Modeling Health Related Quality of Life among Cancer Patients Using an Integrated Inference System and Linear Regression

Lazim Abdullah and Jia Yu Low
School of Informatics and Applied Mathematics/University Malaysia Terengganu, Kuala Terengganu, Malaysia
Email: lazim_m@umt.edu.my, mikki0719@hotmail.com

***Abstract*—Health Related Quality of Life (HRQL) is one of the increasing subjects used for assessing health condition among patients who suffer from specific diseases or ailment. It has been assumed that identification of the variables is able to mirror the one's overall health conditions. However, devising the extent of contribution of multiple variables towards overall health conditions is not straight forward as the arbitrary nature of HRQL variables. This paper aims to model the relationship between HRQL variables using an integrated model of inference system and linear regression. An experiment was conducted to measure the strength of the relationship between the variables and health indices among cancer patients. To model this relationship, thirty outpatients with cancer were recruited from a government funded hospital in Kuala Terengganu, Malaysia. Linguistic data were collected via guided interview and fed into the fuzzy inference system to yield HRQL indices. Multi linear regressions were then undertaken to establish the relationship between the variables and HRQL indices. The model shows that the variable of emotion was identified as the highest risk factor for cancer patients. The use of integrated model, fuzzy inference system and multi linear regressions were successfully identified the strength of the relationship between the multi variables of HRQL and the health status.**

***Index Terms*—Quality of Life, Patients with Cancer, Fuzzy Inference, Linear Relationship, Linguistic Variable**

I. Introduction

In recent years, there has been an increasing interest in discussing the concept of quality of life and its related domains. The conceptual definitions of quality of life are indeed varies depending on social status and local preferences. Some of the definitions of quality of life are extended specifically to health related issues and normally refers as health related quality of life (HRQL). Most of the conceptual definitions of HRQL refer to a person or group's perceived their physical and mental health. According to Hays and Reeve [1], HRQL concerns how well people are able to function and how they feel about physical, mental, and social dimension of their lives. Philips [2] defines HRQL as a multi-dimensional concept that reflects a person's perception of their physical, psychological, and social function and health status. Bowling [3] defines HRQL as optimum level of physical role, for example, work, career, parent and social functioning, including relationships and perceptions of health, fitness, life satisfaction and well-being. Sharma [4] defines HRQL as a multi-dimensional dynamic concept that has developed from the need to estimate the impact of diseases, which includes economic welfare, characteristic of community and environment, and health status.

Some chronic diseases do not typically cause death but impacts of chronic diseases on quality of life are significant. Many studies have been embarked to provide evidence that chronic disease may give substantial adverse effects on general health, fitness, and physical, emotional, and social functioning have major implication to health status and well-being. Zakaria et al., [5], for example, employed a tool in measuring knee osteoarthritis patients and found that the patients were poor in quality of life due to the impact of the disease. A study by Lam and Lauder [6], confirmed that many common chronic diseases adversely affected the quality of life of Chinese patients, as they did for Caucasian patients. Depression, Osteoarthritis of the knee, other joint diseases, stroke, asthma and hypertension were each associated with a 30–200% relative increase in the risk of disability or ill health measured. One of the most frightening chronic diseases nowadays is cancer related diseases. Like many other chronic diseases, cancer becomes one of the major health issues in the world and not just it associates with life expectancy but also it substantially impacts on patients' quality of life. The cancer patients were hypothesized to less able performs physical activity hence some were forced to leave workplace. From psychological point of view, cancer patients also may suffer some extent of emotional disturbances.

Different approaches have been suggested to investigate impact of cancer disease to social and psychological variables. Fan, et al., [7] for example, employed inferential statistics to explain emotional factors of lung cancer patients at preclinical stage. Out of six psychosocial factors, three were positively associated with lung cancer patients. To make further investigation into the role of emotions among cancer patients, Hiatt et al., [8] proposed a framework to integrate the study of the biological nature of cancer and its clinical applications

Manuscript received January 2, 2015; revised March 2, 2015.

with the behavioral and social influences on cancer using quantitative approach. Meisel et al., [9] assessed quality of life among breast cancer survivors using non parametric statistics of Wilcoxon rank sum test, the Fisher exact test and Spearman correlation coefficients. In their study, although the overall quality of life is good in the studied population, there is a subset of women who are dealing with significant anxiety and depression, and a larger group who are experiencing burdensome sadness, hopelessness, and apprehension about their disease. Zebrack and Landier [10] examined whether childhood cancer survivors' perceptions of the impact of cancer are related to quality of life and psychological distress. Hierarchical linear regression models were used to analyze the independent effects of perceived impacts of cancer on distress and quality of life. At the same time, there are relatively few attempts to integrate inference-based intelligent tools with multivariate statistics approaches in solving the impacts of diseases toward patients.

