Quality of Life Research (2006) 15: 877–887
DOI 10.1007/s11136-005-4828-x

Rhinitis Symptom Utility Index (RSUI) in Chinese subjects: A multiattribute patient-preference approach

Phoebe S.Y. Lo[1], Michael C.F. Tong[1], Dennis A. Revicki[2], Ching Chyi Lee[3], John K.S. Woo[1], Henry C.K. Lam[1] & C. Andrew van Hasselt[1]

[1]*Division of Otorhinolaryngology, Department of Surgery, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong SAR, China (E-mail: sylo@cuhk.edu.hk);* [2]*Centre for Health Outcomes Research, MEDTAP Institute, Bethesda, MD, USA;* [3]*Department of Decision Sciences and Managerial Economics, The Chinese University of Hong Kong, Hong Kong SAR, China*

Accepted in revised form 4 November 2005

Abstract

Background: The Rhinitis Symptom Utility Index (RSUI), originally developed in the United States, consists of a patient-preference weighting scheme and a 10-item questionnaire measuring the severity and frequency of rhinitis related symptoms over a 14-day period. This study aimed to determine whether the Chinese RSUI could adopt the US-based multi-attribute utility function (MAUF) in scoring rhinitis symptoms. *Methods*: In a Hong Kong study, 116 Chinese adults with allergic rhinitis completed the RSUI questionnaire and 36-item Short-Form Health Survey (SF-36) after they had been seen by two otorhinolaryngologists for disease-severity ratings. Respondents then completed computer-administered direct preference measures, i.e., visual analogue scale (VAS) and standard gamble (SG) assessments. The VAS and SG data were used to estimate a MAUF for the Chinese-based RSUI. *Results*: The derived MAUF was somewhat different than the one developed for the US RSUI. Test–retest reliability for the Chinese RSUI was satisfactory (ICC = 0.71, $p < 0.001$). Scores differentiated among cases with mild, moderate, and severe symptoms ($p < 0.001$); and between those who did and did not require medications to control symptoms ($p = 0.031$). Findings were significantly correlated with SF-36 domain scores ($r = 0.19$ to 0.37; $p = 0.041$ to < 0.001). When the US-based scoring function was applied to the Chinese subjects, the resulting mean RSUI score was significantly lower ($p < 0.001$). Comparisons between directly measured VAS and SG scores between the US and Chinese samples, demonstrated significant differences (all $p < 0.05$), with the US subjects consistently rating rhinitis symptoms as worse than Chinese subjects. *Conclusions*: The Chinese RSUI has good measurement properties that reflect patient preferences from the Chinese. Results suggest that there are differences in preference rating between US and Chinese subjects and that use of the US-based preference function for the RSUI would bias the measurement of rhinitis symptom outcomes in Chinese subjects.

Key words: Health utility measurement, Patient preference, Rhinitis Symptom Utility Index, Standard gamble

Introduction

Over the past two decades, concern regarding limited healthcare resources and rising expenditures for new treatments have increased the application of health economic studies. These cost-effectiveness studies require valid and responsive measures of health outcomes to evaluate 'effectiveness.' Various methods have been developed to compare medical interventions with respect to

their outcomes, including generic and disease specific health status assessments and utility measures.

There are two approaches for health status or health-related quality-of-life (HRQL) evaluation, the psychometric- and utility-based measures. Using a psychometric method, HRQL is assessed with multiple domains such as physical, social, role and emotional function and well-being. A utility-based (preference-based) technique aggregates all health dimensions into a single index on a scale measuring from 0 (dead or the worst health) to 1 (perfect health or the best health). Utility-based assessments can be classified as direct and multi-attribute measures [1]. The direct method requires each individual to value either hypothetical health states or subjectively defined current health states. Common techniques for utility elicitations are standard gamble (SG) and time trade-off (TTO) exercises. The SG method, based directly on the axioms of utility theory [2, 3], requires the respondent to choose between two alternatives: (1) life with a specific condition, or (2) a gamble with either perfect health (probability p) or immediate death/worst health (probability $1-p$). The TTO requests the respondent to consider what length of life they are willing tradeoff to get perfect health. The multi-attribute method captures people's health preferences in an indirect approach with the use of a utility-weighted health classification system and scoring algorithm. This method facilitates the collection of preference-based data in a convenient way by asking subjects to complete a questionnaire based on this health classification system.

As a chronic disease, allergic rhinitis affects no less than 20 percent of the population [4–7], and each year billions of dollars are spent on pharmaceutical products for symptom control [8, 9]. New drugs are increasingly introduced into the health care system, but little is known about patient preferences associated with the effects of these drugs. To help clinicians to select cost-effective treatments, a utility-based instrument is required for measuring patient outcomes for economic evaluations. To date, the Rhinitis Symptom Utility Index (RSUI) developed by Revicki et al. [10], is the only available preference-weighted outcome measure for the evaluation of rhinitis symptoms. Outcome scores are generated using a multi-attribute utility function (MAUF) and rhinitis status classification system based on patient preferences in the United States.

