ORIGINAL ARTICLE

Clinical and Psychometric Validation of the Quality of Life Assessment System for Advanced Gastric Cancer Based on Traditional Chinese Medicine[*]

QUAN Peng (全　鹏)[1], ZHENG Pei-yong (郑培永)[2], YOU Sheng-fu (尤圣富)[2], HUA Yong-qiang (花永强)[2], SONG Yi (宋　毅)[2], LIU Tao (柳　涛)[2], WAN Chong-hua (万崇华)[1], and LU Jin-gen (陆金根)[2]

ABSTRACT **Objective**: To establish questionnaire scaling and reliability and examine the clinical and psychometric validity of the quality of life assessment based on Traditional Chinese Medicine for advanced gastric cancer (QLASTCM-Ga). **Methods**: The QLASTCM-Ga was developed based on programmed decision procedures with multiple nominal and focus group discussions, in-depth interview, pretesting and quantitative statistical procedures. The questionnaire was administered to 240 patients diagnosed with advanced gastric cancer before and after treatment. Structured group methods were employed to establish a general and a specific module respectively. The psychometric properties of the scale were evaluated with respect to validity, reliability and responsiveness. **Results**: The three identified scales of the QLASTCM-Ga and the total score demonstrated good psychometric properties. Test-retest reliability of the total scale and all domains ranged from 0.90 to 0.94, and internal consistency ranged from 0.86 to 0.93. Correlation and factor analysis demonstrated good construct validity. Significant difference in the subscales and the total score were found among groups differing in traditional Chinese medicine syndrome, supporting the clinical sensitivity of the QLASTCM-Ga. Statistically significant changes were found for each scale and the total score. Responsiveness was also good. **Conclusions**: The QLASTCM-Ga demonstrates good psychometric and clinical validity to assess quality of life in patients with advanced gastric cancer undergoing traditional Chinese medicine therapy. This study is an important first step for future research in this area.

KEYWORDS quality of life, advanced gastric cancer, traditional Chinese medicine

Gastric cancer is common in developing countries, and half of the world total occurs in Eastern Asia (mainly in China).[1] Stomach cancer incidence and death rate are at least twice as high in Chinese compared with whites.[2] Although treatment of gastric cancer has been greatly improved, the prognosis is not optimistic, and gastric cancer continues to be the second leading cause of cancer death worldwide.[3] The most common complaints of gastric cancer patients are dysphagia, eating restrictions, reflux, and abdominal pain.[4] How to improve quality of life (QOL) has become the focus of gastric cancer research. Clinical researchers are choosing measures of QOL as primary and secondary outcomes in clinical trials. Measurement of QOL has developed over the past decade and there are several questionnaires designed for patients with gastric cancer, such as the Functional Assessment of Cancer Therapy–Gastric Cancer (FACT-Ga),[5] and European Organization for Research and Treatment of Cancer (EORTC) QLQ-STO52 (The QLQ-STO 22 in conjunction with the QLQ-C30, the questionnaire module to assess health-related QOL in gastric cancer).[6-8] Although some studies have yielded useful clinical data, none of these questionnaires have been sufficiently powered to detect clinically significant QOL differences between groups of patients undergoing traditional Chinese medicine (TCM) therapy.

Culture contributes to QOL. There is an increasing need to measure QOL from a TCM perspective. In

TCM, the metaphoric views of the human body based on observations of nature are fully articulated in the theory of yin-yang. Within TCM philosophy, cancer results from a disturbance of yin-yang balance, and is a group of syndromes in which there is disharmony in the body–spirit–environment network.[9] In TCM view, some environmental factors such as climatic condition and geographical locality are responsible for illness.

