

Measuring Quality of Life in Chinese Cancer Patients: A New Version of the Functional Living Index for Cancer (Chinese)

Y B Cheung,¹PhD, G Y Ng,²BA, L C Wong,³MBBS, FRCR, W H Koo,⁴MBBS, MRCP, E H Tan,⁴MBBS, M Med, MRCP, M H Tay,⁵MBBS, MRCP, D Lim,⁵MBBS, MRCP, D Poon,⁵MBBS, MRCP, C Goh,⁶FAMS, MRCP, PhD, S B Tan,⁷PhD

Abstract

Introduction: Since its translation into Chinese, the Functional Living Index for Cancer (FLIC) has not been widely received due to some of its difficulties. We modified its visual analogue scale (VAS) to an ordered categorical scale and changed some of the wording in the instrument. This study examined the measurement properties of the modified FLIC. **Materials and Methods:** The modified version of FLIC and the Functional Assessment of Cancer Therapy (FACT-G Chinese version 4) were filled in by 140 patients recruited from the National Cancer Centre Singapore. The patients' FLIC scores were compared with their clinical characteristics to establish known-group validity. Convergent and divergent validity of FLIC were examined by correlation analysis with FACT-G and its sub-scales. Cronbach's alpha and relative efficiency were also examined. **Results:** FLIC and most of its sub-scales could indicate a clear and statistically significant difference of quality of life (QOL) according to patients' performance status and treatment status. FLIC strongly correlated with FACT-G. The Physical, Psychological, and Symptoms sub-scales of FLIC converged to and diverged from FACT-G sub-scales as conceptually expected. Cronbach's alpha indicated a satisfactory level of reliability. FLIC appeared to be more efficient than FACT-G, meaning that a smaller sample size will be required for FLIC than for FACT-G to achieve the same research purpose. **Conclusions:** The modified version of FLIC was found to have achieved satisfactory measurement properties. This is a user-friendly alternative to the original FLIC.

Ann Acad Med Singapore 2003; 32:376-80

Key words: Cancer, Chinese, Quality of life

Introduction

Health-related quality of life (QOL) is recognised as an important aspect of patient care. In oncology studies, it may stand as the primary end-point.^{1,2} Most QOL instruments are developed in English, although about one-fifth of the world's population is ethnic Chinese.³ Singapore is a multi-ethnic society. In 2000, 77% of its population were ethnic Chinese, 14% were Malays, and 8% were Indians (<http://www.singstat.gov.sg>). Chinese language is a major medium of communication, especially among the older population,⁴ who are at a higher risk of developing cancer.⁵ There is an obvious need for translating English QOL instruments into Chinese.

The Functional Living Index for Cancer (FLIC) is widely used in oncology research.⁶ It was translated into Chinese by one of the authors and others in Singapore, in collaboration with the FLIC developers.⁷ Nevertheless, utilisation of this Chinese version of FLIC has been limited. From our experience, some clinicians and researchers are discouraged by two factors. First, FLIC uses the visual analogue scale (VAS), which is felt to be more difficult to explain and be filled in by patients as compared to the more commonly used Likert or adjectival scales. This is especially so in the older or less educated patients. Furthermore, the VAS poses more difficulty for data entry, as it involves using a ruler to measure the location of the answers that are

¹ Biostatistician

² Research Coordinator

⁷ Senior Biostatistician

Division of Clinical Trials and Epidemiological Sciences

³ Registrar

Department of Therapeutic Radiology

⁴ Senior Consultant

⁵ Registrar

Department of Medical Oncology

⁶ Head

Department of Palliative Medicine

National Cancer Centre

Address for Reprints: Dr Cheung Yin-Bun, Division of Clinical Trials and Epidemiological Sciences, National Cancer Centre, 11 Hospital Drive, Singapore 169610.

E-mail: ctecyb@nccs.com.sg

in the form of vertical lines drawn by patients. Secondly, although the translation is correct word for word, some of the resulting translated sentences do not come across in a manner that agrees with common usage. For instance, in the translation of the item “Rate how often you feel discouraged about your life”, the Chinese version was faithful to the English version in that the word “rate” was retained. But the resulting sentence structure is not in a form that is commonly used in everyday Chinese. A Chinese equivalent that reads “How often do you feel discouraged about your life” would seem more appropriate.

