Development and Validation of the Chinese Version of the Massachusetts General Hospital Acupuncture Sensation Scale: an exploratory and methodological study

David TW Yu[1], Alice YM Jones[2]*, Marco YC Pang[2]

[1]David Tai Wai YU, Physiotherapy Department, Queen Elizabeth Hospital, Hong Kong, China

[2]Alice Yee Man JONES, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China.

[2]Marco Yiu Chung PANG, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China.

E-mail: David TW Yu – yutwd@ha.org.hk; Alice YM Jones* - alice.jones.hk@gmail.com; Marco YC Pang - Marco.Pang@inet.polyu.edu.hk.

*Corresponding author

Correspondence to Professor Alice Jones, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. Fax: 852-23308656, Voice: 852-27666749, Email: alice.jones.hk@gmail.com

Abstract

Background: The Massachusetts General Hospital Acupuncture Sensation Scale (MASS) is a tool to measure needle sensations. The aims of the present study were to develop a Chinese version and to assess its psychometric properties.

Methods: This study was a methodological and exploratory study. The English version of the MASS was translated into Chinese using standardized translation procedures. Content validity was conducted by 9 acupuncture experts. The prefinal Chinese version (C-MASS) was then administered to 30 acupuncture naïve healthy subjects. Electroacupuncture was performed on right LI4 and LI11 for 30 min. Test-retest reliability measurement was administered again 1-2 weeks later. Construct validity was examined by comparing C-MASS and the Short-form McGill Pain Questionnaire (SF-MPQ). The construct validity was further assessed by the principle component analysis.

Results: C-MASS demonstrated a content validity ratio on relevance and importance from -0.04 to 1.00. Convergent validity was demonstrated by its significant association with the sensory dimension of SF-MPQ (γ=0.63, p<0.05). Discriminant validity was demonstrated by its low association with the affective dimension of SF-MPQ (γ=-0.3, p=0.111). A five-factor structure of C-MASS was established by factor analysis. C-MASS demonstrated good internal consistency (Cronbach's α = 0.71) and test-retest reliability (ICC= 0.92). Since the descriptors "sharp pain" was not a valid needle sensation related to *deqi*, this was removed from C-MASS. We renamed as the Modified MASS-Chinese version (C-MMASS).

Conclusion: A 12 descriptors C-MMASS was established and shown to be a reliable

and valid tool in reporting needle sensations associated with *deqi* among healthy young Chinese people.

INTRODUCTION

Acupuncture has been widely used in China for thousands of years.[1-3] One of the fundamental characteristics of acupuncture is to "obtain *qi*" during acupuncture, a sensation referred to as "*deqi*". Based on the concept of Traditional Chinese Medicine (TCM), *qi* must flow in correct strength and quality along the meridians so that health can be maintained. It was believed that restoration of health can only be achieved if the acupuncture technique is able to elicit *deqi*, thereby allowing the flow of *qi* to be altered.[3-4] While investigations of the relationship between therapeutic effectiveness of acupuncture and deqi experience was reported in literature, [3,5-7] scientific evidence to support such relationship is still lacking.[7]

When *deqi* occurs, both the acupuncturist and the subject may experience some "unusual" sensations around the needle. The acupuncturist may perceive heaviness or tenseness around the needle when *qi* arrives.[8] However, the feelings of the acupuncturist are more often subjected to biased preconceptions of "what one ought to feel" and thus hold greater likelihood of a biased subjective report.[9] In recent years, researchers focused more on sensations perceived by the subjects. Standard needle sensations associated with *deqi* experienced by the subjects are described as soreness, distension, heaviness and numbness.[10-12] Nevertheless, there is still no standard method in measurement and quantifying *deqi* sensations at present.

Over the last two decades, a number of studies have attempted to quantify *deqi* sensations.[10-12] Vincent and colleagues adopted the McGill Pain Questionnaire and created a new of 20 adjectives describing possible acupuncture sensations.[10] In 2005, Park et al modified Vincent and colleagues' work by adding five extra descriptions of the sensation scale.[11] However, these scales have limitations because both of these instruments measure a range of sensations that include pain as well as *deqi*. In 2006, a panel of 29 international acupuncture experts was involved in MacPherson and Asghar's study who categorized 25 sensations during acupuncture as associated with *deqi* and acute pain at the needling site.[12] However, the above study quantified *deqi* sensations based on the acupuncturists rather than the patient/subject's perception.