The present study proposes an integration of intelligent decision making tool and a multi-variable statistical approach in searching the interactions between physical and emotional factors of patients with cancer. Specifically, this paper aims to model the relationship between HRQL indices of patients with cancer and the selected emotional and physical variables using the integrated approach of fuzzy inference system and multiple linear regressions.

II. Research Framework

This framework is designed to aid in conceptualizing how the experiment was conducted. It is aimed to search interaction between input variables and health status among patients with cancer. Thirty patients with cancer at a government funded hospital in Peninsular Malaysia were interviewed to tap their health related quality of life conditions specifically in the three areas. They were asked to respond in linguistics scales over the extent of 'emotion', 'unemployment' and 'daily activities' affect their quality of life. These input variables are fed into the fuzzy inference system to established HRQL indices. These indices are then used as response variable in multi linear regression model. The input variables were regressed with the indices to construct a multiple linear regression equation. The model is expected to be able to capture the most influential variable affecting quality of life. In short, this research flow is divided into two phases. Phase I explains the step-wise procedures to obtain HRQL indices using fuzzy inference system. Phase II presents the proposed step-wise procedures in obtaining multi linear regression equation. This equation could be used to identify the most effective predictor of HRQL among patients with cancer. Summarily, the research framework can be depicted in Fig. 1.

Computations for the data are implemented in the following section.

III. Computations and Results

The section describes detailed steps of the computation in accordance with the research framework. Phase I explains the stepwise procedures to obtain HRQL indices using FIS.

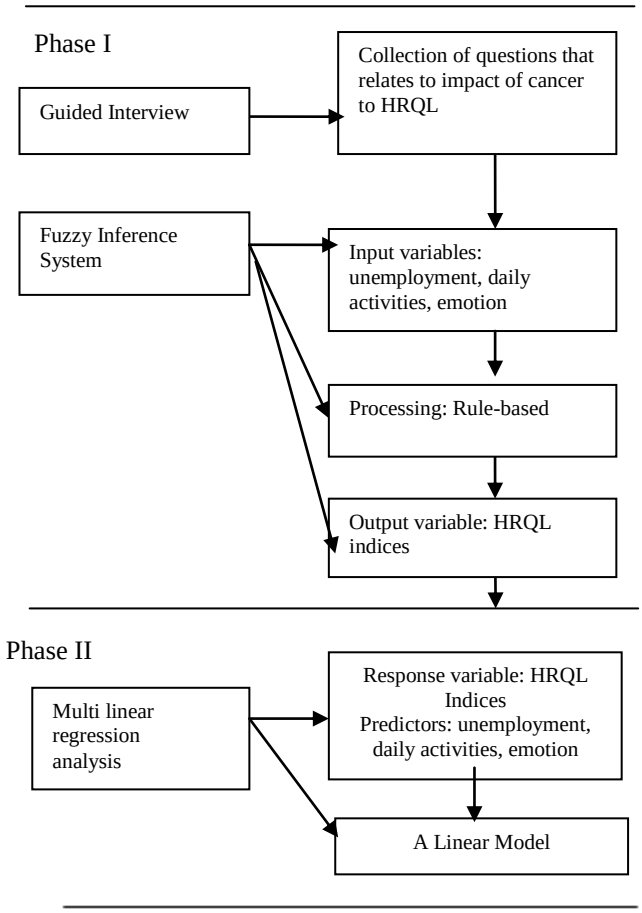

Figure 1. Framework of the research.

- Phase I

Step 1: Identify Input and Output Variables.

Input variables: Unemployment, Daily activities, Emotion

Fig. 2 depicts all the three membership functions of input variables.

Output variable: HRQL Index.

Fig. 3 shows the defined membership function of output variable.

Step 2: Fuzzify the variables.

