Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Validation of the Wisconsin upper respiratory symptom survey-24, Chinese version

Yuanyuan Wang[a,b]*, Zehui He[c]*, Simin Chen[d,b], Yuntao Liu[b,e], Fang Li[b], Bruce Barrett[f], Zhongde Zhang[b,e], Guobin Su[g,h,a,i] and Cecilia Stålsby Lundborg[a] (ID)

[a]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden; [b]Department of Emergency, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, PR China; [c]Department of Clinical Epidemiology, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, PR China; [d]Intensive Care Unit, Traditional Chinese Medicine Hospital Dianjiang, Chongqing, PR China; [e]Guangdong Provincial Key Laboratory of Research on Emergency in TCM, Guangzhou, PR China; [f]Department of Family Medicine and Community Health, University of Wisconsin, Madison, WI, USA; [g]Department of Nephrology, Guangdong Provincial Hospital of Chinese Medicine, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, PR China; [h]Department of Nephrology, National Clinical Research Center for Kidney Disease, State Key Laboratory of Organ Failure Research, Guangdong Provincial Clinical Research Center for Kidney Disease, Guangdong Provincial Institute of Nephrology, Guangdong Provincial Key Laboratory of Renal Failure Research, Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Nanfang Hospital, Southern Medical University, Guangzhou, PR China; [i]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

ABSTRACT

Background: The Wisconsin upper respiratory symptom survey (WURSS) is a validated English questionnaire to evaluate the quality of life and severity of upper respiratory tract infections (URTIs). We aimed to develop a Mandarin Chinese version of WURSS-24 (WURSS-24-C) and evaluate its reliability, validity and minimal important difference (MID).

Methods: The WURSS-24-C was developed using the forward-backward translation procedure. People with URTIs' symptoms within 48 h of onset were recruited and asked to fill in the WURSS-24-C daily for up to 14 d. Exploratory and confirmatory factor analyses were used to suggest domains. The 8-Item Short Form Health Survey (SF-8) assessing general mental and physical health was used to assess validity. Reliability estimated by Cronbach's alpha and mean day-to-day change for those indicating minimal improvement as MID were evaluated.

Results: The WURSS-24-C was found to be acceptable, relevant, and easy to complete in cognitive debriefing interviews. A total number of 300 participants (age 28.4 ± 9.3, female 70%) were monitored for 2500 person-days. Four domains (activity and function, systemic symptoms, nasal symptoms and throat symptoms) of the WURSS-24-C were confirmed (comparative fit index [CFI] = 0.93). The reliability of this 4-domain-structure is good (Cronbach's alphas varied from 0.849 to 0.943). Convergent validity is moderate (Pearson correlation coefficients between daily WURSS-24-C and the SF-8 were -0.780 and -0.721, for the SF-8 physical and mental health, respectively). Estimates of MID for individual items varied from -0.41 to -1.14.

Conclusions: The WURSS-24-C is a reliable and valid questionnaire for assessing illness-specific quality-of-life health status in Chinese-speaking patients with URTIs.

KEY MESSAGES

- The Wisconsin upper respiratory symptom survey (WURSS) series are patient-oriented questionnaire instruments assessing the quality of life and severity of upper respiratory tract infections (URTIs).
- The WURSS-24 was translated into Mandarin Chinese using the forward-backward translation procedure, and evaluated its validity, reliability and minimal important difference (MID) in 300 Chinese participants with URTIs.
- The WURSS-24 Chinese version (WURSS-24-C) seems to be a reliable and valid questionnaire for assessing illness-specific quality-of-life health status in Chinese patients with URTIs.

Abbreviations: WURSS: the Wisconsin Upper Respiratory Symptom Survey; WURSS-24-C: the Chinese version of Wisconsin Upper Respiratory Symptom Survey with 24 items; URTIs: upper

CONTACT Guobin Su ✉ guobin.su@ki.se 🖃 Department of Nephrology, Guangdong Provincial Hospital of Chinese Medicine, the Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, PR China; Zhongde Zhang ✉ doctorzzd99@163.com 🖃 Department of Emergency, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, 111 Dade Road, Yuexiu District, Guangzhou City, 510120 PR China
*The authors contributed equally to this work.

ⓘ Supplemental data for this article can be accessed here.

respiratory tract infections; HRQoL: health-related quality of life; SF-8: the 8-item short-form; MID: minimal important difference; EFA: exploratory factor analysis; CFA: confirmatory factor analysis; CFI: comparative Fit Index; GFI: goodness of fit index; AGFI: adjusted goodness of fit index; NFI: normed fit index; SRMR: standardized root mean square residual; RMSEA: root mean square error of approximation; SD: standard deviations; NSAIDs: non-steroidal anti-inflammatory drugs

Background

Upper respiratory tract infections (URTIs), including the common cold and influenza, are among the most common human illnesses [1,2]. Although generally mild and self-limited, these illnesses can lead to pneumonia and can even be life-threatening, especially in elderly persons and those with underlying disease, such as chronic obstructive pulmonary disease, heart failure and chronic kidney disease, etc. [3–5]. In addition, URTIs lead to increases in inappropriate antibiotics use [6,7], hospitalizations and mortality [8], resulting in substantial economic burden.

Patients-oriented outcomes are important for outcome assessment in both clinical practice and clinical trials [9,10]. Despite laboratory measures of URTI, like counts of white blood cells and identification of virus are sometimes useful, none of them correlate well with specific symptoms and functional impairments [11]. Those with URTIs usually have symptoms, such as sore throat, rhinitis, rhinorrhoea, cough and malaise, which are associated with poorer health-related quality of life (HRQoL) [2]. However, limited tools or scales are available to assess patient-oriented outcomes, such as specific symptoms and functional impairments in patients with URTIs.

