lam for the inter-rater reliability ratings. My grateful thanks are also extended to the staff of the Voice Research Laboratory for their support and assistance throughout the study.

Appendix

Figures and illustrations of skin surface stimulation

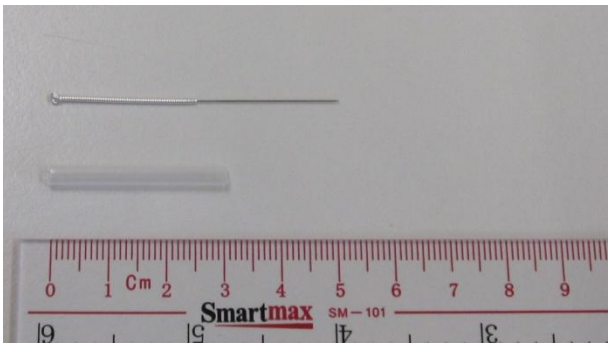

Figure 1. The needle and guide tube

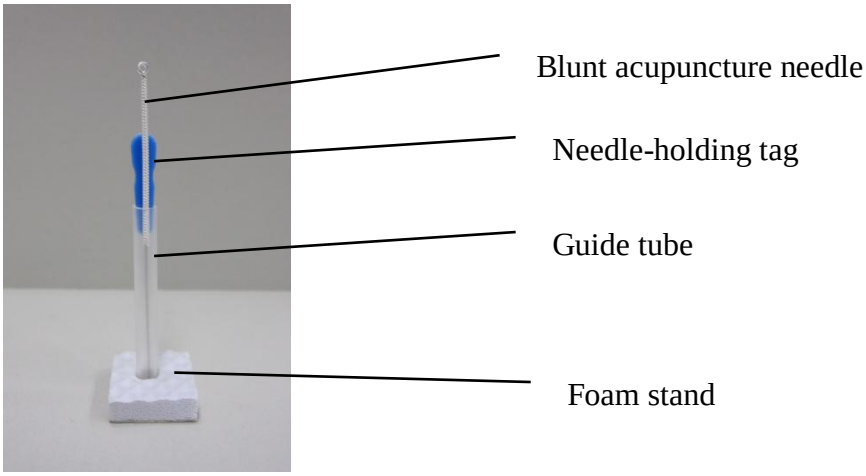

Figure 2. Skin surface stimulation setup with stand, needle and needle-holding tag

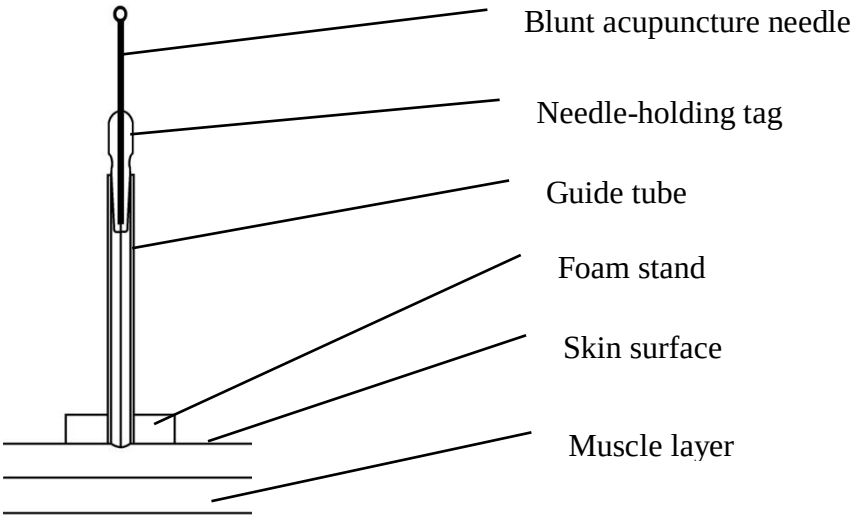

Figure 3. Illustration of skin surface stimulation

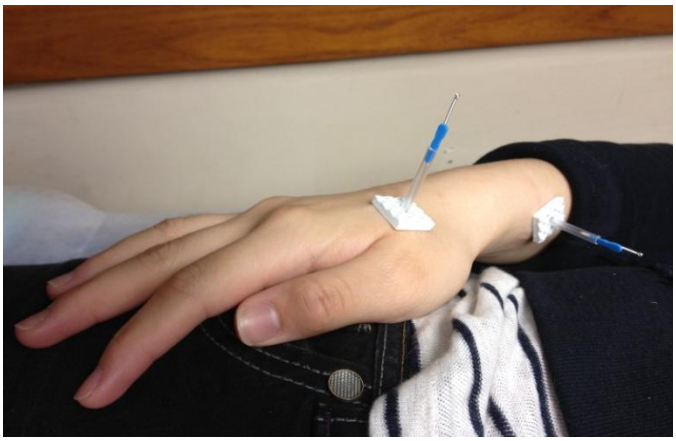

Figure 4. Skin surface stimulation setup placed *in-situ* at *Hegu* (Li4) and *Lieque* (Lu7)

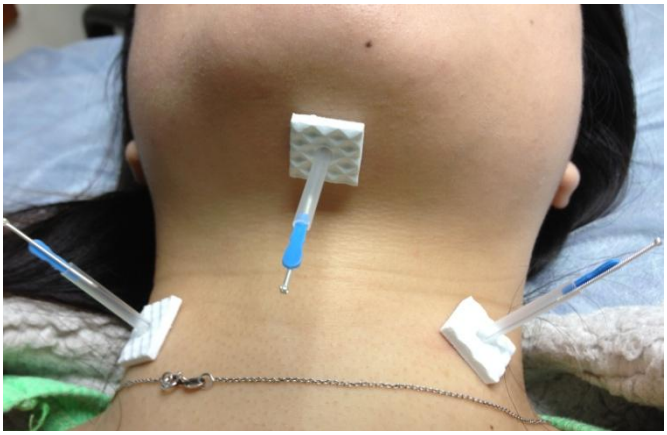

Figure 5. Skin surface stimulation setup placed *in-situ* at *Lianquan* (CV23) and *Renying* (St9)

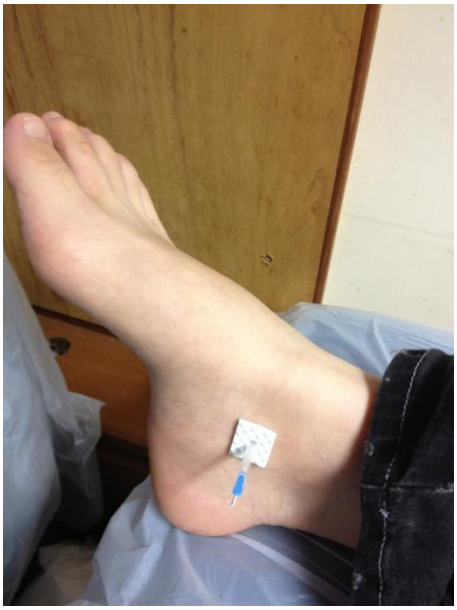

Figure 6. Skin surface stimulation setup placed *in-situ* at *Zhaohai* (Ki6)