To date, there is currently no validated measure to quantify QOL among advanced gastric cancer patients from TCM perspective. It is necessary to develop a psychometrically sound TCM measurement tool to quantify advanced gastric cancer patients' QOL. First of all, this newly developed scale would cover some deficiency of the existed ones, such as the domain related to the adaptation of the human body to natural environment, which is not included in other questionnaires. Secondly, it would be useful for both clinical and research purposes, especially for TCM therapy, for instance, to measure the difference in QOL scores between the initial and follow-up of different TCM syndromes. The aim of this study was to develop a disease-specific health status instrument to measure patients' QOL, and to test the psychometric properties and clinical validity in a large sample size.

METHODS

The formal QOL based on traditional Chinese medicine (QLASTCM, 43 items) is composed of two basic elements: (1) a core questionnaire, QLASTCM-general module (QLASTCM-GM, 34 items) which can be used with all types of cancer patients, comparing QOL across various disease states; (2) an additional disease-specific questionnaire for advanced gastric cancer (QLASTCM-Ga, 9 items), which is a supplementary module to the general module, and designed to focus on domains, characteristics, and complaints most relevant to advanced gastric cancer. The general module is supplemented by the specific module to increase sensitivity and specificity.

Establishment of the General Module

Based on a literature review of TCM, the GM was designed to measure QOL using two subscales: "correspondence between man and universe" and "unity of the body and spirit". The first subscale (correspondence between man and universe, 23 items) assessed the adaptation of the human body to natural environment. The second subscale (unity of the body and spirit, 11 items) was intended to measure whether the body is an organic whole. An overview of the theoretical construct of the QLASTCM-GM is shown in Figure 1.

The QLASTCM-GM was developed based on interviews with Chinese cancer patients and TCM professionals to generate 278 potentially candidate questions and to create a conceptual framework for how clinical manifestations of advanced gastric cancer impact the lives of patients. To establish the validity, reliability, and responsiveness, 625 cancer patients were enrolled in pilot-test. To assess experts' perceptions of the importance of each potential item, the item pool was also administered to 50 clinical experts with a rating questionnaire that provided 5 responses, ranging from "not important"[1] to "extremely important".[5] An open-ended response item was also included so that experts could add issues that were not included on the original list. The potential number of questions was 34 after distilling (T14, T25–27, T29–31, T42 were deleted).

Establishment of the Specific Module

The development of specific module was similar to general module, 27 items about a few important categories such as physical symptoms, side effects, psychology and social factors were proposed. After interviewing, importance analysis, and several rounds of expert group discussions, 11 items were selected and coded as W1 to W11. After that, 316 advanced gastric cancer patients were enrolled for the pilot-test. The variation coefficient, correlation coefficient, cluster analysis and the Cronbach coefficient method were used to screen each potential item. W6 and W11 were deleted (Electronic Supplementary Material 1).