The input variables are fuzzified by determining the degree to which they belong to each of the appropriate fuzzy set via membership functions. Input variable of emotion, for example, is fuzzified into five linguistics from 0 to 5 (see Fig. 2). Index for HRQL (output variable) is defuzzified into five linguistics of 'very high' to 'very low' in the range [0, 50] (see Fig. 3).

Step 3: Creating Rules

Based on input variable "Unemployment", "Daily Activity" and "Emotion", 40 possible and relevant 'IF–THEN' rules was created. Parts of the rules are shown in Fig. 4.

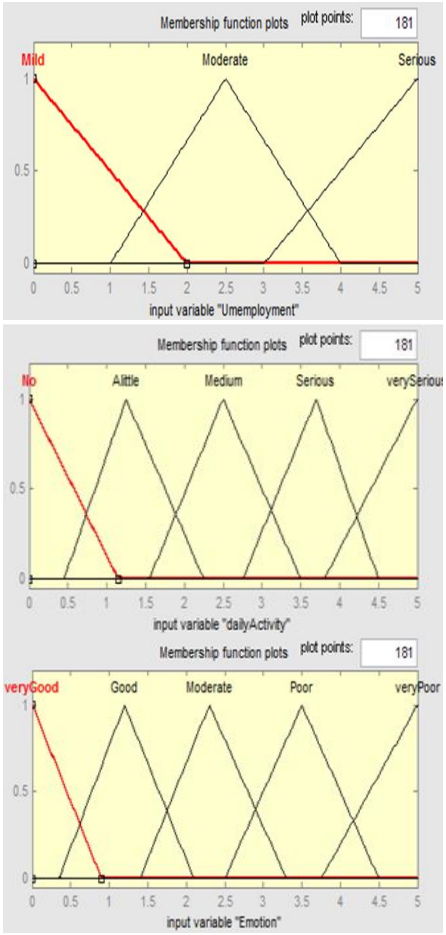

Figure 2. Membership function of input variables.

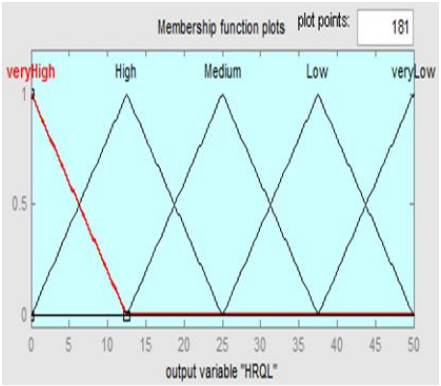

Figure 3. Membership of output variable.

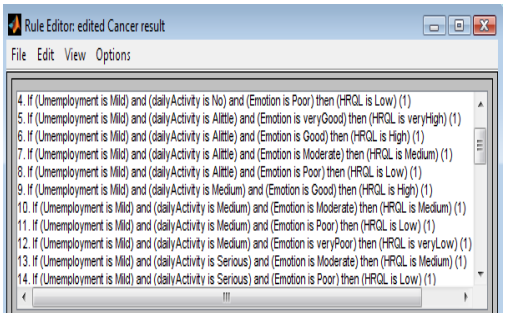

Figure 4. Examples of the rules.

Step 4: Defuzzification

Output can be calculated by inserting the data into rule viewer. The process of aggregating the variables into crisp number is called as defuzzification. Fig. 5 shows one of the many examples on how defuzzification is computed. For example, if the input of Unemployment = 2, Daily Activity = 4, and Emotion = 2, then resulted output of HRQL = 37.5.

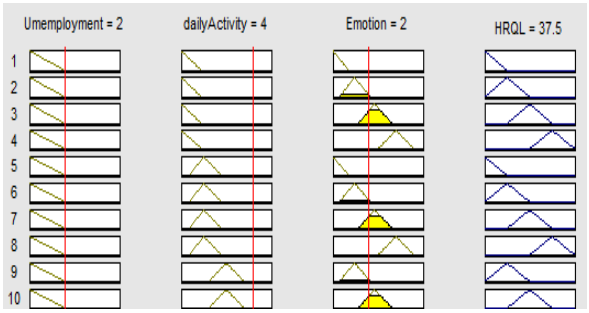

Figure 5. Defuzzification process.

With the similar steps, the indices for other sets of data were obtained. The linguistic used and indices for all patients are summarized in Table I.