In this study, we revised the two aspects of the Chinese version of FLIC. We examined the measurement properties of the new version. We also extended the work to compare its relative validity and relative efficiency with the Functional Assessment of Cancer Therapy – General (FACT-G Chinese version 4),⁸ another widely used instrument, but not translated in Singapore.

Materials and Methods

The VAS of the FLIC requires patients to draw a vertical line on a horizontal scale with equi-distance marks from 1 to 7. Written descriptions are given at the extreme ends of the scale. For instance, in the item “Rate how often you feel discouraged about your life”, the descriptors “Always” and “Never” are given below the numbers 1 and 7, respectively. In the new version, patients are asked to circle a number in the range from 1 to 7, with the same descriptors at the ends. Ten items of the English version begin with the word “rate”, such as the one given above. We removed the Chinese equivalent of the word from the items, so they read like “How often do you feel discouraged about your life” in the new Chinese version. A copy of the new questionnaire is available from the corresponding author.

Patients were recruited from the outpatient clinics of the National Cancer Centre Singapore from December 2001 to February 2002. Patients were screened by oncologists in the clinics according to the inclusion and exclusion criteria. The inclusion criteria were: able to understand written Chinese, aged 18 years or older, and physically fit for self-administering a questionnaire. The exclusion criteria were: evidence of brain metastasis, psychosis or severe depression. Eligible patients were referred to a research psychologist. The psychologist explained the study in detail, obtained informed consent, as well as delivered and collected the questionnaires. The respondents self-administered the questionnaires, but they could ask the research psychologist for clarification if they have a question. The study was approved by the Ethics Committee of the Centre.

Assuming the true Pearson’s correlation coefficient between the new version of FLIC and FACT-G was 0.75, a sample size of 110 subjects were required in order to

achieve a precision that the 95% confidence interval (CI) would have a total width of 0.2 (based on Fisher’s Z transformation).⁹ The actual sample size we obtained was larger because our secondary purpose not detailed in this manuscript involved factor analyses of the two instruments, which required a larger number of subjects.¹⁰ So the sample size here was larger than required for the present purpose.

Performance status of the patients was evaluated by the referring clinicians, using the Karnofsky performance score. They also classified the patients’ treatment status as currently under chemotherapy, radiotherapy, or neither. The questionnaire package began with a demographic section, followed by the FACT-G, and ended with the new Chinese version of FLIC. Convergent and divergent validity were assessed by correlation analysis between FLIC and FACT-G and between their sub-scales. For instance, a strong correlation between FLIC Physical sub-scale and FACT-G Physical sub-scale was taken as evidence of convergent validity; a weak correlation between FLIC Physical sub-scale and the other FACT-G sub-scales was taken as evidence of divergent validity. Known-group validity was assessed by analysis of variance (ANOVA) using QOL according to Karnofsky performance score and treatment status. For instance, we would expect patients under chemotherapy or radiotherapy to have a significantly lower QOL. Ratios of F-values in ANOVA were used to compare relative efficiency of the two instruments.¹⁰ A higher efficiency suggests that the measure is less subject to random variation, and a smaller sample size is required to detect a significant difference between different patient groups. The bootstrap method was used to test the hypothesis of equal relative efficiency.¹¹ Cronbach’s alpha was used to evaluate reliability of the FLIC and its 5 sub-scales as summarised by Ruckdeschel and Piantadosi.¹²

Results

A total of 152 patients were approached. Six refused to participate and another 6 were found unable to understand written Chinese. Among the 140 respondents, 84 returned the QOL questionnaires with no missing values, 47 with one missing value and 9 with two. The most common missing value was from the FACT-G item on satisfaction with sexual life. We applied the FACT-G scoring algorithm for missing values to both the FLIC and FACT-G items. That is, a missing value was replaced by the mean of the items in the same sub-scale if valid answers were given to at least half of the items in the same sub-scale. There was no missing value in the scale or sub-scale scores after using this algorithm. The total FLIC score was a summation of the 5 sub-scale scores.