Validation of the QLASTCM-Ga
Patients

We enrolled 240 (150 males) advanced gastric cancer patients in the formal survey. The mean age was 59.3 ± 11.7 years with a range of 27–92 years. Among them, 116 patients had junior middle school education, 59 had senior middle school education, 39 had community college education, and 24 had 4-year college education. Patients were prospectively registered before treatment and were eligible to participate if they had a diagnosis of advanced gastric cancer. The study population was limited to patients who were able to read and understand the questionnaires at any stages and treatments. The
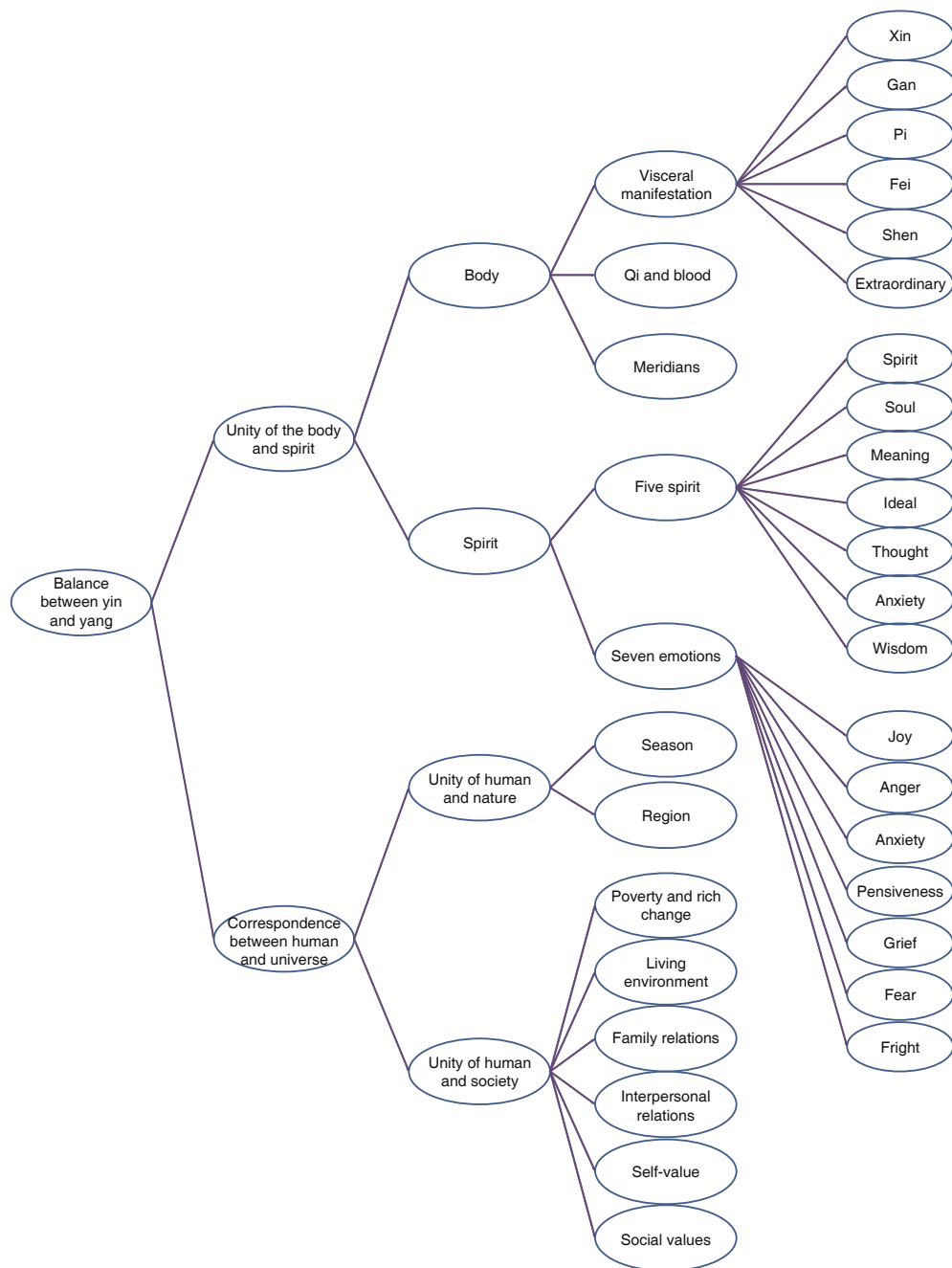
Figure 1.   Theoretical Construct of the QLASTCM-GM

participating investigators explained the trial and the scale to the patients. Written informed consent was obtained. Investigators obtained ethical committee approval to be eligible to recruit patients into this study. Compliance rates were 100%, questionnaires were well accepted and less than 4% of items had missing data.

Data Collection and Scoring

Each patient was asked to answer the questionnaires at the time of admission to the hospital by themselves. After giving a brief introduction

and explanation, the full version of QLASTCM-Ga (43 items) was administered to patients. In order to get test-retest reliability, 46 randomly selected patients received a retest 2–3 days later. To evaluate responsiveness of the questionnaire, each patient received four longitudinal measures (before starting therapy, 1 week, 3 months, and 6 months later). Answers were checked immediately each time by the investigators in order to ensure its integrality. If missing values were found, the questionnaire would be returned to the patients to fill in the missing item.