The Wisconsin upper respiratory symptom survey (WURSS), developed by Barrett et al. [12,13], is such an illness-specific questionnaire instrument assessing URTIs emphasizing patient-oriented outcomes. WURSS has different language versions including English, Spanish, French, German, Korean, etc. The English and Korean versions have been tested for reliability and validity in previous studies [14,15]. WURSS-24 is one of the widely used versions of WURSS aiming to assess influenza-like illness and has been used for an extensive assessment of patients with URTIs in clinical practice [16,17].

However, the lack of the WURSS Chinese version limits its potential application in the Chinese population, which is one of the most common languages used by approximately 16% of the global population [18]. And the lack of the WURSS-24 Chinese version (WURSS-C) limited its use in multinational studies. In this study, we developed the WURSS-24-C and evaluated its validity, reliability and minimal important different (MID) in Chinese patients with URTIs.

Methods

This study was designed as a prospective observational study. It was approved by the ethical committee of Guangdong provincial hospital of Chinese medicine, Guangzhou, China (GPHCM; B2016-090-02). Written informed consent was provided by all participants.

Translation of the WURSS-24

A three-step linguistic validation procedure was employed [19].

Forward translation

Independent forward translations were performed by two bicultural native speakers of Chinese. One is a professional of clinical medicine, comprehending health care terminology and the content area of the construct of the instrument. Another is familiar with colloquial phrases, health care slang and jargon, idiomatic expressions and emotional terms in common use. Then reconciliation of two forward translations was made by a third bilingual and bicultural independent translator regarding ambiguities and discrepancies of words, sentences and meanings.

Back translation

Following forward translation, a literal back translation was completed by two native speakers of American English, who have the same characteristics described above in forward translation. They were blind to the original version of the survey.

The translation reconciliation was finished by a committee, including all the four bilinguals and a respiratory clinician. They compared the back-translated version with the source to highlight and investigate discrepancies between the source and the translation. After the proofing/formatting of the reconciled translation, the committee detected possible translation discrepancies between different language versions, and harmonized linguistic solutions across both languages.

Cognitive debriefing interviews and finial review

A group of five subjects were recruited in mainland China to evaluate alternative wording and to check understanding, interpretation and cultural relevance of

the translation. In the cognitive debriefing, five participants were interviewed to comment on questionnaire items and instructions. All of them are native speakers residing in mainland China who were experiencing a common cold illness.

The final revision of WURSS-24-C was adapted according to the review of cognitive debriefing results and reporting. This process involves comparing each subject's interpretation of the translation with the source version to highlight discrepancies and adjust translations as needed; as well as, compiling and documenting the detailed accounts of each subject interview.

Participants

Participants with new-onset URTIs were recruited from smartphone online application and in outpatient in Guangdong provincial hospital of Chinese medicine. This hospital has four hospital branches located in different districts of Guangzhou city, Guangdong province, South China. Participants were recruited at outpatient clinics, or by advertisement on WeChat, a widely used chatting app in China with function of messaging and social networking. Respondents were invited by research assistants to meet for screening and informed consent.

Inclusion, exclusion and termination criteria

Participants who satisfied the following criteria were included: (1) age 18 or older; (2) diagnosis of URTIs, such as "common cold" or "influenza" by research physicians according to Jackson scale [20,21], with at least one symptom from the following cold symptoms (or synonyms): a) nasal discharge (runny nose), b) nasal obstruction (plugged nose, stopped up nose and stuffiness), c) sneezing or d) sore throat (scratchy throat). (3) With onset of symptoms within 48 h. (4) Willing to participate and able to complete the questionnaire.

The exclusion criteria were: (1) Any symptoms likely due to allergy, or other non-URTI cause, such as asthma, allergic rhinitis, etc.; or (2) those patients who require hospitalization at enrolment; or (3) those with previous history of chronic respiratory diseases, such as chronic obstructive pulmonary disease, bronchiectasis, sinusitis recurring more than twice per year, anatomical nasal obstruction or deformity, otitis and exudative pharyngitis.

During the follow-up period, the observation of the participants was terminated if one of the following criteria was fulfilled: (1) The disease of the URTIs progressed to lower respiratory tract infection, such as bronchitis or pneumonia, etc., confirmed by chest X-ray or physical examination by physicians. (2) The condition of the URTIs was getting worse or developing into complications that require hospitalization. (3) The course of the illness exceeded 14 d.

Sample size

According to one guideline for cross-cultural health care research [22], the sample size should be at least 10 subjects for each item of the instrument scale and for item analysis and exploratory factor analysis (EFA). Therefore, a sample of at least 240 ($=24 \times 10$) participants were considered the minimum required.

Measures

Participants were asked to fill in the WURSS-24-C every day until they answered "Not sick" two days in a row to the question, "How sick do you feel today?", otherwise to a maximum of 14 d. Summary scores for WURSS-24-C were calculated by summing scores of individual item scores, excluding the first and last items, which had categorically different reference domains, and were analysed separately. Another health-related questionnaire, the 8-Item Short Form Health Survey (SF-8), 24 h recall [23], was simultaneously applied to the study population along with the WURSS-24-C daily to assess validity. The SF-8 is a short form 24-h recall version of the widely used SF-36 [24], and yields separate summary scores for physical and mental health, calculated using algorithms and scoring software provided by OptumInsight Life Sciences, Inc (QualityMetric Health Outcomes (TM) Scoring Software version 4.5; Boston, MA).

Medicine was permitted during the study. Protocol adherence was supported by regular message contact. Participants were followed up *via* telephone or message on Day 3, 5, 7 and 11, respectively, or until the survey was accomplished to ensure all the questionnaires were filled in properly. WURSS-24-C and SF-8 were returned at an in-person exit interview or mailed to the researchers after the cold ended.