Table I shows the descriptive characteristics of the patient sample. Table II examines the convergent and

divergent validity of FLIC and its sub-scales in terms of Pearson's correlation coefficient (r ; 95% confidence interval). FLIC was strongly correlated with FACT-G ($r = 0.71$; 0.62 to 0.78). FLIC Psychological sub-scale was strongly correlated with FACT-G Emotion (GE) ($r = 0.76$;

0.68 to 0.82), but not with the others (r 's between 0.26 to 0.53). FLIC Physical sub-scale was correlated with FACT-G Physical (GP) and FACT-G Functional (GF), the correlation coefficients being 0.70 (0.61 to 0.77) and 0.62 (0.51 to 0.71), respectively. FLIC Symptoms sub-scale correlated with GP ($r = 0.68$; 0.58 to 0.76) and, to a lesser extent, GF ($r = 0.41$; 0.29 to 0.53). These 3 sub-scales converged to and diverged from FACT-G sub-scales as conceptually expected. However, FLIC Family and FLIC Social sub-scales did not correlate well with FACT-G Social (GS) ($r = 0.16$ and 0.26, respectively).

Table II also shows the Cronbach's alpha values (α), as well as means and standard deviations (SD) of the FLIC and its sub-scales. FLIC ($\alpha = 0.92$) and its Psychological ($\alpha = 0.82$), Physical ($\alpha = 0.86$) and Family ($\alpha = 0.69$) sub-scales had adequate levels of reliability. The Symptoms and Social sub-scales only had a moderate alpha ($\alpha = 0.59$ and 0.57, respectively). As many researchers and Cronbach himself pointed out,^{10,13} the alpha value is partly dependent upon the number of items within a scale or sub-scale. It is therefore not fair to the sub-scales with a smaller number of items. The Spearman-Browne formula was used to adjust the alpha values so that all the measures were comparable as if all of them contained 7 items;^{10,13} the expected alpha values ranged from 0.73 to 0.86 for all of the FLIC total and sub-scales.

Table III shows findings on known-group validity in terms of the ANOVA results by the Karnofsky performance score. Apart from FLIC Social sub-scale, all the FLIC total and sub-scale scores clearly discriminated levels of performance status and were significant at 0.01 or 0.05 level. In contrast, although FACT-G and all sub-scales did show different mean levels of QOL by performance status, mean GS and GE scores were not significantly different between the groups ($P = 0.28$ and 0.30, respectively). Patients with the highest Karnofsky performance score had a mean FACT-G, mean GP and mean GF of at least 45% higher than that of patients with performance score 70 or below. In contrast, the corresponding figure for GS was

TABLE I: DESCRIPTIVE SUMMARY OF PATIENT PROFILE

Variable	Mean or frequency*	SD or %*
Age	53	11
Sex		
Female	77	55.0%
Male	63	45.0%
Education		
Primary incomplete	37	26.4%
Primary complete	44	31.4%
Secondary school	50	35.7%
Post-secondary	9	6.4%
Marital status		
Married	109	77.9%
Single	14	10.0%
Divorce/Separated	9	6.4%
Widowed	8	5.7%
Tumour type		
Breast	50	35.7%
Colorectal	42	30.0%
Stomach	11	7.9%
Lung	16	11.4%
Others	21	15.0%
Karnofsky score		
100	58	41.4%
90	61	43.6%
80	13	9.3%
70 or below	8	5.7%
Treatment status (receiving chemotherapy or RT)		
Yes	50	35.7%
No	90	64.3%

RT: radiotherapy; SD: standard deviation

* Mean and SD for continuous variable; frequency and percentage for categorical variables

TABLE II: CORRELATION AMONG FUNCTIONAL LIVING INDEX FOR CANCER SCALES AND FUNCTIONAL ASSESSMENT OF CANCER THERAPY – GENERAL SCALES; CRONBACH'S ALPHA AND DESCRIPTIVE SUMMARY OF FLIC SCALES