As a collateral measure, the Chinese version of EORTC QLQ-STO52 was also distributed to patients at the same time. At the baseline assessment, patients were asked to complete a short debriefing questionnaire covering questions about the need for help in completing the questionnaires and querying if any of the items appeared confusing, difficult to answer or upsetting. Sociodemographic data were also recorded at the baseline assessment.

The scoring method was similar to EORTC QLQ-STO52 consisting of 5-point Likert responses. The anchors for each item were modified according to the wording of the item (e.g., from never to always). The positively stated items were directly scored from 1 to 5. The negatively stated items were scored in reverse. Scores in each domain were obtained by adding the within-domain item scores, and the total score was obtained by adding scale scores. Twelve items in general module (T18, T19, T28, T32, T33, T34, T35, T36, T37, T38, T39, and T40) were directly scored, and the other items were scored in reverse. The raw scores of 1 to 5 were transformed to a 0 to 100 scale, where a score of 0 indicated the most severe symptoms and a score of 100 indicated no limitation. Higher scores indicated better QOL status.

Statistical Analysis for Psychometrics

Research methods for content validity included: (a) a thorough review of the literature to construct an item pool and (b) Delphi expert panel consensus. We conducted a series of analyses to generate an initial version of the instrument which was then used in the validation, reliability and responsiveness testing reported in this manuscript. We began with a content specification and item generation phase, followed by a process of item reduction and refinement, in which the instrument was judged by 57 experts.

Construct validity was examined using the correlation between each item and its designated scale as well as correlations between each item and the other scales. A scaling success for an item would be seen when the correlation between an item and its own scale was significantly higher than its correlation with any other scale.

To test convergent and divergent validity, Pearson's correlation coefficient was computed among scores on the QLASTCM-Ga and the EORTC QLQ-STO52. We hypothesized that the QLASTCM-Ga domains that assess "unity of the body and spirit" would correlate strongly with the EORTC QLQ-STO52 physical and emotional functioning domains, but poorly with the social functioning domain.

Known-group comparisons were examined by exploratory analysis of the clinical validity of the QLASTCM-Ga using the method of known-group comparison i.e., to explore the extent to which the questionnaire scores are able to discriminate between subgroups of patients differing in terms of their TCM syndromes. As per clinician's assessments, known-groups used for this comparison were five clinical categorizations: "weakness of the Pi (Spleen) and Wei (Stomach) syndrome", "phlegm-dampness syndrome", "qi stagnation and blood stasis syndrome", "deficiency of both qi and blood" and others.

Internal consistency was examined using reliability coefficients (Cronbach α coefficients), which was calculated on the data from the first measure. The interclass correlation coefficient (ICC) was used to assess test-retest reliability, which was calculated from 46 patients' data collected from the second test in the next day.

Standardized response means (SRM) was used to assess clinically meaningful changes. Totally, 240 patients were randomly selected for retest to evaluate the responsiveness after 3 month of treatment. SRM was calculated by using the paired t-test.

RESULTS

Content Validity

The item pool was deemed to reflect the opinion of World Health Organization regarding the connotation of QOL.[10] The screening of items was also strictly programmed to achieve good content validity.

Construct Validity

Correlations between each item and its designated scale were all significant at $P<0.01$. Table 2 shows correlations between each item and its designated scale in bold type as well as correlations between each item and the other scales in normal type (unity of the body and spirit: T1–13, T15–24; correspondence between man and universe: T28, T32–41; specific module: W1–5, W7–10). The item-to-scale correlation of each item was high for the designated scale and weak for any other

scale. T41, which showed high correlation with "unity of the body and spirit" domain, is an exception. It hinted that T41 was included in the wrong domain, or its meaning would lead to misunderstanding. Finally, for all items, the item-to-total score correlation was higher than all item-to-scale correlations except the designated scale. This suggests distinct separation of scales (Electronic Supplementary Material 2).