To address the complexity involved in accurately assessing the acupuncture sensations, Kong and coworkers developed the "Subjective Acupuncture Sensation Scale" (SASS), an inventory incorporating various sensations associated with *deqi*.[13] The SASS listed 9 descriptors of sensations associated with *deqi* sensations reported in traditional literature.[13] Each of the elements was presented on a 10-cm bar with the anchor words "none", "mild", "moderate", and "severe" spaced evenly on the continuum. Subjects were asked to quantity their sensations at each point by rating the intensity with which they experienced after acupuncture. The authors demonstrated

that there were significant correlations between numbness and soreness items of the SASS.[13] In order to apply SASS to a wider range of research projects, Kong and his coworkers further expanded the descriptors in the SASS, and this became the Massachusetts General Hospital Acupuncture Sensation Scale (MASS) (Appendix 1).[7] The descriptors in MASS include soreness, aching deep pressure, heaviness, fullness/distension, tingling, numbness, sharp pain, dull pain, warmth, cold, throbbing, plus one blank supplementary row left for subjects to describe perceptions in their own words. The MASS Index is then calculated and used to quantify the intensity of the needle sensations experienced by a subject. The MASS also included two supplementary scales for measurement of spreading of the acupuncture sensation and anxiety.[7]

Establishment of the MASS involved extensive review of relevant literature and appears to be the most comprehensive assessment tool in measuring the needle sensations. MASS has subsequently been used for measurement of acupuncture sensations.[14-15]. The validity and reliability of the MASS, however, have not been properly established. Furthermore, precise Chinese language terms that best describe the needle sensation have not been determined. If optimal needle sensation is essential for achieving positive acupuncture effects, it is necessary to have a valid and reliable instrument to be used in the clinical setting to systematically quantify and document

the needle sensation associated with *deqi*, as this could be an important factor which determines the treatment effect of acupuncture intervention. The objectives of the study therefore were to establish a Chinese version of the MASS and to test its psychometric properties.

METHODS

The study was an exploratory, methodological study that involved two phases. Ethics approval was granted by the Human Subjects Ethics Committee of the university involved. The nature of the study was explained to the subjects and written consent obtained prior to data collection. All procedures were conducted in accordance with the Declaration of Helsinki.

Phase I: Cultural adaptation of the Massachusetts General Hospital Acupuncture Sensation Scale and Assessment of Content Validity

Permission was obtained from the original authors of the MASS before the initiation of the study. The translation process was divided into 5 stages and was mainly based on the guidelines that described by Beaton et al [16] and Wild et al.[17]

The first stage was the forward translation of the MASS into two independent Chinese versions by a physiotherapist and a professionally trained translator whose

mother tongue is Chinese. The professionally trained translator has no medical or clinical background (naïve translator).

The second stage involved the synthesis of the results of the translations. The translated versions were compared and any ambiguous wordings were identified. Then, a single Chinese version of the MASS was first established.

The third stage was the back translation stage. It involved another two independent translators with physiotherapy and TCM backgrounds and they translated the Chinese version into English.

In the fourth stage, a panel of committee consisting of one professor from a local university, 7 physiotherapists and 2 nursing staff examined the preliminary versions in terms of the degree of agreement on the importance and relevance of each item on the Chinese version of the MASS to *deqi* sensation. All the committee members are bilingual individuals with Chinese as their mother tongue. Six of them had a Master level in acupuncture studies and 4 had a Bachelor degree in TCM. They had an average of 9 years of clinical experience in acupuncture. They were asked to rate each item on a 5-point scale from "strongly agree" to "strongly disagree". The Content Validity Ratio (CVR) was then computed. The CVR = (ne – N/2) / (N/2), where 'ne' refers to the number of subject matter that the experts considered as essential

measurement items, and N is the total number of experts in the panel. The CVR range

between –1.00 and +1.00, where a CVR of 0.00 means that 50% of the panel believes

the item to be essential. With an expert panel of 10 members, a 0.60 CVR was

required to meet a 0.05 significance level of importance.[18] The prefinal version was

then established based on the results on the degree of agreement on the importance

and relevance in content validity measurement.