	FLIC-Psy	FLIC-Phy	FLIC-Sym	FLIC-Fam	FLIC-Soc	FACT-G	GP	GS	GE	GF	No. of Items	Alpha	Adjusted Alpha	Mean	SD
FLIC	0.87**	0.90**	0.72**	0.80**	0.61**	0.71**	0.70**	0.30**	0.63**	0.58**	22	0.92	0.78	120.7	21.4
FLIC-Psy	1.00	0.68**	0.45**	0.61**	0.56**	0.63**	0.53**	0.26**	0.76**	0.44**	7	0.82	0.82	37.2	7.5
FLIC-Phy		1.00	0.61**	0.60**	0.43**	0.69**	0.70**	0.31**	0.48**	0.62**	7	0.86	0.86	38.1	8.4
FLIC-Sym			1.00	0.62**	0.21*	0.49**	0.68**	0.15*	0.29**	0.41**	3	0.59	0.74	16.7	3.9
FLIC-Fam				1.00	0.48**	0.49**	0.50**	0.16*	0.44**	0.42**	3	0.69	0.82	16.8	3.8
FLIC-Soc					1.00	0.39**	0.25**	0.26**	0.39**	0.28**	2	0.57	0.73	11.8	2.5

* $P < 0.05$; ** $P < 0.01$

FLIC-Fam: FLIC Family; FLIC-Phy: FLIC Physical; FLIC-Psy: FLIC Psychological; FLIC-Soc: FLIC Social; FLIC-Sym: FLIC Symptoms; GE: FACT-G Emotion; GF: FACT-G Functional; GP: FACT-G Physical; GS: FACT-G Social; SD: standard deviation

TABLE III: MEAN QUALITY OF LIFE SCORES BY LEVEL OF KARNOFSKY PERFORMANCE SCORE, TESTED BY ANALYSIS OF VARIANCE

KPS	FLIC	FLIC-Psy	FLIC-Phy	FLIC-Sym	FLIC-Fam	FLIC-Soc	FACT-G	GP	GS	GE	GF
100	129.5	39.5	41.5	18.1	17.9	12.4	84.0	25.4	20.2	18.6	19.8
90	118.1	36.0	37.9	16.4	16.3	11.4	77.6	22.6	19.8	17.9	17.4
80	106.9	34.5	31.2	13.9	15.5	11.9	71.5	19.7	18.8	16.9	16.1
≤ 70	99.4	33.4	26.1	13.3	15.5	11.1	56.3	17.5	15.4	15.8	7.6
F value	9.5	3.8	15.3	8.1	3.0	1.9	8.7	13.1	1.3	1.2	10.1
P value	<0.01	0.01	<0.01	<0.01	0.04	0.13	<0.01	<0.01	0.28	0.30	<0.01

FLIC-Fam: FLIC Family; FLIC-Phy: FLIC Physical; FLIC-Psy: FLIC Psychological; FLIC-Soc: FLIC Social; FLIC-Sym: FLIC Symptoms; GE: FACT-G Emotion; GF: FACT-G Functional; GP: FACT-G Physical scale; GS: FACT-G Social; KPS: Karnofsky performance score

TABLE IV: MEAN QUALITY OF LIFE SCORES BY TREATMENT STATUS, TESTED BY ANALYSIS OF VARIANCE

Receiving chemotherapy or radiotherapy	FLIC	FLIC-Psy	FLIC-Phy	FLIC-Sym	FLIC-Fam	FLIC-Soc	FACT-G	GP	GF	GS	GE
Yes	112.9	35.9	35.5	14.8	15.5	11.1	76.1	21.6	16.7	19.8	18.0
No	125.0	37.8	39.6	17.8	17.6	12.3	79.8	24.1	18.3	19.5	18.0
F value	11.0	2.1	8.1	20.7	9.8	7.0	1.6	9.2	1.7	0.1	0.0
P value	<0.01	0.15	<0.01	<0.01	<0.01	<0.01	0.21	<0.01	0.19	0.80	0.99

FLIC-Fam: FLIC Family; FLIC-Phy: FLIC Physical; FLIC-Psy: FLIC Psychological; FLIC-Soc: FLIC Social; FLIC-Sym: FLIC Symptoms; GE: FACT-G Emotion; GF: FACT-G Functional; GP: FACT-G Physical scale; GS: FACT-G Social

31% and that for GE was only 18%, suggesting these 2 sub-scales were less discriminative. Relative efficiency of FLIC and FACT-G was estimated in terms of ratio of F-values. FLIC was not more significantly efficient than FACT-G in differentiating different levels of performance status, the relative efficiency being 1.09 ($P > 0.1$).