Few studies have examined similarities and differences in preference-based measures between Western and Asian respondents, and it is uncertain whether utilities are generalizable across cultures. Could the US-based multi-attribute preference weighted function be applied to scoring rhinitis symptoms in Chinese patients? This study aimed to answer the question by comparing the RSUI scores derived from the US- and Chinese-based preference functions using data collected from Chinese subjects.

Methods

Rhinitis Symptom Utility Index (RSUI)

The RSUI is a preference-weighted measure of common rhinitis symptoms [10]. In the questionnaire (Appendix), five rhinitis-related symptoms (stuffy or blocked nose, runny nose, sneezing, itchy-watery eyes, and itchy nose or throat) are assessed with four-point Likert scales in regard to their frequency and severity during a 14-day period. The frequency of each symptom is judged as either 'not at all,' '1–3 days,' '4–7 days,' or '8–14 days.' The severity of each symptom is evaluated as 'not applicable,' 'mild,' 'moderate' or 'severe.' The levels within each symptom (attribute) as listed in Table 1 vary from no symptoms to severe symptoms for 8–14 days. The RSUI is a single index of the overall rhinitis state on a scale where the best state (no symptoms) has a score of 1 and the worst rhinitis state (severe symptoms for 8–14 days) has a score of 0.

The RSUI questionnaire was translated into Chinese using an iterative forward-backward translation process, as defined by the International Quality of Life Assessment Project Group [11]. The method for translation followed an iterative process of forward translation by independent translators, review and testing, back translation by independent translators, and review and testing until a linguistic and culturally acceptable translated version was developed. The final Chinese translation of the RSUI was pilot tested with 20 subjects using qualitative methods to examine the

Table 1. Person-mean single-attribute utility functions

Level	Symptom severity and days	Attribute				
		Stuffy nose	Runny nose	Itchy eyes	Itchy throat	Sneezing
1	None	1.00	1.00	1.00	1.00	1.00
2	Mild, 1–3	0.95	0.95	0.98	0.97	0.98
3	Mild, 4–7	0.92	0.92	0.96	0.95	0.96
4	Mild, 8–14	0.88	0.88	0.94	0.93	0.94
5	Moderate, 1–3	0.84	0.84	0.90	0.90	0.90
6	Moderate, 4–7	0.79	0.79	0.86	0.86	0.87
7	Moderate, 8–14	0.72	0.73	0.83	0.82	0.83
8	Severe, 1–3	0.59	0.59	0.73	0.72	0.74
9	Severe, 4–7	0.33	0.33	0.44	0.44	0.44
10	Severe, 8–14	0.00	0.00	0.00	0.00	0.00

level of understanding and comprehension of the translated version and to ensure that it captured the content of the original instrument. The pilot study confirmed the quality of the translation and cultural adaptation of the RSUI into Chinese.

Patient sample

Study subjects were a convenience sample of patients who attended the ENT clinic at a university-affiliated hospital in Hong Kong. Ten patients participated in a pilot test of visual props (i.e., colored chance cards and probability wheels) for assessing utilities, and another 10 patients participated in a pilot test of computer-interactive interviews. Based on the pilot results, we determined that the computer-administered approach was acceptable and worked best given that the respondents demonstrated better comprehension and concentration in the computer administration. One hundred and sixteen Chinese patients who were aged 18 or older and had a history of allergic rhinitis were recruited for the field test. Subjects with sinusitis, cognitive disorder, cold or influenza in the past 14 days, or a nasal surgery in the past 6 weeks were excluded from the study. In this study, 84% of eligible patients contacted agreed to participate. The most frequent reason for non-participation was the study time commitment.

Clinical and health-related measures

Disease severity

Prior to the questionnaire administration and preference interviews, subjects were seen by two otorhinolaryngologists for disease-severity ratings. Lebel et al.'s Rhinitis Symptom Score (RSS) [12] was used to confirm the diagnosis and determine disease severity. RSS scores were rated from 0 to 12 where higher scores indicated more severe rhinitis symptoms. Based on the RSS rating and clinical judgment, each physician independently determined the patient's severity of rhinitis on a six-point Likert scale from 1 (very mild rhinitis) to 6 (most severe rhinitis). The mean of the two physician-ratings were used to define disease severity where scores <3 were defined as mild, scores 3–4 as moderate, scores >4 as severe.

In addition, the respondent was asked to rate his/her current rhinitis status from 0 (the worst possible rhinitis state) to 100 (perfect health). VAS 67–99 was considered as mild, 34–66 as moderate, and 0–33 as severe symptoms. Patients also reported whether or not they required medications to control their symptoms in the past two weeks.