In the fifth stage, the prefinal version was pilot tested on ten normal healthy

young subjects who were naïve to acupuncture and had no prior knowledge on the

original MASS. They received electroacupuncture on right Hegu (LI4) and Quichi

(LI11) (see paragraph in Phase II below) and they were asked to complete the prefinal

Chinese version of the MASS. Subjects were asked to select from this prefinal version

that best describe the sensations during electroacupuncture. In computing the MASS

index, the subject's individual MASS rating scales were ordered from highest to

lowest intensity, and the MASS index was calculated according to the following

equation:

$$\text{MASS Index} = \frac{\sum_{i=1}^{n} (1/2)^i R_i}{1 - (1/2)^n}$$

Where R indicates the ratings for different sensations from highest to lowest; n

represents the number of *deqi* sensations on the MASS questionnaire.[7]

Based on the feedback obtained from the subjects who participated in the pilot testing, minor modifications were made to the inventory to further improve its clarity. The final Chinese version, the C-MASS, was adopted and was used to examine the reliability and construct validity in this study.

Phase II: Validation of the C-MASS

Subjects' selection and sample size estimation

Subject inclusion criteria included people of under 40 years of age, right-handed, with normal health, naïve with acupuncture, and able to understand both Chinese and English, as the subjects were required to complete the Short-Form McGill Pain Questionnaire (English version). Subjects with known neurological, cardiovascular and psychological disorders were excluded. Convenience sampling method was adopted. They were invited to a face-to-face interview and demographic data were collected.

The software PASS 2008 was used for sample size calculation. Sample estimation was based on the intraclass correlation coefficient (ICC) obtained during the pilot test-retest reliability study. Preliminary results showed that the test-retest reliability was good with the ICC exceeding 0.95. Therefore, if the ICC was set at 0.95 at Type I error of 5% and power of 80%, 30 subjects were required.

Procedures for electroacupuncture administration

The acupuncture points Hegu (LI4) and Quichi (LI11) of the right arm were selected as appropriate points for stimulation. All subjects were in a sitting position with both arms supported by a pillow. After the acupoints were located, isopropyl alcohol was applied for disinfection. Then two 40mm × 0.25mm single-use, sterile, prepacked stainless needles with guide tubes were inserted into the L14 and LI11 points. The needles were inserted for about 0.5 inch and then manually manipulated until *deqi* sensation was reported. In this study, *deqi* was defined as sensations perceived by the subjects as "soreness / numbness / distension / heaviness" at the intensity around a moderate level. This was to ensure that proper acupoints were located and the subjects could identify types of sensations when *deqi* arrived. In order to standardize the stimulation throughout the stimulation period, electroacupuncture was used in the present study. Two electrodes were attached to the needles and connected to an electroacupuncture device (ITO, EX-160, Hannover, Germany). Electrical stimulation was then applied for 30 minutes at a frequency of 2Hz, pulse duration at 0.6 – 0.8 ms.[13] Intensity of the electrical stimulation was increased gradually to moderate level using the 10-point Visual analog scale.[19] Subjects were checked every 5 minutes and the intensity was readjusted so as to ascertain that *deqi* sensation was present.

After the electroacupuncture stimulation, the subjects were asked to quantify their acupuncture sensations by first completing the C-MASS and then the Short-Form McGill Pain Questionnaire (SF-MPQ).[20] Those who experienced *deqi* sensation at a moderate level during electroacupuncture in this session were invited back one to two weeks later for test-retest reliability measurement. This period was chosen because it was relatively long enough to minimize the chance of recalling the answers but short enough to minimize the probability of the subjects having undergone real change on the domain of interest.[21] Similar to the first assessment session, the subjects received electroacupuncture on right LI4 and LI11 in the second session and then required to quantify the *deqi* with the C-MASS. The MASS Index obtained at the first visit was compared with the second visit for the test-retest reliability assessment.

The Short-Form McGill Pain Questionnaire (SF-MPQ)

The SF-MPQ is designed to provide a brief measurement of pain.[20] It is a 2-factor inventory (sensory and affective) constructs for measurement of pain. The SF-MPQ consists of 15 descriptors (11 sensory and 4 affective) rated on an intensity scale as 0 = none, 1 = mild, 2 = moderate or 3 = severe.