Statistical analysis

Data were double entered, with a resolution of discrepancies by comparison to paper questionnaires. Missing data and outliers were checked and corrected if appropriate. Data were presented in frequencies for categorical variables and as mean and standard

deviation (SD) or median and interquartile range for continuous variables, depending upon their distribution.

EFA was used to extract the factor structure using the principal components analysis method [25]. It began with Day 3 data, chosen because this day represented the breadth of symptomatic and functional impairment which was aligned with the validation analysis of original WURSS [13]. The data of Day 3 was used for this analysis based on the usual duration and recovery process of the common cold. Factorial structures were assessed for 22 items (excluding the first and last items measuring global severity and global change) using principal component analysis.

Confirmatory factor analysis (CFA) was followed to assess the model fit using the EFA dataset, according to the method of Mulaik [26]. Several goodness-of-fit indices were reported including goodness of fit index (GFI), adjusted goodness of fit index (AGFI), normed fit index (NFI), comparative fit index (CFI), standardized root mean square residual (SRMR) and root mean square error of approximation (RMSEA).

Reliability as the measure of internal consistency was estimated by Cronbach's alpha coefficient using the data of Day 3 with the factor structure conducted by EFA.

Convergent validity was assessed by Pearson's correlation coefficients between WURSS-24-C and SF-8.

MID was considered as day-to-day change for those indicating minimal improvement. Using methods developed by Jaeschke et al. and Guyatt et al. [27,28], MID was defined as the average amount of instrument-assessed change for all subjects who rated themselves as "a litter better" or "somewhat better".

Statistical analysis was performed using SPSS version 17.0 (SPSS Inc., Chicago, IL) and LISREL version 8.8 (Scientific Software International, Chicago, IL).

Results

Translation of WURSS-24 and cognitive debriefing interviews

Information about the translation and back translation is provided in the Supplementary Materials (Supplementary Item 1). No major difficulties were experienced during the forward translation process; minor discrepancies were harmonized after discussion by the committee. After a comparison of the back-translated version with the original English version, there were no significant differences in linguistic and conceptual contents. We then adapted a reconciled translation across both languages and used it in the

cognitive debriefing interviews. There were five subjects (three females and two males, ageing from 21 to 69) interviewed who reported no difficulty in understanding the items or answering the questions. The final version was produced after iterative refinement, reconciliation and proof-reading (Supplementary Item 2).

Baseline characteristics and descriptive data

A total of 315 participants were initially enrolled from May 2018 to February 2019. Eleven participants did not return their questionnaires, four participants' information were incomplete and therefore 300 were included in the analysis, for 2500 person-days covered by this study.

Table 1 describes sociodemographic and clinical characteristics for the study participants. The mean age of participants was 28.4 years (SD = 9.3) and 70% of participants were women. Almost all of the participants (99.3%) were Han population, well-educated (76% with above high school education level, and did not smoke (89%). Most of the participants (96.3%) were healthy people without comorbidities. The mean time from first symptom to enrolment were 20.8 h (SD = 14.6). A majority of 181 (60.3%) had body temperature that was normal at enrolment; 77 (25.7%) had low fever (37.3–38 °C) , 35 (11.7%) had temperature of 38.1–39 °C and 7 (2.3%) were higher than 39 °C. A total of 140 out of 300 participants had not taken medicine before enrolment, others had taken Chinese medicine (115), antipyretic analgesics (31), antivirals (27) and antihistamine (3).

The total WURSS-24-C scores of all participants from the 1st to the 14th day are shown in Table 2. All of the participants filled in the questionnaire for at least 3 d. A total of 27(9%) participants continued to report being sick at the end of their 14-d monitoring period. The symptoms and functions improved over time. The mean score ranged from 54.11 (SD = 30.28) in the first day to 7.69 (SD = 10.33) in the 14th day.

Factor analysis

The scree plot of the EFA indicated that the 4-domain structure was accepted which explained 68.25% of the total square deviation. The 4 domains were defined as activity and function, systemic symptom, nasal symptom and throat symptom. The factor loading coefficients of individual items in each domain are shown in Table 3, only those with promax rotation factor loaded greater than 0.4 were accepted. All of the

Table 1. Demographics of the study participants who had common cold during May 2018 to February 2019 visited Guangdong provincial hospital of Chinese medicine, China.

Variable	Value
Number of participants	300
Age, years	
Mean (SD)	28.4 (9.3)
Gender, no./total (%)	
Female	210 (70)
Ethnicity, no./total (%)	
Han	298 (99.3)
Education, highest, no./total (%)	
High school or less educated	76 (25.4)
College degree	121 (40.3)
Graduate degree	103 (34.3)
Tobacco use, no./total (%)	
Current	31 (10.3)
Past	2 (0.7)
Nonsmoker	267 (89)
Comorbidities/disease history, no./total (%)	
No	289 (96.3)
Yes	11(3.7) 3 pharyngitis, 2 anaemia and 1 chronic Bronchitis, 1 hypertension 1 hypothyroidism and 1 Optic neuritis, 1 calculus of kidney and 1 Appendicitis
Time from first symptom to enrolment (hours)	
Mean (SD)	20.8 (14.6)
Inter-quartile range	24–30
Maximum body temperature before enrolment	
Mean (SD)	37.3 (0.7)
$\leq$37.2 °C, no./total (%)	181 (60.3)
37.3–38 °C, no./total (%)	77 (25.7)
38.1–39 °C, no./total (%)	35 (11.7)
>39 °C, no./total (%)	7 (2.3)
Having taken medicine before enrolment, no.	
No medicine	140
Chinese medicine	115
Antipyretic analgesics	31
Antivirals	27
Antihistamine	3
Other	6

Table 2. The WURSS-24-C item and summary descriptive statistics.