Generic health status measure

The Medical Outcomes Study 36-item Short-Form Health Survey (SF-36) Chinese Hong Kong version [13] was used to examine convergent validity of the RSUI. In the SF-36 [14], one item measures health transition. The remaining 35 items are grouped into eight domains: physical functioning, role limitation due to physical problems, bodily pain, general health, vitality, social functioning, role limitation due to emotional problems, and mental health. Each domain score ranges from 0 to 100, where higher scale scores reflect better functioning and well-being. The SF-36 has extensive evidence supporting reliability and validity, and

has been used in multiple community-based clinical studies.

Preference-based measures

The preference-weighting scheme used to generate the Chinese RSUI was based on the methods in the original US RSUI study [10]. Five single symptoms, each with 10 levels (see Appendix), were rated on the 0–100 point VAS. Five corner symptom states and five hypothetical multiple symptom states (Table 2) were assessed using both VAS rating scales and the standard gamble (SG). A corner symptom state was defined as one symptom at the worst possible level of severity (i.e., most severe) and frequency (i.e., 8–14 days) while the remaining four symptoms were absent. Multiple symptom states were combinations of different rhinitis symptoms with various levels of severity and frequency.

For the VAS ratings, a vertical Feeling Thermometer (FT) marked with numerical scales from 0 to 100 units was presented to the patient in the form of computer image. The FT was anchored at one end by the worst rhinitis state (0) and at one end by the best rhinitis state (100). For the within attribute (symptom) VAS ratings, the 'worst rhinitis state' was the most severe and frequent relevant symptom (i.e., runny nose on 8–14 days) and the best rhinitis state was no symptom (i.e., no runny nose). For the VAS ratings for the corner- and multisymptom-states, the worst rhinitis state was defined as the all-worst state (i.e., all five symptoms most severe on 8–14 days) and the best rhinitis state was a state where none of the symptoms were present. During the VAS assessment, various hypothetical rhinitis states were displayed adjacent to the FT, and the respondent indicated the valuation of each hypothetical rhinitis state by moving the level mark of the thermometer. Ties were allowed and the relative spacing between pairs of states had to reflect the subject's judgment about the relative differences in preferences. The respondent could view previous ratings on the screen. Following the ranking of the single symptom states, corner states (one severe symptom each), and multiple symptom states, the respondent rated his/her own health in current rhinitis status. Time was allowed for the respondent to re-adjust the ratings at the end of each session. Changes were allowed even up to the end of the FT measures. The numerical value of the FT adjacent to the placement of each rhinitis state was immediately saved in an electronic record during the process. To enhance the patient's understanding of the rating tasks, a warm up exercise was performed with the use of paper FT and rhinitis state arrow-cards before the computer administration.

For the SG exercises, two colored circles (Option A and Option B) were displayed on a computer screen. Option A was a certain hypothetical rhinitis state (e.g., a corner state) whereas Option

Table 2. Person-mean (estimated with 10% trimmed mean) of measured VAS preferences, SG utilities and RSUI-derived utilities for the rhinitis states

State	VAS Person-mean	SG Person-mean	RSUI
Corner states			
Severe stuffy nose	0.27	0.61	0.61
Severe runny nose	0.27	0.61	0.61
Severe sneezing	0.33	0.76	0.69
Severe itchy eye	0.32	0.75	0.68
Severe itchy throat	0.32	0.75	0.68
*Multisymptom states**			
Moderate stuffy nose and runny nose (8–14 days)	0.42	0.75	0.73
Moderate stuffy nose and itchy eyes (1–3 days)	0.56	0.88	0.86
Severe stuffy nose/moderate runny nose and itchy eyes (4–7 days)	0.26	0.64	0.61
Severe stuffy nose/moderate runny nose, itchy eyes and itchy throat (1–3 days)	0.40	0.76	0.72
Severe stuffy nose, itchy eyes and itchy throat/moderate runny nose and sneezing (4–7 days)	0.15	0.46	0.40

*If a symptom is not mentioned in the multisymptom state, it is described as mild in the health state.

B was a gamble of an uncertain rhinitis outcome which varied the probabilities of either the best (p) or the worst rhinitis state ($1-p$). The best rhinitis state was described as none of the symptoms present and the worst rhinitis state was the all-worst state (see above). During the computer-interactive interviews, the participant was asked to state his/her preference choice between Option A and Option B. The gamble in Option B started with 100% chance of the best state (no symptoms) and 0% chance of the worst rhinitis state for the following two weeks. If the respondent selected Option B, a click on the B button would make the gamble change to 50% chance of the best state and 50% chance of the worst state. Again, the respondent was asked to state his/her preference between Option A and Option B. The process went on and the probabilities of the best (p) and the worse state ($1-p$) were varied in an iterative interval-division manner [15] until the respondent was indifferent to the choices between Option A and Option B. This exercise was repeated for each of the corner states and multi-symptom states. Patient preferences under uncertainty for the health states were represented on a scale from 0 to 1, where 0 stands for the worst rhinitis state and 1 stands for the best rhinitis state. A warm-up exercise with the aid of a probability wheel was carried out before the actual computer-based interview. The pre-test exercise aimed to introduce or reinforce the concept of probability and most important to ensure subject comprehension on the meaning of 'indifference' between the two alternatives in the SG task.