Statistical Analysis

Descriptive statistics was performed to indicate the intensity of each needle sensation. Construct validity was examined by comparing the C-MASS and the SF-MPQ. It was hypothesized that items in C-MASS would be highly correlated with the sensory component but not with affective component of the SF-MPQ. Correlation matrices were created between the items of C-MASS and SF-MPQ. The data were compared by Pearson product-moment coefficient of correlation γ. A correlation coefficient above 0.50 to 0.75 indicates a moderate to good relationship, above 0.75 was considered as good to excellent relationship. The observed values of γ should be greater than or equal to the tabled value to achieve a significance level.[21]

The construct validity was further assessed by principle component analysis (PCA) of the C-MASS. A correlation matrix of all the items was created. The factor was extracted for factor analysis by two criteria: factors with eigenvalues greater than 1 and the Scree test criterion, obtained by plotting the eigenvalues against the number of factors. A factor loading of greater than 0.40 was considered significant. Orthogonal rotation using varimax rotation was used for the rotation of factors to improve the spatial structure of the variables so that distinct factors would be more visible.[21] Naming of factors was based on those items which had the highest factor loading in each factor.[21]

For reliability, the internal consistency of the C-MASS was assessed using Cronbach's coefficient α which was considered reliable if it approached 0.70.[21] To examine the test-retest reliability, ICC two-way random model were computed. ICCs above 0.75 were indicative of good reliability whereas those between 0.5 and 0.75 were considered moderate reliability.[21] The level of significance was set at 0.05 for all analyses. All of the statistical analyses were performed using SPSS 17.0 software (SPSS, In., Chicago, Illinois, USA).

RESULTS

Phase I: Assessment of content validity

The degree of agreement on the importance and relevance of each item associated with *deqi* sensation is shown in Table 1 and Table 2. It was found that the CVR of all items ranged from -0.4 to 1.0.

Table 1. Results of content validity of the prefinal Chinese version of the MASS ("Importance" content associated with *deqi* sensation) (n=11)

	Strongly agree (%)	Agree (%)	Total % of agree	Neutral (%)	Disagree (%)	Strongly disagree (%)	Total % of disagree	CVR
Soreness	70	30	100	0	0	0	0	1.0
Aching	20	60	80	0	20	0	20	0.6
Deep pressure	20	40	60	10	20	10	30	0.2
Heaviness	40	50	90	10	0	0	0	0.8
Fullness/distension	50	40	90	10	0	0	0	0.8
Tingling	40	40	80	10	10	0	10	0.6
Numbness	50	30	80	20	0	0	0	0.6
Sharp pain	20	30	50	20	20	10	30	0.0
Dull pain	60	40	100	0	0	0	0	1.0

	Strongly agree (%)	Agree (%)	Total % of agree	Neutral (%)	Disagree (%)	Strongly disagree (%)	Total % of disagree	CVR
Warmth	10	40	50	20	30	0	30	0
Cold	10	20	30	30	40	0	40	-0.4
Throbbing	40	40	80	20	0	0	0	0.6
Other (subject defined)	70	30	100	0	0	0	0	1.0
Overall agreement	38.5	37.7	76.2	11.5	10.8	1.5	12.3	

CVR = Content Validity Ratio

Table 2. Content validity of the prefinal Chinese version of the MASS ("Relevance" to content of *deqi* sensation) (n=11)

	Strongly agree (%)	Agree (%)	Total % of agree	Neutral (%)	Disagree (%)	Strongly disagree (%)	Total % of disagree	CVR
Soreness	80	20	100	0	0	0	0	1.0
Aching	30	40	70	10	20	0	20	0.4
Deep pressure	40	10	50	20	20	0	20	0.0
Heaviness	70	20	90	10	0	0	0	0.8
Fullness/distension	70	20	90	10	0	0	0	0.8
Tingling	50	30	80	20	0	0	0	0.6
Numbness	60	30	90	0	10	0	10	0.8
Sharp pain	30	0	30	10	20	40	60	-0.4
Dull pain	70	30	100	0	0	0	0	1.0
Warmth	40	30	70	30	0	0	0	0.4
Cold	30	10	40	20	30	10	40	-0.2
Throbbing	40	40	80	10	10	0	10	0.6
Other (subject defined)	90	10	100	0	0	0	0	1.0
Overall agreement	53.8	22.3	76.1	11.5	8.5	3.8	12.3	