Day	N	Mean (SD)	Min	Max
1	300	54.11 (30.28)	4	137
2	300	49.82 (27.67)	0	130
3	300	41.74 (27.39)	0	108
4	298	30.89 (27.14)	0	125
5	290	21.00 (22.05)	0	123
6	264	13.78 (18.35)	0	112
7	223	10.16 (15.67)	0	82
8	168	8.79 (14.55)	0	87
9	119	8.41 (14.23)	0	71
10	81	8.02 (12.73)	0	66
11	57	7.70 (12.83)	0	66
12	41	9.59 (14.47)	0	66
13	32	10.63 (14.03)	0	45
14	27	7.96 (10.33)	0	30

Data shown represent the average and the range of the total score from Day 1 to 14.

items were loaded into the 4-domain structure. Besides, item 11 (feeling tired) was grouped in both activity and function and systemic symptom domain, while item 17 (breath easily) was grouped in activity and function and nasal symptom. (Table 3). CFA was

Table 3. The factor loading of WURSS-24-C items.

Item	Symptom	1	2	3	4
1	Runny nose	–	–	0.867	–
2	Plugged nose	–	–	0.759	–
3	Sneezing	–	–	0.812	–
4	Sore throat	–	–	–	0.721
5	Scratchy throat	–	–	–	0.798
6	Cough	–	–	–	0.686
7	Hoarseness	–	–	–	0.654
8	Head congestion	–	0.563	–	–
9	Chest congestion	–	0.552	–	–
10	Feeling tired	0.505	0.586	–	–
11	Headache	–	0.786	–	–
12	Body aches	–	0.812	–	–
13	Fever	–	0.687	–	–
14	Think clearly	0.693	–	–	–
15	Sleep well	0.636	–	–	–
16	Breath easily	0.426	–	0.664	–
17	Walk/climb stairs	0.828	–	–	–
18	Accomplish daily activities	0.868	–	–	–
19	Work outside the home	0.864	–	–	–
20	Work inside the home	0.839	–	–	–
21	Interact with others	0.845	–	–	–
22	Live your personal life	0.843	–	–	–
Definition of the 4 domains		Activity and function	Systemic symptom	Nasal symptom	Throat symptom
	Cronbach's α	0.943	0.859	0.882	0.849

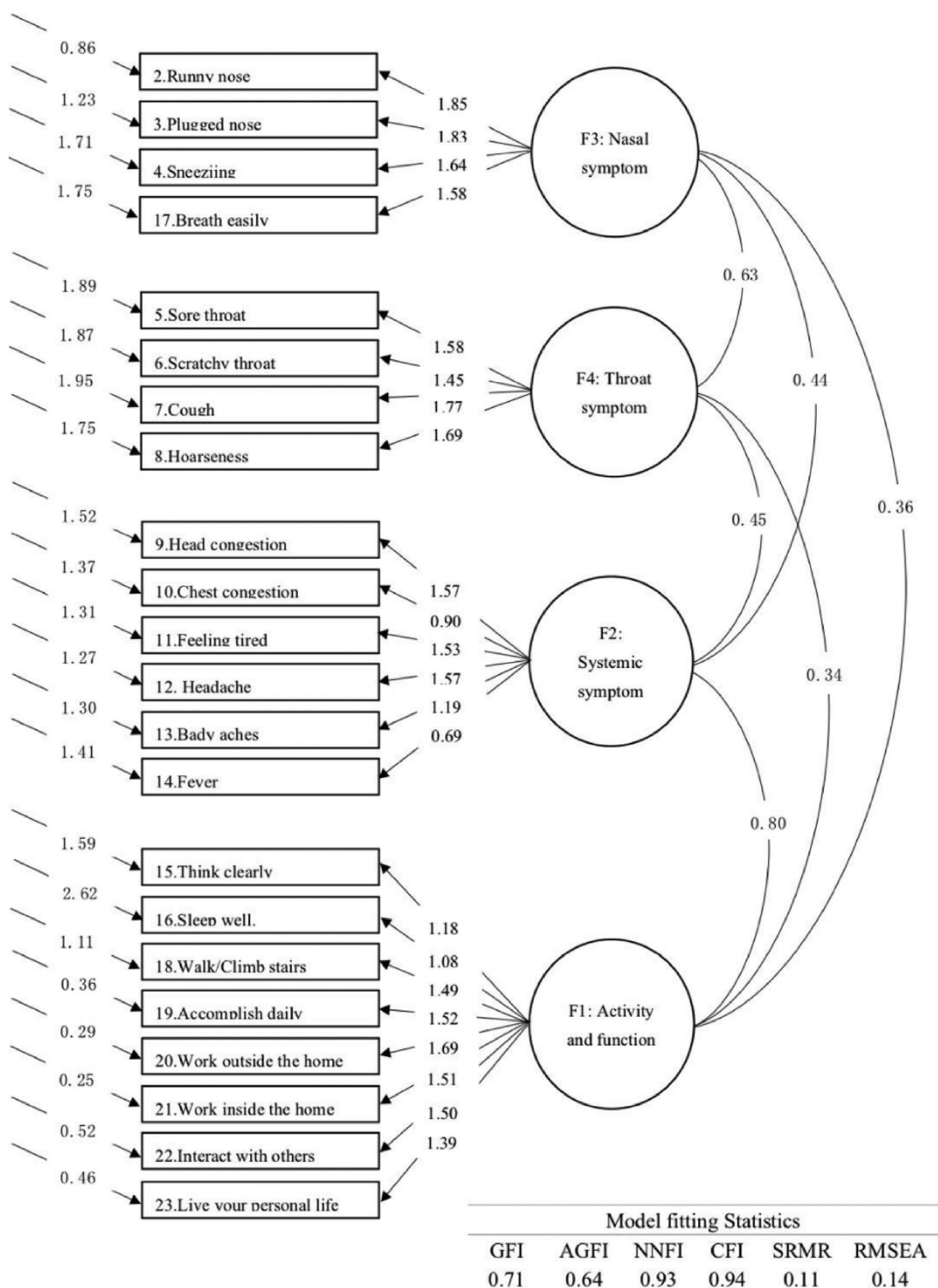Only those with factor loads greater than 0.4 are listed.