Data collection procedures

After obtaining consent, the participant was seen separately by two otorhinolaryngologists for the diagnosis and disease-severity rating. On completion of the RSUI and SF-36, the subject was interviewed for preference measures by a single well-trained interviewer. For the evaluation of RSUI test-retest reliability, subjects completed the RSUI two weeks after the baseline administration. They also responded to a question on the change in rhintis status on a five-point global scale where 1 = better, 2 = slightly better, 3 = more or less the same, 4 = slightly worse, and 5 = worse.

Data analysis

The 100-point VAS values were transformed to a scale of 0 to 1, where 0 indicated the worst symptom(s) and 1 represented no symptom(s). To facilitate comparisons with the US-based RSUI, estimation procedures of the MAUF for the current project followed those in Revicki et al.'s study [10]. The conceptual framework and details on the method were described elsewhere [16, 17].

Developing the multi-attribute utility function (MAUF)

We applied the person-mean approach for the estimation of parameters in structuring the MAUF. The person-mean was estimated by 10% trimmed mean of the direct preference measures so as to minimize central tendency bias caused by outliers [18]. Trimmed means of the VAS values and SG utilities for the single symptom states (VAS values only), corner symptom states and multi-symptom states provided the basis of parameters to construct the mathematical function. A power function was used to establish the fitted disvalue-disutility relationship (disvalue = 1 − value; disutility = 1 − utility) and the log function, $\ln(1-u) = \alpha\,[\ln(1-v)]$ where v equals to values and u equals to utilities. A linear regression model with no intercept was estimated using the data of the five corner states and the five multi-symptom states. The standard form for a five-attribute multiplicative disutility function is shown as follows:

$$\bar{U} = (1/c)\left[\prod_{j=1}^{5}(1 + cc_j\bar{u}_j) - 1\right]$$

$$1 + c = \prod_{j=1}^{5}(1 + cc_j)$$

$\bar{U}$ is the disutility index where the scale is 0.00 for perfect health and 1.0 for the all-worst state. Each $\bar{u}_j$ is the single attribute disutility function for the attribute j on a scale where the best level has a disutility of 0.00 and the worst level has a disutility of 1.0, where c and c_j are the model parameters. Each c_j is a scalar constant in conjunction with the corner state j that also contributes to the weightings of the conditional single attribute utilities.

Psychometric evaluation

Psychometric evaluation of the Chinese-based RSUI included test–retest reliability, discriminant and convergent validity. Test–retest reliability (stability) was estimated by the calculation of the intraclass correlation coefficient (ICC) [19] which reflects both degree of correspondence and agreement among ratings [20]. A coefficient of 0.70 is an acceptable indicator of stability for a patient-reported outcome instrument [21]. An agreement between the predicted (RSUI) and directly measured (SG) utility scores was also tested with ICC. Analysis of variance (ANOVA) procedures were utilized to test the ability of the RSUI to discriminate among different levels of disease severity and between those who did and did not require medications to control symptoms in the past 14 days. Inter-rater reliability in the severity ratings between the two physicians was assessed by the calculation of ICC. Convergent validity was examined Pearson product-moment correlations between RSUI and SF-36 domain scores. A low correlation was considered < 0.35, and a moderate correlation was considered in the range 0.35–0.50 [22]. *t*-Tests were used to compare mean VAS preferences and SG utility scores for the directly measured states between the Chinese and US samples. Mean RSUI scores were calculated based on the Chinese- and US-based MAUF. The significant difference between the outcomes resulting from the two functions within subjects was tested by paired *t*-test. The association between the two outcomes was analyzed with Pearson product-moment correlation. Statistical significances were determined by a *p*-value of < 0.05.

Results

Of the 116 subjects who participated in the study, only 112 (59 females and 53 males) were included in the analysis because the SG interviews of 4 subjects (3.5%) were of poor quality. Data quality was determined by respondent self-assessment and interviewer judgment. The mean age of the study subjects was 37.6 ± 11.5 years (range 18–62 years). About 18% ($n = 21$) of the participants achieved a university degree. Four subjects (3.6%) had never received formal education. The remainder of the participants attained either primary (21.4%) or secondary educational levels (56.3%).

Direct preference measures and MAUF

The 10% trimmed means (5% trimmed off of each end of the distribution) of the direct VAS preferences and SG utilities for the corner symptom states and multiple symptom states are shown in Table 2. All SG utility scores were higher than the VAS values. Severe stuffy nose and severe runny nose were rated worse than the other corner states for both VAS preferences and SG utilities.