Table IV shows the ANOVA results by treatment status. All FLIC total and sub-scale scores indicated a lower level of QOL in the group undergoing chemotherapy or radiotherapy. All of them were found to be statistically significantly for the 2 treatment groups (each $P < 0.01$), except FLIC Psychological sub-scale ($P = 0.15$). In contrast, mean GS and GE scores were almost identical in the 2 groups. FACT-G and 3 of its sub-scales did not show statistically different levels of well-being between the two groups (each $P > 0.1$). Only mean GP was significant ($P < 0.01$). FLIC was more efficient than FACT-G in differentiating treatment status, the relative efficiency being 7.01 ($P < 0.01$; 95% CI = 1.28 to 231.04). Using the lower bound of the 95% CI as a conservative estimate of relative efficiency (i.e., 1.28), it indicates that if one uses FACT-G one needs a sample size of 28% larger than if one uses FLIC, other factors being held constant.

Conclusion

In the past, QOL was taken as a supplementary end-point. Now, its role as a primary end-point in oncology is recognised. As many Singaporeans and many of the world's population use the Chinese language, translating QOL measures into Chinese is important.

The VAS has often been used in medicine, especially in the measurement of pain and psychological states.^{14,15} Its proponents claim that it is simple and intuitive to use, and that it enumerates more precise answers and is more sensitive. However, empirical research has suggested that patients may not find it as simple and intuitive as some researchers think. The use of VAS has been associated with a larger amount of missing values, increased error and decreased sensitivity in some patient groups, such as older patients and patients suffering a reduction in cognitive function.^{13,15} Furthermore, outcomes measured by the VAS and an ordered categorical scale often correlated substantially.^{13,15} As such, there is little reason to prefer the VAS. In a comment on QOL instruments developed for research purposes, Higginson and Carr¹⁶ said: "For a measure to have clinical usefulness, ... it must also be simple, quick to complete, easy to score...". Our experience suggested that some researchers were reluctant to use the earlier Chinese version of FLIC because of the difficulties associated with its VAS. We decided to change it to an ordered categorical scale, keeping its original adjectival labels and the range of seven points. We have also made some minor changes in the wording of the questions.

FLIC and FACT-G were not administered in a randomised order. Some researchers suggested that the order of questionnaire components may cause a learning effect or context effect.¹³ That is, respondents learn from the earlier items and then answer the later items in a different way. However, FLIC and FACT-G have quite different characteristics. For example: FACT-G uses a 5-point scale

printed to the right of the questions, while FLIC uses a 7-point scale printed below each question. FACT-G items are explicitly arranged to form sub-scales, while FLIC items appear in a random fashion, etc. As such, responding to the FLIC should be a new experience to the patients, even though they have just filled out the FACT-G. It is unlikely that the order of the instruments could have an effect in the present case.

Correlation analyses demonstrated the convergent and divergent validity of the Psychological, Physical and Symptoms sub-scales of the FLIC. The total score of FLIC also correlated well with that of FACT-G. However, the Family sub-scale, and to a lesser extent, the Social sub-scale, of FLIC did not show convergence to the Social sub-scale of FACT-G. One problem is that the FACT-G sub-scale consists of both family and social elements. It includes 4 items related to family life, 2 related to friendship and 1 related to sex life. So, it would not be surprising that this sub-scale did not correlate well with the FLIC Social sub-scale. Furthermore, a careful examination of the items in the FLIC Family sub-scale would suggest that they were oriented to the impact of disease on family members, e.g. "The degree to which your illness has imposed a hardship on those closest to you". In contrast, the family items in FACT-G Social sub-scale were oriented to a patient's feeling about the family, e.g. "I get emotional support from my family". Again, the conceptual difference may explain the weak correlation.