Convergent and Divergent Validity

As expected, results in Table 3 show that QLASTCM-Ga "unity of the body and spirit" domain had higher correlations to the EORTC QLQ-STO52 physical (r=0.79) and emotional (r=0.77) functioning domains as compared with the social functioning (r=0.68) domain. Table 3 further confirms the direction of correlation as expected. These results indicated that convergent and divergent validity was high. Conversely, the QLASTCM-Ga "correspondence between man and universe" domain had lower correlations to all EORTC QLQ-STO52 domains. This provided evidence that "correspondence between man and universe" domain was a unique domain which reflected traditional Chinese culture (Electronic Supplementary Material 3).

Clinical Validity

Scales distinguished between clinically distinct groups of patients. Patients in clinically distinct groups reported significant differences in baseline QOL scores in several scales (Table 1). Differences were statistically significant in the "correspondence between man and universe", specific module, GM and total score (P<0.05). The "unity of the body and spirit" scale did not demonstrate statistical differences (P=0.065).

Reliability

Cronbach α coefficients were high (> 0.80). All scales showed good reproducibility with interclass correlations above 0.90 (Table 2).

Responsiveness

After 3 months treatment, significant differences were reported in "unity of the body and spirit", GM, and total score (Table 3). There was no statistical difference in specific module and "correspondence between man and universe", which might attribute to the slow change of the clinical parameters and social factors. The SRM of the EORTC QLQ-STO52 was not statistically significant in all domains.

DISCUSSION

This study describes the development and validation of QLASTCM-Ga. The QLASTCM-Ga was created partly in response to feedback from TCM professionals who felt that existing measures were not suitable for Chinese patients using TCM cancer therapy. QOL is a subjective concept, which was often evaluated through personal feelings or one's own evaluation. It

Table 1. QOL Scores for Gastric Cancer Patients of Different TCM Syndromes after Treatment (Score, $\bar{x} \pm s$)

TCM syndrome	Case	Unity of the body and spirit	Correspondence between man and universe	Specific module	General module	Total score
Weakness of Pi and Wei	154	77.04 ± 14.11	59.33 ± 17.78	83.08 ± 16.18	71.31 ± 13.12	73.77 ± 12.93
Phlegm-dampness	33	72.63 ± 15.19	56.68 ± 13.47	76.35 ± 13.36	67.47 ± 13.46	69.33 ± 13.11
Qi stagnation and blood stasis	17	77.49 ± 13.98	55.21 ± 19.08	82.84 ± 17.10	70.29 ± 13.65	72.91 ± 13.71
Deficiency of both qi and blood	28	71.04 ± 17.45	52.68 ± 14.87	73.51 ± 22.03	65.10 ± 14.47	66.86 ± 15.53
Others	8	65.68 ± 14.25	42.53 ± 13.04	78.17 ± 22.65	58.19 ± 10.16	62.38 ± 12.47
F		2.25	2.48	2.69	2.95	2.99
P		0.065	0.045	0.032	0.021	0.020

Table 2. Reliability of QLASTCM-Ga

Domain/facet	Correlation coefficient r (n=46)	ICC (n=46)	Cronbach α coefficient (n=240)
Unity of the body and spirit	0.91	0.91 (0.84–0.95)	0.93
Correspondence between man and universe	0.92	0.92 (0.85–0.95)	0.86
Specific module	0.90	0.90 (0.83–0.95)	0.89
General module	0.94	0.94 (0.89–0.96)	0.93
Total score	0.94	0.94 (0.89–0.97)	–

Table 3.   Responsiveness of QLASTCM-Ga and QLQ-STO52 (n=240, $\bar{x} \pm s$)