The regression model for estimating utilities from VAS preferences resulted in a R^2 of 0.979 and an unstandardized regression coefficient, of 2.926. This coefficient was used in the power function to estimate utilities for the within attribute (symptom) levels (Table 1).

We then applied the standard form for a five-attribute disutility function. The multiplicative model best fits the preference data, based on the derived parameters $c = -0.814$, $c_1 = 0.40$, $c_2 = 0.40$, $c_3 = 0.32$, $c_4 = 0.32$, $c_5 = 0.31$, This resulted in a multiplicative MAUF for the Chinese-based $\text{RSUI} = 1.228\ (S_1 \times S_2 \times S_3 \times S_4 \times S_5) - 0.228$, where RSUI is the derived utility for the symptom states on a scale that the best symptom state (no symptoms) has a score of 1 and the worst possible symptom state (severe symptoms for 8–14 days) has a score of 0; and Si is the score for the level on symptom i. Table 3 displays all the coefficients of Si for the rhinitis classification system. The derived RSUI is the score calculated from the MAUF and rhinitis classification system.

Reliability

Rhinitis symptoms often may fluctuate from day to day. Twenty-six percent of our subjects reported improvement and 14.3% reported worsening over a 2-week period. In the examination of test–retest reliability, we only included those subjects ($n = 66$) who reported more or less the same rhinitis status when comparing with their condition at baseline. The ICC, indicating the 2-week reproducibility of the RSUI, was 0.71 ($p < 0.001$).

Validity

The direct utility measures and the derived RSUI scores for the corner and multi-symptom states are displayed in Table 2. The differences between the utility scores obtained from the direct SG and those derived by the RSUI MAUF ranged from −0.07 to 0.00. The standard deviations among the differences for the corner states and multi-symptom states were 0.06 and 0.04, respectively. The agreement between the predicted RSUI and directly measured SG utility scores for the multi-symptom states was high (ICC = 0.98, $p < 0.001$).

Using the Chinese-based MAUF in scoring rhinitis symptoms, the generated RSUI mean score for the study group was 0.70 (SD = 0.22) and the median score was 0.76. RSUI scores did not significantly vary by gender ($p = 0.064$) or age ($p = 0.703$). Moreover, no significant correlations were detected between RSUI scores and years experiencing rhinitis ($p = 0.723$) nor age of onset of rhinitis ($p = 0.591$). Table 4 summarizes the mean RSUI scores by physician-rated and patient self-rated severity in rhinitis status. There was a fairly strong association between the physician-rated and patient self-rated severity scores ($r = 0.65$, $p < 0.001$). Mean RSUI scores varied significantly by both physician-rated ($p < 0.001$) and patient-rated severity of rhinitis symptoms ($p < 0.001$). Inter-rater agreement between the two physician-ratings was only moderate (ICC = 0.58, $p < 0.001$). Significant discrimination properties in the RSUI were also detected based on the two independent physician-ratings ($p = 0.004$ and $p < 0.001$, respectively). The RSUI scores were correlated with the patient self-rated VAS scores for their current rhinitis status ($r = 0.59$, $p < 0.001$). Mean RSUI scores differentiated between the patients who did (mean RSUI = 0.68) and did not require medications (mean RSUI = 0.78) to control symptoms in the past 2 weeks ($p = 0.031$).

In the examination of convergent validity, there were weak to moderate correlations between RSUI and the majority of the SF-36 dimensions. Significant correlations were detected between RSUI and bodily pain ($r = 0.25$, $p = 0.009$), general health ($r = 0.37$, $p < 0.001$), vitality ($r = 0.27$, $p = 0.003$), social functioning ($r = 0.25$, $p = 0.009$), role emotional ($r = 0.19$, $p = 0.041$), and mental health ($r = 0.20$, $p = 0.035$). No significant correlations were observed between the RSUI and physical functioning ($p = 0.298$) or role physical score ($p = 0.154$).

Comparison between Chinese- and US-based preferences and RSUI scores

The directly measured VAS preferences and SG utilities for the five corner states and the five multi-symptom states were compared between the Chinese and US samples. There were statistically significant differences in mean VAS preference scores for all of these health states (all $p < 0.05$; not shown). The differences ranged from 0.07 to 0.25 points. For the SG utilities, there were statistically significant mean differences between all of the

Table 3. Multi-attribute utility function on worst possible symptom state to no symptoms scale for RSUI

Level	Symptom severity & days	Symptom (attribute) Stuffy nose S_1	Runny nose S_2	Itchy eyes S_3	Itchy throat S_4	Sneezing S_5
1	None	1.00	1.00	1.00	1.00	1.00
2	Mild, 1–3	0.98	0.98	0.99	0.99	0.99
3	Mild, 4–7	0.97	0.97	0.99	0.99	0.99
4	Mild, 8–14	0.96	0.96	0.98	0.98	0.98
5	Moderate, 1–3	0.95	0.95	0.97	0.97	0.98
6	Moderate, 4–7	0.93	0.93	0.96	0.96	0.97
7	Moderate, 8–14	0.91	0.91	0.95	0.95	0.96
8	Severe, 1–3	0.87	0.87	0.93	0.93	0.93
9	Severe, 4–7	0.78	0.78	0.85	0.85	0.86
10	Severe, 8–14	0.68	0.68	0.74	0.74	0.75