Figure 1. The 4-domains model fit confirmatory factor analysis (CFA) for the WURSS-24-C. Chi-square $= 1673.47$, df $= 203$, p value $= .0000$. GFI: goodness of fit index; AGFI: adjusted goodness of fit index; NNFI: non-normed fit index; CFI: comparative fit index; SRMR: standardized root mean square residual; RMSEA: root mean square error of approximation.

performed to evaluate the 4-factor structure validity of the WURSS-24-C. The goodness-of-fit indices and the structure graphs are shown in Figure 1. The CFI was 0.94, RMSEA was 0.14 and SRMR was 0.11, which suggest acceptable model fit in this four-factorial model (Figure 1).
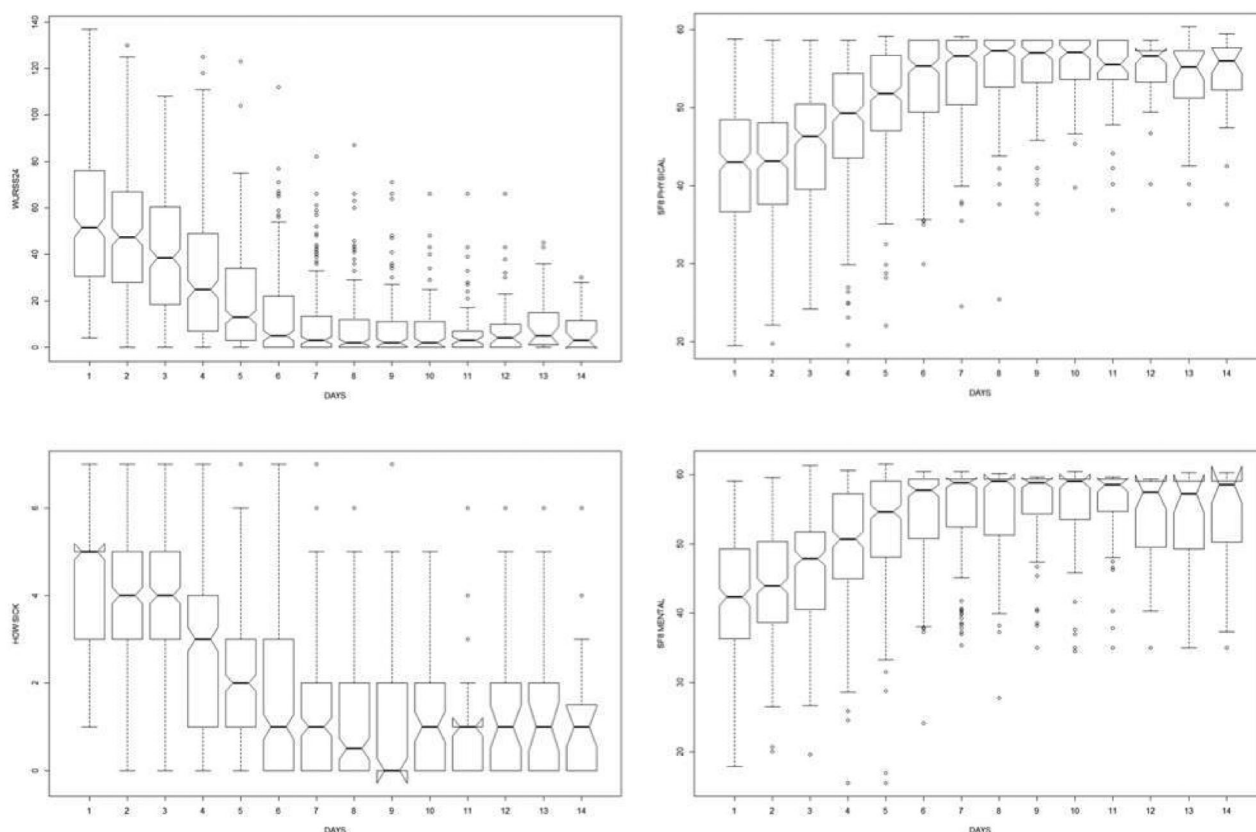
Figure 2. The average daily score of WURSS-24-C and SF-8 from Day 1 to 14. Sample size diminishes as participants' colds resolve, from $N = 300$ on Day 1 to $N = 27$ on Day 14. The centre of the notched boxes is the median summed score for that day. The top of the notched boxes indicates the 25 and 75% percentiles, respectively. The notches portray the median ± 1.57 (interquartile range = IQR)/N^{-2} and thus can be compared to assess difference at the $p = .05$ level of significance. The ends of the vertical lines indicate the last actual data point within 1.5 (IQR) from the 25%ile and 75%ile. The symbols above and below these lines are actual outlying data points.

Reliability

After establishing the 4-domain-structure, we tested the Cronbach's alpha for internal reliability. Cronbach alpha coefficients were 0.943, 0.859, 0.882 and 0.849 for activity and function, systemic symptom, nasal symptom, and throat symptom, respectively (Table 3). In addition, we tested the internal reliability of the whole questionnaire (based on the data of Day 3), of which the Cronbach's alpha was 0.940.

Convergent validity

Figure 2 shows a daily change of illness severity over time as measured by the WURSS-24-C and the SF-8 (both physical and mental health scores). The number of participants decreased as their symptoms decreased, from $N = 300$ on Day 1 to $N = 27$ on Day 14. As measured by the SF-8, the trend of general physical health and mental health was similar during the illness process.

WURSS-24-C had a similar change with the SF-8. All changes were more rapid in the first 7 d than in the later periods. The Pearson correlation coefficient of the WURSS-24-C yielded -0.780 for SF-8 physical health ($p < .001$) and -0.721 for SF-8 mental health ($p < .001$) (Figure 3).