Domain/facet	Baseline	After 3 months' treatment	t	P	SRM
QLASTCM-Ga					
Unity of the body and spirit	75.32 ± 14.96	72.02 ± 8.42	3.27	0.001	0.21
Correspondence between man and universe	57.50 ± 17.03	55.66 ± 12.77	1.65	0.100	0.11
Specific module	80.71 ± 17.28	81.09 ± 6.64	−0.35	0.725	0.02
General module	69.56 ± 13.64	66.73 ± 6.81	3.30	0.001	0.22
Total	71.89 ± 13.67	69.73 ± 6.32	2.50	0.013	0.16
QLQ-C30					
Physical functioning	77.96 ± 18.46	79.97 ± 10.05	−1.75	0.081	0.12
Role functioning	75.22 ± 24.67	72.05 ± 14.33	2.09	0.038	0.14
Emotional functioning	80.21 ± 17.77	72.88 ± 12.21	6.43	0.000	0.42
Cognitive functioning	82.40 ± 16.28	76.44 ± 11.52	4.94	0.000	0.32
Social functioning	67.89 ± 25.00	68.32 ± 10.90	−0.27	0.785	0.02
Global health status/QOL	62.21 ± 22.23	63.61 ± 11.07	−0.99	0.324	0.06
Fatigue	29.65 ± 20.67	32.14 ± 8.96	−1.88	0.061	0.12
Nausea and vomiting	15.09 ± 22.31	12.93 ± 13.89	1.42	0.156	0.09
Pain	18.97 ± 22.03	26.72 ± 13.05	−5.24	0.000	0.34
Dyspnoea	10.49 ± 17.27	2.01 ± 7.95	7.60	0.000	0.50
Insomnia	22.27 ± 24.95	24.14 ± 16.46	−1.18	0.238	0.08
Appetite loss	24.57 ± 26.59	27.16 ± 15.67	−1.60	0.112	0.10
Constipation	15.80 ± 21.02	12.93 ± 16.28	1.88	0.061	0.12
Diarrhoea	8.91 ± 16.91	2.16 ± 8.21	5.98	0.000	0.39
Financial difficulties	33.91 ± 27.21	33.05 ± 11.60	0.50	0.618	0.03
QLQ-STO22					
Dysphagia	11.11 ± 16.02	9.96 ± 7.58	1.05	0.293	0.07
Stomach pain	21.19 ± 21.92	25.75 ± 12.18	−3.28	0.001	0.22
Reflux	17.15 ± 18.31	12.55 ± 7.95	3.88	0.000	0.25
Eating restriction	16.56 ± 16.24	20.22 ± 10.03	−3.41	0.001	0.22
Anxiety	26.87 ± 21.19	31.47 ± 9.20	−3.41	0.001	0.22
Dry mouth	22.99 ± 21.89	28.02 ± 12.99	−3.69	0.000	0.24
Taste change	14.51 ± 20.92	11.93 ± 17.17	1.51	0.131	0.10
Body image	28.30 ± 22.76	31.18 ± 10.75	−1.99	0.047	0.13
Hair loss	13.51 ± 23.61	4.45 ± 12.95	6.11	0.000	0.40

is well known that culture contribute to QOL. Any QOL measures will only apply to a defined community. The basis of TCM diagnosis includes palpitation, upset, choler, lethargy, night sweating, and xerostomia, which is not included in Western medicine. The focus of TCM is on the patient rather than disease and fundamentally aims to promote health while enhancing QOL, with therapeutic strategies for treatment of specific diseases in a holistic fashion.[11]

The validity of the QLASTCM-Ga was evaluated by content validity, construct validity and criterion-related validity. The usual methods to evaluate content validity were correlation coefficient, factor analysis and SEM.[12] As the number of items was not sufficient, SEM was not used in this study.[13] Content validity was evaluated using Delphi method. Based on results from correlation analysis and exploratory factor analysis, the construct validity was good. When no clear gold standard exists for quantifying a property, the most assuring method to establish the validity of the QLASTCM-Ga is convergent validity, in which the new measure is shown to be correlated with other measures that are believed to quantify the same concept. Such correlations are considered to be high when the correlation coefficients are ⩾0.4. Conversely,

divergent validity is demonstrated when domain items that are thought to measure different concepts have low correlations ($r<0.4$). Correlation between the same or similar domain of two questionnaires was higher than that from the different domains. The convergent validity and divergent validity were satisfied in this study.