CVR= Content Validity Ratio

Phase II – Validation of the C-MASS

Subject characteristics

A total of 30 normal healthy subjects were recruited by convenience sampling (18 males, 12 females). The mean age was 34.1 $\pm$3.81 years. Three subjects

experienced only very low level of "soreness / numbness / distension / heaviness" at the intensity and were considered not achieved *deqi* in the first session of electroacupuncture. These subjects were excluded from the analysis as inclusion of vague and uncertain sensations from these subjects was considered inappropriate. Therefore, only 27 subjects were invited to receive the electroacupuncture stimulation again for the test-retest reliability assessment at one to two weeks after the first session. A comparison of the intensity of various *deqi* sensations during the first and second sessions is displayed in Table 3. The result showed that fullness/distension, soreness, numbness, dull pain and aching scored highest among the 13 sensations. No subject added any other sensations in the blank row provided. The mean MASS Index obtained in the first and second assessment sessions were 4.4(SD=1.3) and 4.7(SD=1.1) respectively.

Table 3. Comparison of different *deqi* sensations intensity between Time 1 and Time 2 and Test-retest Reliability of the C-MASS (n=27)

	Sensation Intensity		Test-retest Reliability of the C-MASS			
			$ICC_{3,1}$	95% CI		p-value
	Time 1	Time 2		Lower	Upper	
	Mean ($\pm$SD)	Mean ($\pm$SD)		Bound	Bound	
Soreness	6.9 ($\pm$1.3)	7.6 ($\pm$0.9)	0.62	0.16	0.83	0.009*
Aching	5.9 ($\pm$1.6)	6.3 ($\pm$1.4)	0.98	0.95	0.99	<0.001*
Deep pressure	3.2 ($\pm$1.5)	3.4 ($\pm$1.5)	0.96	0.91	0.98	<0.001*
Heaviness	6.0 ($\pm$1.4)	6.6 ($\pm$0.9)	0.55	0.01	0.88	0.024*
Fullness/distension	7.4 ($\pm$1.2)	7.9 ($\pm$0.8)	0.73	0.41	0.88	0.001*
Tingling	4.0 ($\pm$1.8)	4.4 ($\pm$1.4)	0.89	0.77	0.95	<0.001*
Numbness	6.3 ($\pm$1.3)	6.5 ($\pm$1.1)	0.97	0.92	0.98	<0.001*

Sharp pain	3.2 (±1.7)	4.0 (±1.7)	0.79	0.54	0.91	<0.001*
Dull pain	6.1 (±1.5)	5.9 (±1.4)	0.98	0.95	0.99	<0.001*
Warmth	0.2 (±0.4)	0.2 (±0.4)	0.96	0.91	0.98	<0.001*
Cold	0.2 (+0.4)	0.2(±0.4)	0.96	0.91	0.98	<0.001*
Throbbing	3.4 (±1.4)	3.1 (±1.2)	0.96	0.91	0.98	<0.001*
Other (subject defined)	0.0	0.0	0.92	0.83	0.96	<0.001*
Overall MASS Index	4.4 (±1.2)	4.7 (±1.1)	0.92	0.83	0.96	<0.001*

CI = Confidence Interval, ICC = Intraclass correlation coefficient

* indicate statistical significance (p<0.05)

Validity analysis

The C-MASS showed moderate correlation with the overall SF-MPQ score with the Pearson's product moment correlation coefficient γ=0.56, p=0.001. A significant correlation was observed when the C-MASS was compared with the sensory dimension of the SF-MPQ (Pearson's product moment correlation coefficient γ=0.63, p<0.05) but not with the affective dimension of the SF-MPQ (Pearson's product moment correlation coefficient γ=-0.3, p=0.111).

For factor analysis, based on the criterion of an eigenvalue of more than 1, the data yield 5 factors and explained 77% of variance. Therefore, the results suggested there were 5 factors among the 13 sensations of the C-MASS (Table 4). Naming of factors was based on locating the highest factor loading in the factor. Therefore, Factor 1 was labelled as "Soreness"; Factor 2 was labelled as "Heaviness"; Factor 3 was labelled as "Fullness"; Factor 4 was labelled as 'Dull pain" and Factor 5 was

labelled as "Numbness".

