APPLICATION OF FUZZY LINEAR REGRESSION MODELS FOR PREDICTING TUMOR SIZE OF COLORECTAL CANCER IN MALAYSIA'S HOSPITAL

MUHAMMAD AMMAR BIN SHAFI

A thesis submitted in
fulfillment of the requirement for the award of the
Degree of Master of Science

Faculty of Science, Technology and Human Development
University Tun Hussein Onn Malaysia

JULY 2015

DEDICATION

This thesis is dedicated exclusively to my lovely parents that:

Haji Shafi Bin Haji Shaid

Hajah Hasidah Binti Haji Razali

Blessing both of you always accompanies me now and hereafter.

My beautiful wife:

Nur Azia Hazida Binti Mohamad Azmi

My siblings:

Muhammad Ismail Bin Haji Shafi

Nur Syafiqah Binti Haji Shafi

Muhammad Firdaus Bin Haji Shafi

And my sister in-law

Nik Suhaily Binti Ismail

Always support and also to the loved ones who are always there for me.

ACKNOWLEDGEMENT

"In the name of Allah SWT, the Merciful and Beneficent"

Praise is to Allah S.W.T for the blessing and giving me the strength and health to complete this thesis. First and foremost, I would like to thanks to my kindness supervisor, Dr. Mohd Saifullah Bin Ruisman for his invaluable advice, comments and constant guidance during the progress of my thesis. I would like to express my special acknowledgement to my co-supervisor, Dr.Mahathir Bin Mohamad and Dr. Siti Noor Asyikin Binti Mohd Razali for giving me an opportunity to advised and teaches me about the statistical methods and research methodology. The cooperation given by the general hospital in Kuala Lumpur is also highly appreciated. Sincere appreciations to my beloved family, especially my parents Hj. Shafi bin Hj. Shaid and Hajah. Hasidah Binti Haji Razali who are given full supports and encouragement in the process to finish this thesis, my thank you will be given to my beautiful wife Nur Azia Hazida Binti Mohamad Azmi that give encouragement and help me during thesis process and all my siblings Muhammad Ismail, Nur Syafiqah and Muhammad Firdaus and my sister in-law Nik Suhaily who are never give up giving me an advised and supported. Last but not least, special thanks to all beloved friends who are shared their experiences, time and commitment in finishing this thesis. No such valuable words than "Thank You" to appreciate your eagerness and kindness. More than words could express on paper or could be spoken in words. Hopefully this research thesis would be guidance in future. Thank You.

MUHAMMAD AMMAR BIN HJ.SHAFI

ABSTRACT

Fuzzy linear regression analysis has become popular among researchers and standard model in analysing data vagueness phenomena. These models were represented by five statistical models such as multiple linear regression, fuzzy linear regression (Tanaka), fuzzy linear regression (Ni), extended fuzzy linear regression under benchmarking model (Chung) and fuzzy linear regression with symmetric parameter (Zolfaghari). A case study in colorectal cancer (CRC) data at the general hospital in Kuala Lumpur was carried out using the five models as mention above. Secondary data of 180 colorectal cancer patients who received treatment in general hospital were recorded by nurses and doctors. Twenty five independent variables with different combination of variable types were considered to find the best models to predict the size of tumor colorectal cancer. The quality of life among CRC patients which is to detect the early CRC stage is still very poor, not implemented and divulged as a nationwide programme. The main objective of this study is to determine the best model by predicting the size of tumor of CRC. Moreover, this study wants to identify the factors and symptoms that contribute the size of tumor. The comparisons among the five models were carried out to find the best model by using statistical measurements of mean square error (MSE) and root mean square error (RMSE). The results showed that the fuzzy linear regression with symmetric parameter (Zolfaghari) was found to be the best model, having the lowest MSE and RMSE value by 98.21 and 9.91. Hence, the size of tumor could be predicted by managing twenty five independent variables.