Scoring formula

$RSUI = 1.228\ (S_1 \times S_2 \times S_3 \times S_4 \times S_5) - 0.228$

Table 4. Mean RSUI scores for patient categories as classified by physician-rated and patient self-rated rhinitis severities

Rhinitis symptoms	Physician-rated severity		Patient-rated severity	
	Mean RSUI (95% CI)*	n	Mean RSUI (95% CI)*	n
Mild	0.82 (0.76–0.88)	37	0.83 (0.78–0.88)	39
Moderate	0.70 (0.65–0.75)	56	0.69 (0.63–0.74)	52
Severe	0.49 (0.40–0.58)	19	0.50 (0.42–0.59)	21
p-value	< 0.001		< 0.001	

*95% confidence interval.

corner and multi-symptom states (all $p < 0.05$), except for the itchy throat corner state ($p > 0.05$) and two of the multi-symptom states (i.e., Severe stuffy nose/moderate runny nose and itchy eyes; and Severe stuffy nose, itchy eyes and itchy throat/moderate runny nose and sneezing) (both $p > 0.05$). Differences in SG utilities ranged from 0.02 to 0.14.

When the US-based MAUF [10] was used to score rhinitis symptoms for the Chinese subjects in the current study, the mean RSUI score was lower, 0.63 (SD = 0.21). There was a statistically significant mean difference of 0.07 (range from −0.06 to 0.23; $p < 0.001$) when compared to the RSUI scores based on the Chinese MAUF. A highly significant correlation was observed between the RSUI scores from the US-based and Chinese-based functions ($r = 0.96$, $p < 0.001$), and the ICC was 0.90 ($p < 0.001$).

Discussion

This study developed the RSUI scoring system based on patient preferences from a Chinese Hong Kong sample. Few study subjects were unable to complete the utility assessment interview. The derived RSUI scores appeared to have slightly under-predicted the directly measured utilities as many of the MAUF-SG differences were negative. The differences however were very minimal. Predictive validity of the MAUF model should be evaluated with the use of inter-survey groups in future research. In the evaluation of reliability, the Chinese RSUI exhibited good reproducibility over 2 weeks in subjects that remained clinically stable (based on self-reports). In the validity evaluation, we demonstrated that the RSUI scores varied significantly by rhinitis-related severity rated by both clinicians and patients. In all cases, patients reporting more severe rhinitis status also had significantly lower (worse) RSUI scores compared to those with less severe rhinitis. The consistency in association between rhinitis severities rated by patients and physicians provides strong supportive evidence for the validity of the RSUI scores. The RSUI scores also were significantly worse in patients who required medication therapy to control their rhinitis symptoms, providing additional supportive evidence for validity.

There are few studies comparing the preferences of Asian and Western subjects for different health states. Furlong and colleagues [23] examined the results from the literature in regard to the MAUFs constructed for the generic instruments between different Western countries and concluded that these functions were fairly similar. However, it is less certain that this conclusion can be generalized to Asian countries. In this article, we compared the directly measured preference scores between the US and Hong Kong samples and found significant differences in the VAS and SG scores. There may be cultural differences between the US and Chinese respondents, as well as differences in demographic and clinical characteristics. For example, while the age and gender distributions were similar between the two samples, the US subjects were more educated. Forty-four percent of the US subjects held a university degree (and 32% had some college) compared with 18% of the Chinese sample. The differences in the preference scores could also be due to the variations between the preference measuring methods, e.g. in-person interviews vs. computer-interactive interviews. We tried to minimize the differences between the two modes of administration by simulating Revicki et al.'s traditional assessment method [10] in the warm up exercises.

For the VAS preference ratings of corner states, Chinese subjects consistently rated the states as 9–12% more preferable than the US subjects. For the SG utilities, the scores obtained from the Chinese and US populations varied and differed substantially for the state involving itchy-watery eyes. These observed difference in the scoring patterns of the five corner states in SG utilities may be due to geographic variations in disease characteristics, where the US subjects rated itchy-watery eyes as worse and the Chinese rated stuffy nose and runny nose as worse. Seasonal allergic rhinitis is commonly seen in the US [9, 24] and it usually induces relatively more ocular symptoms especially during the pollen season [6, 25, 26]. Cases of seasonal allergic rhinitis are rarely seen in Hong Kong. Instead, perennial allergic rhinitis mostly induced by house dust mite is common [27]. In the Hong Kong clinical experience, the most frequent complaints from patients are nasal obstruction and runny nose. This is consistent with the patterns of SG utilities seen in this study. In the current study, 30% of our subjects did not have any eye-related problems, and this is the highest free-of-symptom ratio among all other symptoms (Table 5). Therefore, because of the lower prevalence of itchy-watery eyes in these patients, they may be of less concern.