Table 4.   Factor analysis on C-MASS (n=27)

Items	Factors				
	Soreness	Heaviness	Fullness	Dull pain	Numbness
Factor 1	-	-	-	-	-
Soreness	0.90[§]	-	-	-	-
Aching	0.81	-	-	-	-
Pressure	0.53	-	-	-	-
Factor 2	-	-	-	-	-
Heaviness	-	0.87[§]	-	-	-
Warmth	0.43	0.67	-	-	-
Factor 3	-	-	-	-	-
Throbbing	-	-	0.65	-	-
Fullness	-	0.59	0.92[§]	-	-
Factor 4	-	-	-	-	-
Pressure	-	-	-	0.54	-
dull pain	-	0.40	-	0.67[§]	-
Factor 5	-	-	-	-	-
Tingling	0.41	0.86	-	-	0.45
Numbness	-	-	-	-	0.79[§]
sharp pain	-	-	-	-	0.77

(Only factor loading > 0.4 was displayed); [§]highest factor loading in the factor

Internal consistency and test-retest reliability measurements

The overall Cronbach's α of the C-MASS was calculated as 0.71 (Table 5). For test-retest reliability, 3 subjects did not meet the criteria of achieving *deqi* sensation during the first session of electroacupuncture and were thus excluded from the test-retest reliability assessment. Based on the data of the 27 subjects, the test-retest

reliability of individual items of the C-MASS was moderate to excellent ($ICC_{3,1}=$ 0.55-0.98) and that of the MASS index was excellent ($ICC_{3,1} = 0.92$) (Table 3).

Table 5. Item-Total Statistics of the C-MASS (n=27)

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Soreness	27.23	52.94	0.39	0.76	0.69
Aching	27.43	52.12	0.38	0.87	0.69
Pressure	27.97	55.62	0.31	0.69	0.70
Heaviness	26.80	54.03	0.40	0.69	0.69
Fullness	27.53	46.33	0.47	0.72	0.68
Tingling	29.17	56.28	0.48	0.82	0.69
Numbness	26.20	50.99	0.48	0.56	0.68
Sharp pain	28.43	56.19	0.40	0.44	0.69
Dull pain	26.43	51.29	0.34	0.73	0.70
Warmth	29.20	54.92	0.36	0.71	0.69
Cold	30.20	63.41	-0.10	0.61	0.73
Throbbing	29.13	56.46	0.30	0.82	0.70
Other (patient specific)	30.26	61.03	0.14	0.62	0.71

DISCUSSION

This study examined the measurement properties of the translated version of the MASS as an inventory of needle sensations associated with *deqi*. The results demonstrated that the translated version is a valid and reliable instrument for the assessment of needle sensations in Hong Kong Chinese people receiving electroacupuncture.

In the forward translation process, two translators carried out independent forward translations of the instrument to ensure that the translations could be compared, enabling detection of errors and divergent interpretation of ambiguous items in the original, thus reducing the potential bias.[17] Our study also adopted two back translations that were carried out in parallel so as to ensure that the quality of the translation is literally and conceptually the same as the original one.[16-17] The major discrepancy was on the translation of "aching" and "dull pain" as both terms are semantically and pragmatically similar in Chinese translation. The reconciliation of between the two forward translators come to a conclusion that both "aching" and "dull pain" should be used in the Chinese version.

The convergent validity and the discriminant validity of the C-MASS were established by comparing with the SF-MPQ. The overall C-MASS score has moderate correlation with the overall score of the SF-MPQ. Since 4 of the items in the sensory dimension of the SF-MPQ have the same description as the MASS, therefore, a stronger correlation was observed if the C-MASS was compared with the sensory dimension of the SF-MPQ (convergent validity). On the contrary, as we do not expect any affective effect from acupuncture, it is reasonable that only minimal correlation was observed when the C-MASS was compared with the affective dimension of the SF-MPQ (discriminant validity).

The original developers of the SASS conducted a Principle Component Analysis (PCA) which supported the partition of 3 factors among the 9 sensations.[7] When PCA was performed on the C-MASS in this study, 5 factors were extracted under the "eigenvalues greater than 1" rule. These 5 factors explained 77% of the total variance. Factor analysis performed in this study confirmed that the C-MASS is a 5-factor scale. The factors we identified (soreness, heaviness, fullness, dull pain and numbness) are in fact similar to those reported in most TCM literature as the "standard" *deqi* sensations perceived.[3-4,7] Results of this study showed that the test-retest reliability of the C-MASS was high.