Frequency, severity and MID

Table 4 displays the pattern of experienced symptoms and functional limitations in the first 3 d. The most reported items were *Feeling tired* (98%) and *Sleep well* (91.7%), followed by nasal symptoms: *Plugged nose* (87.7%); affected abilities: *Think clearly* (86.3%); and throat symptoms: *Sore throat* (85.7%). The most severe symptom was *Plugged nose* with a score of 3.50, followed by *Feeling tired* (3.37), *Sore throat* (3.34) and *Runny nose* (3.31). MID was presented item-by-item for the WURSS-24-C, ranged from -0.41 (*Chest congestion*) to -1.14 (*How sick*).
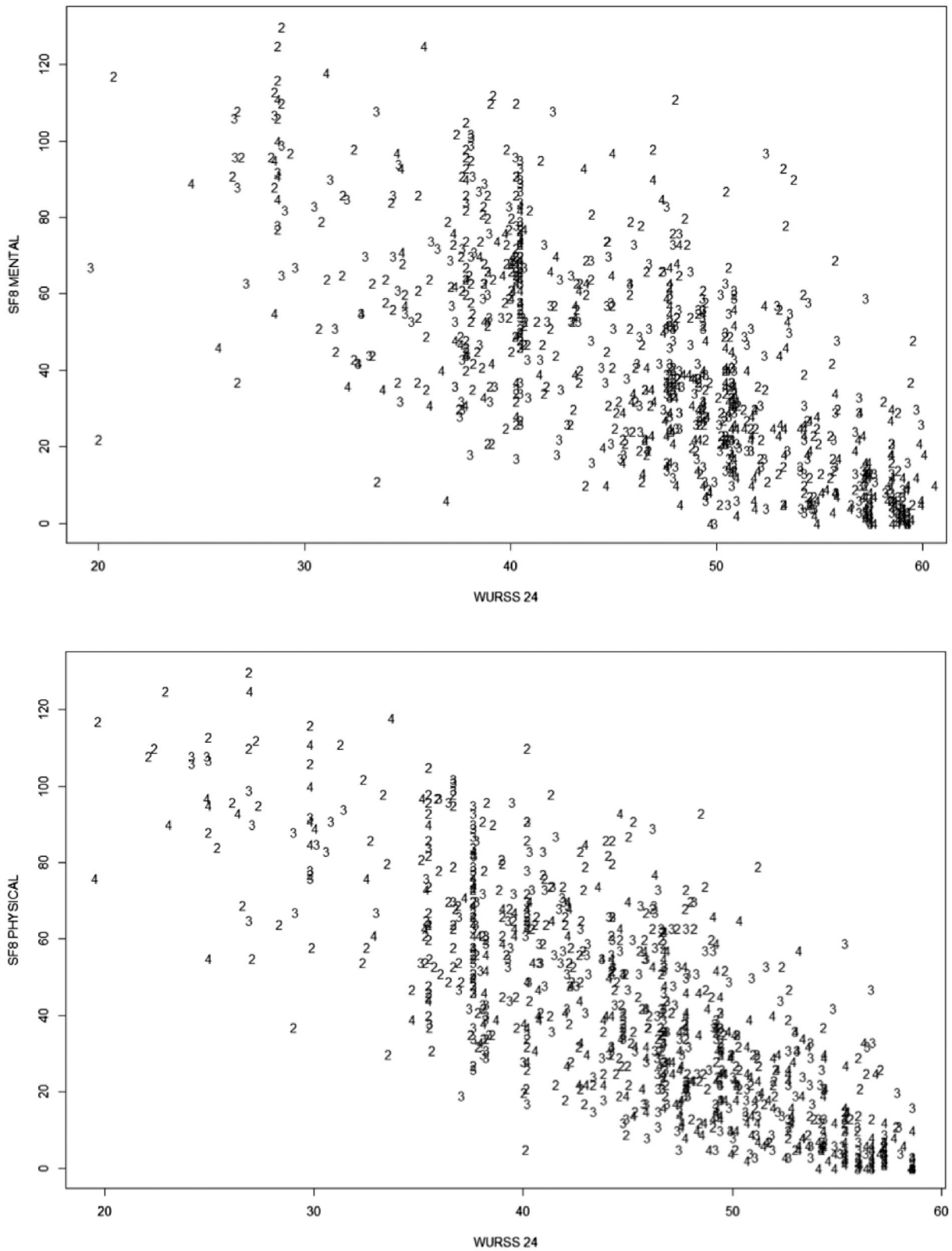
Figure 3. The scatterplot correlations of the WURSS-24-C with SF-8. Data shown represent Days 2–4, where sample size was $N = 300$, $N = 300$ and $N = 298$, respectively. The WURSS-24 correlated more statistically with physical than mental health, yielding Pearson correlation coefficient of -0.780 for SF-8 physical health and -0.721 for SF-8 mental health.

Table 4. Frequency, severity and minimal important difference of WURSS-24 Items.