TABLE I. HRQL INDICES FOR PATIENTS

No of Patients	Unemployment	Daily Activity	Emotion	HRQL index
1	Moderate	Serious	Serious	37.5
2	Mild	Medium	Serious	33.0
3	Moderate	High	A little	35.1
4	Serious	Low	Moderate	38.1
…	…	…	…	…
…	…	…	…	…
…	…	…	…	…
29	Serious	Very Serious	Poor	45.3
30	Moderate	Serious	Very Good	25.0

The HRQL index was successfully established by the system. The averaged index of thirty patients is 34.81. Based on the defined linguistic, this index can be interpreted as 'Low' with degree of membership 0.7. It is, therefore the HRQL status among cancer patients are not in satisfactory condition and should be given extra attention among medical fraternity and health care management.

- Phase II

The FIS successfully composited all the three input variables to yield the HRQL indices. Nevertheless, the effect of every single input variable toward the indices is still inconclusive. As to overcome with this problem, the analysis is resumed with the multiple linear regression. The effects of the input variables toward HRQL index could be dealt from the multi linear regressions equation. Multiple linear regressions is a method used to model the linear relationship between a dependent variable and one or more independent variables. The dependent variable is sometimes called the response variable and the

independent variables are called as the predictors. As to make the analysis more organized, the following steps are proposed.

Step 1: Identify predictors and response variable.

Predictors: Emotion, Daily activity, unemployment

Response variable: HRQL indices

(Output from Phase I).

Step 2: Assumption of normality.

All variables are subjected to fulfil the normality assumption. The normal P-P Plot of regression is plotted as to observe the data distribution. Fig. 6 shows the data distributions where the data are almost linear in P-P plot.

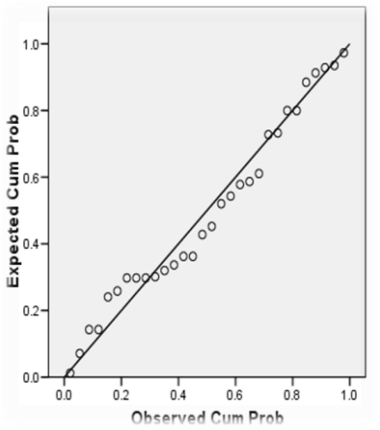

Figure 6. FPP-Plot of multiple linear regression.

The horizontal axis represents the observed cumulative probability and the vertical axis represents expected cumulative probability. The relationship between the observed and expected cumulative probability is almost linear. Therefore, assumption of normality for multiple regression analysis is fulfilled.

Step 3: Testing the relationships.

The relationship between predictors and HRQL indices are pre-tested using F-test based on the following hypotheses at significant level of =0.05.

H0: The model is not suitable for predicting HRQL of cancer patients

H1: The model is suitable for predicting HRQL of cancer patients

Summary of the test is shown in Table II.

TABLE II. F-TEST FOR THE RELATIONSHIP

	Sum of Squares	df	Mean Square	F	Sig.
Regression	2108.722	3	702.907	36.781	0.000(a)
Residual	496.873	26	19.110		
Total	2605.595	29			

With the F-value at 36.781 and 0.000 probability of rejection, the null hypothesis is rejected at 0.05 level of significance. There is a relationship between predictors and HRQL indices. Therefore, the multi linear regressions can be used in predicting the HRQL among cancer patients.

Step 4: Obtain the multi linear regression equation.

Multi linear regression equation is used in explaining the behaviors of predictors toward HRQL indices among cancer patients. The coefficients of the regression equation are presented in Table III.

TABLE III. REGRESSION COEFFICIENTS

	Unstandardized Coefficients		t	Sig
	B	Std. Error		
(Constant)	8.224	2.822	2.914	.007
Emotion	5.028	0.825	6.093	.000
Daily activity	2.445	0.733	3.337	.003
Unemployment	1.754	0.794	2.208	.036

As can be seen from the Table II, the linear regression equation can be written as,

$$\hat{Y} = 8.224 + 5.028\,(\text{Emotion}) + 2.445\,(\text{Daily activity}) + 1.754\,(\text{Unemployment}),$$

where $\hat{Y}$ is the expected values of HRQL indices.