From the opinions of the content experts, most of the items in the C-MASS were found to be important and relevant in relation to the content of *deqi* measurement, with "deep pressure", "sharp pain", "warmth" and "cold" being the exception. These items have a low CVR. Previous studies supported that "pressure" appeared to be an important characterization of *deqi*,[15,22,23] whereas "warmth and coolness" sensations are frequently considered as important associates of health in ancient TCM literature (*Suwen*, Chapter 54).[7] According to Suwen (Chapter 54), a patient with "excess" syndrome would feel coldness under the needle when the *yin qi* arrives; on the other hand, a patient with "deficient" syndrome would feel warmth under the needle when the *yang qi* arrives.[7] Our study showed that our subject cohort did not

experience any "cold" sensation or only experienced only low level of "warmth" sensation during acupuncture needling. This is probably because our subjects were young and healthy and thus less sensitive to these sensations, which dominate only when there is an imbalance of *qi* in their health status. In view of the above, we considered deep pressure, warmth and cold sensations are important components of the *deqi* sensations to be included in the instrument.

On the other hand, although the original MASS includes "sharp pain" as one of the descriptors, it is not considered as a sensation associated with *deqi* by the original authors of MASS and others.[7,12,24]. A recent pilot survey on the perception of *deqi* by Chinese and American acupuncturists also revealed that 50% participants classified "sharp pain" as "not *deqi*" and 42% believed it was harmful.[24] Our study showed that with "sharp pain" removed from the scale, the overall Cronbach's α only decreased by 0.02, indicating that the overall homogenicity of the scale was not affected (Table 5). The test-retest reliability was still maintained excellent with the $ICC_{3,1}$ increased from 0.92 to 0.93 (p<0.001). The modified version also showed moderate correlation with the overall SF-MPQ score (γ=0.53, p=0.03) and sensory dimension of the SF-MPQ (γ=0.6, p<0.001). The 5-factor structure was also preserved and it explained 79.6% of the total variance. Since the validity and reliability of the modified scale was not much affected if "sharp pain" was removed, we consider that

it would be appropriate for the descriptor "sharp pain" to be removed from the C-MASS, for assessment of needle sensation associated with *deqi*. As the instrument now is one descriptor short of the original MASS, we thus renamed this Chinese version the "Modified" Massachusetts General Hospital Acupuncture Sensation Scale Chinese version (C-MMASS). The C-MMASS is a 12-item measure that includes soreness, aching, deep pressure, heaviness, full/distension, tingling, numbness, dull pain, warmth, cold, throbbing, and the one supplementary row at the end for subjects to describe perceptions in their own words (Appendix 2).

D*eqi* sensations are complex and highly individualized. In addition, the intensity and nature of *deqi* sensations are determined by many factors such as patient constitution, types of needle, acupoints, needling and allied techniques.[25] In our study, we have minimized all these confounding factors by including only young healthy subjects one type of acupuncture needle was used, only LI4 and LI11 were chosen for stimulation, acupuncture performed by the same trained physiotherapist, and a fixed protocol used in electroacupuncture.

Limitation of the study

There are several limitations in our study. Firstly, we have only recruited healthy subjects at age below 40. Therefore, generalization of the result findings to other age

groups of Hong Kong Chinese would not be deemed appropriate. Secondly, the present study only adopted the LI4 and LI11 acupoints, as they are the classical acupoints commonly used for investigation of *deqi*. It may be possible that other distinct needle sensations could be elicited if different acupoints of other meridians were stimulated. Thirdly, the present study did not use non-acupoints for comparison. Hence, we could not distinguish the sensations related to skin piercing and tissue damage from real acupuncture sensation elicited by acupoints stimulation. Lastly, manual and sham acupuncture were generally used in acupuncture studies that involved *deqi* sensations measurement.[13,25] However, our study had not included these groups for comparison. We anticipate that subjects may rate the needle sensations differently since manual, sham and electroacupuncture may work through different mechanism and, thus, may affect the same individual on the perception of needle sensations.[13]

CONCLUSION

Deqi is an important concept as it may well influence the therapeutic effect of acupuncture. Developing a valid and reliable tool to quantify needle sensations is an important step towards a better understanding of the basic mechanisms underlying the reported therapeutic effects. While there has been no standardized, valid and reliable

tool to quantify the *deqi*, our study has established a Chinese version from the MASS and modified it. We demonstrated that this is a valid tool for assessment of needle sensation associated with *deqi*. This study has also shown that this Chinese version of MASS has good repeatability, producing consistent scores over a period of time, as well as good internal consistency. The 5-factor structure of the translated version was also established. Further study with the involvement of manual acupuncture and sham acupuncture as study groups may confer greater applicability of the C-MMASS in quantifying *deqi* in acupuncture in the future.