Item	Symptom	Frequency%	Severity	MID
1	How sick	100.0	4.11 (1.20)	−1.14
2	Runny nose	85.0	3.31 (1.57)	−0.88
3	Plugged nose	87.7	3.50 (1.58)	−0.95
4	Sneezing	84.3	3.10 (1.54)	−0.91
5	Sore throat	85.7	3.34 (1.57)	−0.95
6	Scratchy throat	83.3	2.84 (1.54)	−0.73
7	Cough	82.7	3.18 (1.68)	−0.67
8	Hoarseness	75.3	2.81 (1.67)	−0.74
9	Head congestion	83.7	2.80 (1.68)	−0.75
10	Chest congestion	59.0	2.03 (1.57)	−0.41
11	Feeling tired	98.0	3.37 (1.56)	−1.06
12	Headache	83.7	2.70 (1.71)	−0.73
13	Body aches	68.0	2.36 (1.48)	−0.61
14	Fever	60.7	2.19 (1.40)	−0.58
15	Think clearly	86.3	2.54 (1.42)	−0.76
16	Sleep well	91.7	2.99 (1.51)	−0.93
17	Breathe easily	84.0	3.10 (1.65)	−0.83
18	Walk/climb stairs	84.3	2.49 (1.52)	−0.73
19	Accomplish daily activities	81.0	2.24 (1.41)	−0.67
20	Work outside the home	81.7	2.60 (1.57)	−0.73
21	Work inside the home	78.3	2.22 (1.44)	−0.61
22	Interact with others	79.3	2.40 (1.45)	−0.67
23	Live your personal life	77.7	2.29 (1.32)	−0.63

The last item in the WURSS-24-C which assesses global change (change since yesterday) was not included.
Frequency: Scored above zero at least once in first 7 d of monitoring.
Severity: mean ± SD; averaged over the first 3 d; calculated only for those with symptom present all 3 d. To weight each person's responses equally, data were first averaged within-person-over-time, then averaged among participants.
MID: Minimal important difference: mean day-to-day change for those indicating minimal improvement.

Discussion

This study confirmed that the Chinese version of WURSS-24-C demonstrated highly acceptable reliability and validity, similar to that found in validation studies of the original English language instrument [12,13].

Exploratory and confirmatory procedures found that there were four domains (activity and function, systemic symptom, nasal symptom and throat symptom) in WURSS-24-C., which was different from the original WURSS-21 with three domains (activity and function, nasal symptom and throat symptom) [13]. In our EFA results, two items were grouped in different domains. It might be explained by domains overlap. *Feeling tired* (item 11) can be regarded as either activity and function, or systemic symptom domain. Likewise, *Breath easily* (item 17) can be explained by either activity and function or nasal symptom. Similar conditions were found from the original WURSS studies as well [13]. One item can be grouped in different domains, which reflects one symptom can affect more than one aspect of functional life. In the WURSS-21 study, the following items *Head congestion*, *Chest congestion* and *Feeling tired* was in the nasal symptom domain, which might not be suitable. These three items, together with another influenza-like item

(*Headache*, *Body aches* and *Fever)* were included in the new domain of WURSS-24-C as Systemic symptoms (Figure 1). CFA then further indicated that individual items of the WURSS-24-C, agreed well with this 4-dimensional structure model, as fit indices met criteria suggested by Hu and Bentler [29].

As a measure of the reliability of a scale, Cronbach alpha has been widely used since it was developed by Lee Cronbach [30]. According to Tavakol and Dennick [31] and Vogelzang [32], the alpha values of 0.8–0.9 are excellent, but if a coefficient alpha greater than 0.9, it may suggest redundancies and show that the test length should be shortened. We tested the four domains of the WURSS-24-C, and the Cronbach alpha perfectly ranges from 0.849 to 0.882, except for the "activity and function", which is up to 0.943. Compared with the original version of WURSS-21, for which the value is 0.961[13], we believe this 9-item dimension of WURSS-21 might be further reduced or adjusted. We further tested the reliability of the whole WURSS-24-C and found the value was over 0.90. This result suggested that the WURSS-24 might allow leeway for a shorter version, such as the WURSS-11developed by English-language WURSS researchers [33].

Convergent validity was evaluated by the Pearson correlation coefficient. According to Guyatt et al. [28], correlations ranging from 0.50 to 0.75 are moderate to good; and values greater than 0.75 are considered good to excellent. The WURSS-24-C yielded correlation coefficient was over 0.75 when compared with SF-8 physical. The association was stronger than that between the WURSS-24-C with SF-8 mental, for which the coefficient was −0.721. The results were similar to that of the original version of the WURSS-21.

MID refers to the smallest difference in a score that is recognized as worthwhile or important [27]. For clinicians, MID could be used to determine meaningful clinical change in patients. Our results estimated MID for individual items of the WURSS-24-C, which could be considered clinically relevant in patients with URTIs. Compared with the data of the original WURSS-21, the instruments yielded similar MIDs indices, which indicated the WURSS-24-C could be sensitively reflected changes in items.

Strengths of this study include its large sample size, careful translation of WURSS-24-C using three-step linguistic validation procedure, and in-depth statistical analysis. Our study should, nonetheless, be interpreted along with several limitations. First, all participants were living in the southern part of China with Mandarin as the official language, and were mostly young females. Although only a small and unknown

proportion of them came from other parts of China, extrapolation of our results to other regions, or age groups, should be done with caution. Second, the test–retest reliability and other reliability and validity assessments were not evaluated in this study. Third, the same dataset was used to do EFA and CFA. The result of EFA required to be confirmed in other Chinese population in future study. Therefore, participants from a broader range of age groups from other parts of China and additional instruments as for the evaluation might be needed in future studies.

Conclusion

In summary, the 24-item Chinese version of WURSS (WURSS-24-C) seems to be a reliable and valid instrument applicable to Chinese URTIs patients for assessment of HRQoL. With multi-language version of WURSS-24 available, the WURSS-24-C could thus be incorporated into future multinational research.

Acknowledgements

The authors thank the Wisconsin Alumni Research Foundation, University of Wisconsin-Madison, for the permission to use the WURSS.

Ethics approval and consent to participate

The study protocol was approved by the ethical committees of Guangdong provincial hospital of traditional Chinese medicine, Guangzhou, China (GPHCM; B2016-090-02). All methods involved in this study were performed in accordance with the relevant guidelines and regulations. Written informed consent was provided by all participants.