The diversity of patient preferences in SG measures contributed to the different weightings for the utility-scoring functions between the two cultures. The MAUF for the Chinese RSUI was somewhat different from that derived based on the original US study sample. This resulted in a significant difference in the rhinitis outcome in the Chinese subjects between using the US- and Chinese-based RSUI. Our findings revealed that use of the US-based scoring algorithm would overestimate the severity of rhinitis outcomes for Chinese subjects. Based on the US RSUI calculation, the majority of our subjects (89%) had comparatively worse scores, while 9% of the study group (n = 10) had better rhinitis scores. It was observed that these 10 patients either reported no eye-related symptoms (n = 5) or very mild itchy-watery eyes (n = 5). Results indicated that subjects with moderate or severe itchy-watery eyes might have been overestimated as more severe on their overall rhinitis index.

In conclusion, the present study results support the reliability and validity of the Chinese-based RSUI for the use in the Chinese population. The RSUI allows the incorporation of patient preferences into outcome measurements of rhinitis symptoms. The developed MAUF appears to reflect directly assessed preferences for rhinitis related health states. This study also provides some preliminary evidence that MAUFs and preferences may vary between US and Asian countries. Additional research is needed to examine the nature and extent of differences in preferences for different health states between Asian and Western cultures. In the present study, we are not able to rule out whether or not the differences between the Chinese- and US-based RSUI are associated with subject characteristics or different preference measurement methods. However, we have observed that patient preferences seem to be affected by geographical variations in disease characteristics. Therefore, caution should be taken before the application of a disease-specific preference function from another country. Our results suggested that use of the US-based preference function might bias the measurement of rhinitis outcomes in the Chinese. Based on this study, there is evidence that the Chinese RSUI is reproducible and valid. Further research is needed to demonstrate the predictive validity and responsiveness of the RSUI to clinically meaningful changes in rhinitis status. The RSUI may be useful as a preference-weighted outcome measure for clinical and economic studies comparing different rhinitis treatments.

Table 5. Symptom distributions of the study subjects (n = 112)

Symptom	Stuffy nose n (%)	Runny nose n (%)	Itchy eyes n (%)	Itchy throat n (%)	Sneezing n (%)
None	11 (9.8)	16 (14.3)	34 (30.4)	12 (10.7)	11 (9.8)
Mild, 1–14 days	49 (43.8)	44 (39.3)	55 (49.1)	60 (53.6)	59 (52.7)
Moderate, 1–14 days	41 (36.6)	42 (37.5)	18 (16.1)	34 (30.4)	33 (29.5)
Severe, 1–14 days	11 (9.8)	10 (8.9)	5 (4.5)	6 (5.4)	9 (8.0)

Appendix: Rhinitis Symptom Utility Index

(English Version)

I would like to ask you some questions about different symptoms of rhinitis and how often you were bothered by these symptoms in the past 2 weeks.