Testing demonstrated that QLASTCM-Ga was reliable and sensitive to changes in QOL as well as being able to discriminate between clinically distinct groups of patients with advanced gastric cancer. Physicians' clinical assessment of their patients' syndrome was explicitly collected on the case report forms. Patients with a particular TCM syndrome were compared. The results showed a good clinical validity.

Internal consistency, or reliability, examines the consistency of items within a scale and quantifies the degree to which each item is measuring aspects of the same underlying domain.[14] In this analysis, the internal consistency of the QLASTCM-Ga and its subscales was examined using Cronbach α coefficient, in which a value of 0.90 or higher is excellent and 0.80 or higher is sufficient.[15] The test-retest reliability was evaluated by ICC, which measures how stable responses are over time ($\geqslant 0.7$), which was calculated from 46 patients' data from the second test during the following day, in which patients' QOL status would be less likely to change. The internal consistency reliability and test-retest reliability of the QLASTCM-Ga both exceeded 0.8. It was concluded that the QLASTCM-Ga possesses good reliability and stability.

The responsiveness of an instrument refers to its ability to detect clinically meaningful changes in a patient's QOL status over time. We selected SRM as a measure of the change in instruments' scores within each group and calculated it for all questionnaires. SRM is defined as mean score change divided by the standard deviation of that score change. As a benchmark for the interpretation of SRM, Cohen describes an effect size of 0.20 as small, 0.50 as medium, and 0.80 as large.[16] A series of paired t-tests were conducted to compare changes in scores for all questionnaires. The QLASTCM-Ga was considered to have good responsiveness.

The QLASTCM-Ga shares some common characteristics with the EORTC QLQ-STO52. For example, specific module and general modules were developed at the same time. The items in both measures were rated using a Likert scoring system. However, there were some distinguishing characteristics in the QLASTCM-Ga. From the view of structure, the EORTC QLQ-C30 was constituted by 9 scales and 6 single items.[17] There are a large amount of single items on the reaction symptoms in the QLQ-C30. The QLASTCM-Ga general module was constituted by 34 items from only 2 scales. Moreover, there are 5 scales and 3 single items in the EORTC QLQ-STO22. Only one dimension needs to be analyzed in the QLASTCM-Ga specific module. In other words, scoring for the QLASTCM-Ga is easier. Specifically, items in the QLASTCM-Ga capture significant Chinese cultural characteristics, such as "I feel sore and weak in my waist and knees" (T7), "I'm happy with the surrounding natural environment" (T28), "I am happy with the climate of residence" (T32), "I feel scorching hot in the Wei" (W5), and "I have a dragging sensation in anus" (W11). These items embody "unity of the body and spirit" and "correspondence between man and universe", which are specially focused on in Chinese culture and are not included in other questionnaires.

Although this study successfully developed a new QOL measure and subsequently validated this instrument, there are some limitations in this study. One limitation is that the responsiveness of the QLASTCM-Ga is not high enough. It might be that the scale was too large (there were 23 items in "unity of body and spirit"), and that is counteracts some minor facets under the main domain. The minor components should be subdivided under main domains. Another limitation is that our survey excluded severe patients. Therefore, the results may not be applicable to these special groups. There were also a relatively higher proportion of some TCM syndromes in our sample. For instance, one out of five syndromes ("weakness of the Pi and Wei syndrome") was noted in 64% patients, while another four syndromes were only noted in 36% of the patients. Fortunately the bias is expected to be consistent for all subjects and should not affect our conclusions. Future research is being planned concerning the interpretation and application of the QLASTCM-Ga in different TCM syndrome samples. Our goals are to establish separate norms for different TCM syndromes in order to determine clinically meaningful changes in QLASTCM-Ga scores.