Author contributions

All authors have read and approved the final version of the manuscript. YW contributed to the design, supervised data collection and analysis, and drafted the manuscript. ZH contributed to the design, conducted statistical analysis. SC conducted data collection, entered, cleaned and analysed data. YL contributed to the design, coordinated data collection. FL coordinated data collection.BB contributed to the research idea and give advice for statistical analysis and results interpretation. ZZ contributed to the design, conducted statistical analysis. GS contributed to the design, supervised data collection and analysis. CSL was responsible for supervision or mentorship.

Disclosure statement

The authors declare that they have no competing interests.

Funding

ORCID

Cecilia Stålsby Lundborg 🆔 http://orcid.org/0000-0001-6525-1861

Data availability statement

The datasets generated or analysed during this study are not publicly available due to limitations of ethical approval involving the patient data and anonymity but are available from the corresponding author on reasonable request.

References

[1] Mylotte JM. Epidemiology of respiratory infection. Pathy's Princ Pract Geriatr Med. 2012;1:549–554.

[2] Allan GM, Arroll B. Prevention and treatment of the common cold: making sense of the evidence. CMAJ. 2014;186(3):190–199.

[3] Hurst JR, Donaldson GC, Wilkinson TMA, et al. Epidemiological relationships between the common cold and exacerbation frequency in COPD. Eur Respir J. 2005;26(5):846–852.

[4] Farmakis D, Parissis J, Lekakis J, et al. Acute heart failure: epidemiology, risk factors, and prevention. Rev Esp Cardiol (Engl Ed). 2015;68(3):245–248.

[5] Su G, Xu H, Marrone G, et al. Chronic kidney disease is associated with poorer in-hospital outcomes in patients hospitalized with infections: electronic record analysis from China. Sci Rep. 2017;7(1):11530.

[6] Fletcher-Lartey S, Yee M, Gaarslev C, et al. Why do general practitioners prescribe antibiotics for upper respiratory tract infections to meet patient expectations: a mixed methods study. BMJ Open. 2016;6(10): e012244.

[7] Dal Negro RW, Zanasi A, Turco P, et al. Influenza and influenza-like syndromes: the subjects' beliefs, the attitude to prevention and treatment, and the impact

in Italian general population. Multidiscip Respir Med. 2018;13:7.

[8] Feng L, Shay DK, Jiang Y, et al. Influenza-associated mortality in temperate and subtropical Chinese cities, 2003–2008. Bull World Health Organ. 2012;90(4): 279–288B.

[9] Sacristán JA. Patient-centered medicine and patient-oriented research: improving health outcomes for individual patients. BMC Med Inform Decis Mak. 2013; 13:6.

[10] Aubin D, Hebert M, Eurich D. The importance of measuring the impact of patient-oriented research. CMAJ. 2019;191(31):E860–E864.

[11] Barrett B, Brown R, Voland R, et al. Relations among questionnaire and laboratory measures of rhinovirus infection. Eur Respir J. 2006;28(2):358–363.

[12] Barrett B, Brown R, Mundt M, et al. The Wisconsin upper respiratory symptom survey is responsive, reliable, and valid. J Clin Epidemiol. 2005;58(6):609–617.

[13] Barrett B, Brown RL, Mundt MP, et al. Validation of a short form Wisconsin upper respiratory symptom survey (WURSS-21). Health Qual. Life Outcomes. 2009;7: 76.

[14] Barrett B. K. L. Wisconsin upper respiratory symptom survey (WUR$SS); 2022. Available from: https://www.fammed.wisc.edu/wurss/

[15] Yang SY, Kang W, Yeo Y, et al. Reliability and validity of Wisconsin upper respiratory symptom survey, Korean version. J. Epidemiol. 2011;21(5):313–318.

[16] Zgierska A, Obasi CN, Brown R, et al. Randomized controlled trial of mindfulness meditation and exercise for the prevention of acute respiratory infection: possible mechanisms of action. Evidence-Based Complement Altern Med. 2013;2013:1–14.

[17] Jiwa M, Krejany CJ, Kanjo E, et al. Symptom profile of patients receiving antibiotics for upper respiratory tract infections in general practice: an observational study using smartphone technology. Fam Pract. 2019; 36:560–567.

[18] Statista. The world's most spoken languages; 2022. Available from: https://www.statista.com/chart/12868/the-worlds-most-spoken-languages/

[19] Mapi. Linguistic validation of a patient reported outcomes measure. Lyon, France: Mapi Research Institute; 2005.

[20] Jackson GG, Dowling HF, Anderson TO, et al. T.M.: susceptibility and immunity to common upper respiratory viral infections—the common cold. Ann Intern Med. 1960;53:719–738.

[21] Jackson G, Dowling H, Muldoon RL. Acute respiratory diseases of viral etiology. VII. Present concepts of the common cold. Am J Public Health Nations Health. 1962;52:940–945.

[22] Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract. 2011;17(2):268–274.

[23] Roberts B, Browne J, Ocaka KF, et al. The reliability and validity of the SF-8 with a conflict-affected population in Northern Uganda. Health Qual Life Outcomes. 2008;6:108–110.

[24] Ware JE, 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.

[25] Norris M, Lecavalier L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. J Autism Dev Disord. 2010;40(1): 8–20.

[26] Mulaik SA. Confirmatory factor analysis. Handbook of multivariate experimental psychology. Boston (MA): Springer US; 1988. p. 259–288.

[27] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4): 407–415.

[28] Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis. 1987;40(2):171–178.

[29] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model. 1999;6(1): 1–55.

[30] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297–334.

[31] Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011;2:53–55.

[32] Vogelzang JL. Health measurement scales: a practical guide to their development and use. J Nutr. Educ Behav 2015;47(5):484.e1.

[33] Obasi CN, Brown RL, Barrett BP. Item reduction of the Wisconsin upper respiratory symptom survey (WURSS-21) leads to the WURSS-11. Qual Life Res. 2014;23(4): 1293–1298.