It is evident that of the three predictors, the variable of 'emotion' is the highest contributors toward HRQL indices among patients with cancer. This linear model indicates that psychological disturbances such as emotions are the main health related problem among cancer patients.

IV. CONCLUSIONS

The issue of arbitrary nature of health related quality variables was dealt with the integration model of fuzzy inference system and multi linear regressions. The experiment was conducted to show the relationship between health indices and the three selected variables. Thirty patients with cancer related diseases were chosen as participants in this experiment. The fuzzy inference system provided the health indices whereby these indices were subsequently used as the explanatory variable in the regression analysis. This paper has successfully proposed a linear relationship between the health related quality of life indices and its variables among cancer patients. It is demonstrated that the linear relationship can identify the best predictor for health related quality of life among cancer patients. Emotion of the patients was identified as one of the significant variables in prescribing and diagnosing cancer patients. It is important to note that predictors for quality of life among cancer patients are not limited to the three selected variables. Future research may be undertaken to include greater number of participants and more variables to substantiate these findings.

REFERENCES

[1] R. D. Hays and B. B. Reeve, "Measurement and modelling of health-related quality of life," *International Encyclopedia of Public Health*, pp. 241-252, 2008.

[2] D. Phillips, *Quality of life: Concept, Policy and Practice*. Oxon: Routledge, 2006.

[3] A. Bowling, "What things are important in people's lives? A survey of the public's judgments to inform scales of health-related quality of life," *Social Science & Medicine*, vol. 41, pp. 1447-1462, 1995.

[4] D. Sharma, "Health-related quality of life and its assessment in GI Surgery," *Indian Journal of Surgery*, vol. 66, pp. 323-325, 2004.
[5] Z. F. Zakaria, A. A. Bakar, H. M. Hasmoni, F. A. Rani, and S. A. Kadir, "Health-related quality of life in patients with knee osteoarthritis attending two primary care clinics in Malaysia: A cross-sectional study," *Asia Pacific Family Medicine*, vol. 8, pp. 10, 2009.
[6] C. L. K. Lam and I. J. Lauder, "The impact of chronic diseases on the health-related quality of life of Chinese patients in primary care," *Family Practice*, vol. 17, pp. 159-166, 2000
[7] R. L. Fan, S. H. Zheng, and Z. S. Wu, "Study on the relationship between lung cancer at preclinic stage and psycho-social factors: A case control study," *Chinese Journal of Blood Diseases*, vol. 18, pp. 289-292. 1997.
[8] R. A. Hiatt and N. Breen, "The social determinants of cancer: A challenge for trans disciplinary science," *American Journal of Preventive Medicine*, vol. 35, pp. S141-S150. 2008.
[9] J. L. Meisel., S. M. Domchek, R. H. Vonderheide, A. C. Giobbie-Hurder, and A. H. Partridge, "Quality of life in long-term survivors of metastatic breast cancer," *Clinical Breast Cancer*, vol. 12, pp. 119-126, 2012.
[10] B. J. Zebrack and W. Landier, "The perceived impact of cancer on quality of life for post-treatment survivors of childhood cancer," *Quality of life Research*, vol. 20, pp. 1595-1608, 2011.

Lazim Abdullah is a professor at the School of Informatics and Applied Mathematics, University Malaysia Terengganu. He holds a B.Sc (Hons) in mathematics from the University of Malaya, Kuala Lumpur in June 1984 and the M.Ed in mathematics education from University Sains Malaysia, Penang in 1999. He received his Ph.D. from the University Malaysia Terengganu, (Information Technology Development) in 2004. His research focuses on the mathematical theory of fuzzy sets and its applications in social ecology, environmental sciences, health sciences, and manufacturing engineering. He is interested in the measurement of social indicators, indices of health related quality of life and environmental index using computational intelligence and statistical approaches. His research findings have been published in over two hundred publications including refereed journals, conference proceedings, and chapters in book, and research books. Currently he is a member of editorial boards to several international journals that related to computing and information technology. Besides, he is also a reviewer for number of local and international journals, member of scientific committees of several symposia and conferences at national and international levels. He is an associate member IEEE Computational Intelligence Society, a member of Malaysian Mathematical Society and a member of the International Society on Multiple Criteria Decision Making.

Jia Yu Low is a research student at School of Informatics and Applied Mathematics, University Malaysia Terengganu, Malaysia.