FLIC demonstrated stronger known-group validity. With the exception of FLIC Social sub-scale, all FLIC and its sub-scales indicate a significant difference between patients with different performance status. In contrast, the Social and Emotional sub-scales of FACT-G failed to indicate this difference. With the exception of FLIC Psychological sub-scale, all FLIC and its sub-scales indicated a significant difference between patients of different treatment status. In contrast, only the Physical sub-scale of FACT-G was able to do so. The relative efficiency of FLIC versus FACT-G was significant in this regard ($P < 0.01$).

Two of the FLIC sub-scales (Symptoms and Social) had a reliability level of less than 0.6. However, it was the result of a smaller number of items (3 and 2 items, respectively) rather than a poor quality of the items. When the Spearman-Browne formula was used to adjust the alpha coefficients to a comparable group of 7 items, both of them showed a strong alpha value of 0.74 and 0.73, respectively.

In conclusion, we have found that the modified version of FLIC achieved suitable measurement properties. This is

a user-friendly alternative to the original FLIC. Furthermore, it is even more useful than FACT-G in assessing known group differences. Further research in several aspects are needed, including the assessment of FLIC in a longitudinal setting, and the level of Family and Social sub-scale scores of FLIC in relation to other established instruments of family and social function.

Acknowledgement

This work is funded by the National Medical Research Council (#0546/2001). We thank the oncologists and nurses of National Cancer Centre, Singapore for their help with the patient recruitment.

REFERENCES

1. Cascinu S, Labianca R, Daniele B, Beretta G, Salvagni S. Survival and quality of life in gastrointestinal tumors: two different end points? *Ann Oncol* 2001; 12(Suppl 3):S31-6.
2. Roila F, Cortesi E. Quality of life as a primary end point in oncology. *Ann Oncol* 2001; 12(Suppl 3):S3-6.
3. United Nations. World population prospects: the 2000 revision. New York: United Nations, 2001.
4. Department of Statistics. Singapore census of population 2000 statistical release 2: education, language and religion. Singapore: Department of Statistics, 2002.
5. Chia K S, Seow A, Lee H P, Shanmugaratnam K. Cancer incidence in Singapore 1993-1997. Singapore: Singapore Cancer Registry, 2000.
6. Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients: the Functional Living Index-Cancer: development and validation. *J Clin Oncol* 1984; 2:472-83.
7. Goh C R, Lee K S, Tan T C, Wang T L, Tan C H, Wong J, et al. Measuring quality of life in difference cultures: translation of the Functional Living Index for Cancer (FLIC) into Chinese and Malay in Singapore. *Ann Acad Med Singapore* 1996; 25:323-34.
8. Cella D. FACIT manual: manual of the Functional Assessment of Chronic Illness Therapy (FACIT) measurement system. Evanston: CORE, 1997.
9. Fisher L D, van Belle G. Biostatistics: a methodology for the health sciences. New York: Wiley, 1993.
10. Fayers P M, Machin D. Quality of life: assessment, analysis and interpretation. Chichester: Wiley, 2000.
11. Efron B, Tibshirani R J. An introduction to the bootstrap. New York: Chapman & Hall, 1993.
12. Ruckdeschel J C, Piantadosi S. Quality of life in lung cancer surgical adjuvant trials. *Chest* 1994; 106(6 Suppl):324S-328S.
13. Streiner D L, Norman G R. Health measurement scales. New York: Oxford Univ Press, 1989.
14. Downie W W, Leatham P A, Rhind V M, Wright V, Branco J A, Anderson J A. Studies with pain rating scales. *Ann Rheum Dis* 1978; 37:378-81.
15. McCormack H M, Horne D J, Sheather S. Clinical applications of visual analogue scales: a critical review. *Psychol Med* 1988; 18:1007-19.
16. Higginson I J, Carr A J. Measuring quality of life: using quality of life measures in the clinical setting. *BMJ* 2001; 322:1297-300.