1. How many days were you bothered by a stuffy or blocked nose during the **past 2 weeks**?

   0 Not at all (skip to question 3)
   1 1–3 days
   2 4–7 days
   3 8–14 days

2. On average, how severe was your stuffy or blocked nose during the **past 2 weeks**?

   1 Mild
   2 Moderate
   3 Severe

3. How many days were you bothered by a runny nose during the **past 2 weeks**?

   0 Not at all (skip to question 5)
   1 1–3 days
   2 4–7 days
   3 8–14 days

4. On average, how severe was your runny nose during the **past 2 weeks**?

   1 Mild
   2 Moderate
   3 Severe

5. How many days were you bothered by sneezing during the **past 2 weeks**?

   0 Not at all (skip to question 7)
   1 1–3 days
   2 4–7 days
   3 8–14 days

6. On average, how severe was your sneezing during the **past 2 weeks**?

   1 Mild
   2 Moderate
   3 Severe

7. How many days were you bothered by itching, watery eyes during the **past 2 weeks**?

   0 Not at all (skip to question 9)
   1 1–3 days
   2 4–7 days
   3 8–14 days

8. On average, how severe was your itching, watery eyes during the **past 2 weeks**?

   1 Mild
   2 Moderate
   3 Severe

9. How many days were you bothered by itching nose or throat during the **past 2 weeks**?

   0 Not at all
   1 1–3 days
   2 4–7 days
   3 8–14 days

10. On average, how severe was your itching nose or throat during the **past 2 weeks**?

    1 Mild
    2 Moderate
    3 Severe

References

1. Feeny D, Furlong W, Saigal S, Sun J. Comparing directly measured standard gamble scores to HUI2 and HUI3 utility scores: Group- and individual-level comparisons. Soc Sci Med 2004; 58(4): 799–809.
2. Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care 1989; 5(4): 559–575.
3. Feeny D. A utility approach to the assessment of health-related quality of life. Med Care 2000; 38(9 Suppl): II151–154.
4. Quraishi SA, Davies MJ, Craig TJ. Inflammatory responses in allergic rhinitis: traditional approaches and novel treatment strategies. J Am Osteopath Assoc 2004; 104(5 Suppl 5): S7–S15.
5. Santilli J, Nathan R, Glassheim J, Rockwell W, Gold K. Validation of the rhinitis outcomes questionnaire (ROQ). Ann Allergy Asthma Immunol 2001; 86(2): 222–225.
6. Sibbald B, Rink E. Epidemiology of seasonal and perennial rhinitis: clinical presentation and medical history. Thorax 1991; 46(12): 895–901.
7. Spector SL. Overview of comorbid associations of allergic rhinitis. J Allergy Clin Immunol 1997; 99(2): S773–780.
8. Law AW, Reed SD, Sundy JS, Schulman KA. Direct costs of allergic rhinitis in the United States: Estimates from the 1996 Medical Expenditure Panel Survey. J Allergy Clin Immunol 2003; 111(2): 296–300.
9. Crystal-Peters J, Neslusan CA, Smith MW, Togias A. Health care costs of allergic rhinitis-associated conditions vary with allergy season. Ann Allergy Asthma Immunol 2002; 89(5): 457–462.
10. Revicki DA, Leidy NK, Brennan-Diemer F, Thompson C, Togias A. Development and preliminary validation of the multiattribute Rhinitis Symptom Utility Index. Qual Life Res 1998; 7(8): 693–702.
11. Ware JE Jr, Keller SD, Gandek B, Brazier JE, Sullivan M. Evaluating translations of health status questionnaires. Methods from the IQOLA project International Quality of

Life Assessment. Int J Technol Assess Health Care 1995; 11(3): 525–551.
12. Lebel B, Bousquet J, More A, Chanal I, Godard P, Michel FB. Correlation between symptoms and the threshold for release of mediators in nasal secretions during nasal challenge with grass-pollen grains. J Allergy Clin Immunol 1988; 82(5 pt 1): 869–877.
13. Lam CL, Gandek B, Ren XS, Chan MS. Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 health survey. J Clin Epidemiol 1998; 51(11): 1139s–1147.
14. Ware JE Jr, Sherbourne CD.. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30(6): 473–483.
15. Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber A. The effect of search procedures on utility elicitations. Med Decis Making 1998; 18(1): 76–83.
16. Torrance GW, Feeny DH, Furlong W, Barr R, Zhang Y, Wany Q. Multi-attribute preference functions for a comprehensive health status classification system: Health Uilities index Mark 2. Med Care 1996; 34(7): 702–722.
17. Torrance GW, Furlong W, Feeny D, Boyle M. Multiattribute preference functions. Health Utilities Index. Pharmacoeconomics 1995; 7(6): 503–520.
18. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002; 40(2): 113–128.
19. Ottenbacher KJ, Tomchek SD. Reliability analysis in therapeutic research Practice and procedures. Am J Occup Ther 1993; 47(1): 10–16.
20. Portney LG, Watkins MP. Foundations of Clinical Research. Applications to Practice. Norwalk: Connecticut, 1993.
21. Patrick DL, Wild DJ, Johnson ES, Wagner TH, Martin MA. Cross-cultural validation of quality of life measures. In: Orley J, Kuyken W (eds.), Quality of Life Assessment: International Perspectives. Berlin: Springer-Verlag, 1994: 24.
22. Lehman AF, Postrado LT, Rachuba LT. Convergent validation of quality of life assessments for persons with severe mental illnesses. Qual Life Res 1993; 2(5): 327–333.
23. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. Ann Med 2001; 33(5): 375–384.
24. Nathan RA, Meltzer EO, Selner JC, Storms W. Prevalence of allergic rhinitis in the United States. J Allergy Clin Immunol 1997; 99(6 Pt 2): S808–S814.
25. Lund V. Allergic rhinitis-making the correct diagnosis. Clin Exp Allergy 1998; 28(S6): 25–28.
26. Sly RM. Changing prevalence of allergic rhinitis and asthma. Ann Allergy Asthma Immun 1999; 82(3): 233–248.
27. Tsang RK, Tong MC, Woo JK, Van Hasselt CA. A prospective study on the efficacy of mometasone furoate monohydrate aqueous nasal spray on Chinese patients with allergic rhinitis. Otolaryngol Head Neck Surg 2003; 128(4): 497–502.

Address for correspondence: Phoebe S.Y. Lo, Division of Otorhinolaryngology, Department of Surgery, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong SAR, China
Telephone: +852-26322884; Fax: +852-26324678
E-mail: sylo@cuhk.edu.hk