Patient consent: Obtained

Ethics approval: This study was conducted with the approval of the Human Subjects Ethics Committee of the Hong Kong Polytechnic University (HSEARS20080514001)

Competing interests: None

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Contributors: All authors were involved in the study design. David TW YU and Alice YM Jones were responsible for the conduct of the study and drafting the manuscript. Marco YC Pang performed statistical analysis.

References

1. Diehl D, Kaplan G, Coulter I, et al. Use of acupuncture by American physicians. J Altern Complement Med 1997;**3**:119-26.

2. Eisenberg DM, Davis RB, Ettner SL, et al. Trends in alternate medicine use in the United States. JAMA 1998;**280**:1569-75.

3. Kaptchuk TJ. Acupuncture: theory, efficacy, and practice. Ann Intern Med 2002;**136**:374-83.

4. Vickers A, Wilson P, Kleijnen J. Acupuncture. Qual Saf Health Care 2002;**11**:92-7.

5. Lao L. Acupuncture techniques and devices. J Altern Complement Med 1996;**2**:23-5.

6. Langevin HM, Churchill DL, Fox JR, et al. Biomechanical response to acupuncture needling in humans. J Appl Physiol 2001;**91**:2471–8.

7. Kong JK, Gollub R, Huang T, et al. Acupuncture De Qi, from qualitative history to quantitative measurement. J Altern Complement Med 2007;**13**:1059-70.

8. Seem GW, Cassidy C, Langevin HM, et al. Getting to the point: a forum for master practitioners, researchers and educators. Clin Acupunct and Orient Med 2003;**4**:148-63.

9. Hsu DT. Acupuncture: A review. Reg Anesth 1996;**4**: 361–70.

10. Vincent CA, Richardson PH, Black JJ, et al. The significance of needle placement site in acupuncture. J Psychosom Res 1989;**33**:489-96.

11. Park J, Park H, Lee H, et al. Deqi sensation between the acupuncture-experienced and the naive: a Korean study II. Am J Chin Med 2005;**33**:329-37.

12. MacPherson H, Asghar A. Acupuncture needle sensations associated with De Qi: a classification based on expert's ratings. J Altern Complement Med 2006;**7**:633-7.

13. Kong JK, Fufa DT, Gerber AJ, et al. Psychophysical outcomes from a randomized pilot study of manual, electro, and sham acupuncture treatment on experimentally induced thermal pain. J of Pain 2005;**6**:55-64.

14. Sali N, Bäumler PI, Simang M, et al. Deqi sensations without cutaneous sensory input: results of an RCT. BMC Complement Altern Med 2010;**10**:81.

15. Napadow V, Dhond RP, Kim J, et al. Brain encoding of acupuncture sensation – Coupling on-line rating with fMRI. NeuroImage 2009;**47**:1055-65.

16. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;**25**:3186-91.

17. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and culture adaptation. Value in Health 2005;**8**:94-104.

18. Lawshe CH. A quantitative approach to content validity. Personnel Psycho 1975;**28**:563-75.

19. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health 1990;**13**:227-236.

20. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987;**30**:191-7.

21. Portney LG, Watkins MP. Foundations of Clinical Research: applications to practice. Upper Saddle River, NJ:Pearson Prentice Hall 2009:523-715.

22. Hui KKS, Liu J, Marina O, et al. The integrated response of the human cerebro-cerebellar and limbic systems to acupuncture stimulation at ST 36 as evidenced by fMRI. NeuroImage;**27**:479-496.

23. Hui KKS, Nixon EE, Vangel MG, et al. Characterization of the "deqi" response in acupuncture. Complement Altern Med 2007;**7**:33

24. Hui KKK, Sporko TN, Vangel MG, et al. Perception of *Deqi* by Chinese and American acupuncturists: a pilot survey. Chin Med 2011;**6**:2.

25. Bovey M. Deqi. J Chin Med 2006;**81**:18-29.