Taken together, the QLASTCM-Ga demonstrated

good psychometric properties. The QLASTCM-Ga has significant Chinese cultural characteristic, which was informative in helping to assess primary and secondary outcomes from TCM perspective. In our opinion, the QLASTCM-Ga would make a useful addition to the assessment protocol of clinical trials for advanced gastric cancer.

Conflict of Interest

No conflict of interest declared.

Author Contribution

Quan P and Zheng PY contributed equally to ghis article. Conception and design: Zheng PY; analysis and interpretation: Quan P; data collection: You SF, Hua YQ, Song Y, Liu T; writing the article: Quan P; critical revision of the article: Wan CH; final approval of the article: Lu JG; statistical analysis: Zheng PY; obtaining funding: Lu JG; overall responsibility: Wan CH.

Acknowledgements

Electronic supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s11655-016-2465-6 and is accessible for authorized users.

REFERENCES

1. WHO. Are the number of cancer cases increasing or decreasing in the world? 2008: WHO Online Q&A.

2. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics. Cancer J Clin 2011;61:212-236.

3. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-2917.

4. Zhao L, Shen ZX, Luo HS, Yu JP. Clinical investigation on coexisting of duodenal ulcer and gastric cancer in China. Int J Clin Pract 2005;59:1153-1156.

5. Garland SN, Pelletier G, Lawe A, Biagioni BJ, Easaw J, Eliasziw M, et al. Prospective evaluation of the reliability, validity, and minimally important difference of the functional assessment of cancer therapy-gastric (FACT-Ga) quality-of-life instrument. Cancer 2011;117:1302-1312.

6. Huang CC, Lien HH, Sung YC, Liu HT, Chie WC. Quality of life of patients with gastric cancer in Taiwan: validation and clinical application of the Taiwan Chinese version of the EORTC QLQ-C30 and EORTC QLQ-STO22. Psycho-Oncol 2007;16:945-949.

7. Kobayashi D, Kodera Y, Fujiwara M, Koike M, Nakayama G, Nakao A. Assessment of quality of life after gastrectomy using EORTC QLQ-C30 and STO22. World J Surgery 2011;35:357-364.

8. Blazeby J, Conroy T, Bottomley A, Vickery C, Arraras J, Sezer O, et al. Clinical and psychometric validation of a questionnaire module, the EORTC QLQ-STO 22, to assess quality of life in patients with gastric cancer. Eur J Cancer 2004;40:2260-2268.

9. Xu W, Towers AD, Li P, Collet JP. Traditional Chinese medicine in cancer care: perspectives and experiences of patients and professionals in China. Eur J Cancer Care 2006;15:397-403.

10. Kuyken W, Orley J, Power M, Herrman H, Schofield H, Murphy B. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995;41:1403-1409.

11. Patwardhan B, Warude D, Pushpangadan P, Bhatt N. Ayurveda and traditional Chinese medicine: a comparative overview. Evid-based Complem Alternat Med 2005;2:465-474.

12. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. Psychol Assess 1995;7:238-247.

13. Ullman JB, Bentler PM, eds. Structural equation modeling. Handbook of psychology. John Wiley & Sons, Inc. 2003.

14. Henson RK. Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. Measure Evaluat Counsel Devel 2001;34:177-189.

15. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997;314:572.

16. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-468.

17. Fayers PM, Aaronson NK, Bjordal K, Grønvold M, Curran D, Bottomley A. EORTC QLQ-C30 scoring manual. European Organization for Research and Treatment of Cancer Brussels. 2001.