ABSTRAK

Analisis regresi linear kabur telah menjadi popular di kalangan penyelidik dan menjadi model yang biasa digunakan di dalam fenomena kabur. Model-model ini diwakili oleh lima model statistik seperti regresi linear berganda, regresi linear kabur (Tanaka), regresi linear kabur (Ni), regresi linear kabur lanjutan di bawah model tanda aras (Chung) dan regresi linear kabur dengan parameter simetri (Zolfaghari). Satu kajian kes terhadap data kanser usus (CRC) di hospital umum di Kuala Lumpur menggunakan lima model seperti yang dinyatakan di atas telah dijalankan. Data sekunder daripada 180 pesakit kanser usus yang menerima rawatan di hospital umum telah direkodkan oleh jururawat dan doktor. Dua puluh lima pembolehubah tak bersandar dengan pelbagai jenis kombinasi pemboleh ubah telah digunakan untuk mencari model yang terbaik untuk menjangkakan saiz tumor. Kualiti hidup pesakit CRC yang bertujuan untuk mengesan peringkat awal CRC adalah masih sangat lemah,tidak dilaksanakan dan dihebahkan sebagai satu program di peringkat kebangsaaan.Objektif utama kajian ini adalah bertujuan untuk menentukan model yang terbaik dengan meramal saiz tumor bagi pesakit kanser usus (CRC).Tambahan objektif dalam kajian ini adalah untuk mengenal pasti faktor-faktor dan simptom-simptom yang boleh menyumbang kepada saiz tumor. Perbandingan di antara lima model telah dijalankan untuk mencari model terbaik dengan menggunakan ralat kuasa dua min (MSE) dan ralat punca kuasa dua min (RMSE). Keputusan menunjukkan bahawa model regresi linear kabur dengan parameter simetri (Zolfaghari) adalah menjadi model yang terbaik dengan nilai MSE dan RMSE yang terendah dengan jumlahnya 98.21 dan 9.91. Oleh itu, saiz tumor boleh dijangka oleh dua puluh lima pembolehubah tak bersandar.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/ SYMBOLS

A	- Fuzzy set of A
AHP	- Analytical hierarchy process
ANOVA	- Analysis of variance
B	- Level of existence corresponds to the level membership
BMI	- Body mass index
CI	- Condition index
CI	- Confident interval
CRC	- Colorectal cancer
c_i	- Width of fuzzy parameter
C_{ii}	- ith diagonal element of matrix $(\mathbf{X}^T\mathbf{X})^{-1}$
DM	- Data mining
EFLRBM	- Extended fuzzy linear regression under benchmarking model
EORTC	- European Organisation for Research and Treatment of Cancer
FCM	- Fuzzy c-mean model
FCRM	- Fuzzy c-regression model
FIT	- Fecal immunochemical test
FLR	- Fuzzy linear regression
FLRWSP	- Fuzzy linear regression with symmetric parameter

FOBT	- Fecal occult blood test
$f(x, \mathbf{A})$	- Fuzzy function of set **A**
$f(\mathbf{X}, \mathbf{A})$	- Fuzzy model
gFOBT	- Guaiac based fecal occult blood test
ICU	- Intensive care unit
icd10	- Place where CRC existed by patient
IDA	- Intelligent data analysis
IRFC	- Iterative relative fuzzy connectedness
MANOVA	- Multivariate analysis of variance
MF	- Membership function
MLR	- Multiple Linear Regression
Mm	- Milimetre
MOHM	- Ministry of Health Malaysia
MOS	- Malaysian Oncology Society
MRA	- Magnetic resonance angiography
MRI	- Magnetic resonance imaging
MSE	- Mean sqaure error
MSR	- Mean square regression
NCR	- National Cancer Registry
Q-Q plot	- Plot of quartile
R	- Right reference function
R^2	- Coefficient of determination
RMSE	- Root mean sqaure error
RS2337107	- Gene of polymorphism on risk colorectal cancer
Sig.	- Significant value
SMAD7	- Functional candidate gene for colorectal cancer
SPSS	- Software package for statistical analysis

SSE	- Sum of error
SSR	- Sum of regression
SST	- Sum of total
TNM	- Tumor, nodes and metastases
VIF	- Variance inflation factor
WHI	- Waist circumference to height index
WHO	- World Health Organization
α	- Center for fuzzy parameter
β_i	- Coefficient in multiple linear regression
$\mathbf{B_i}$	- Vector of constants
e	- Index regression number from 1 to n
ε	- Random error of parameter
$\boldsymbol{\varepsilon}$	- Vector of independent normal random variables
$\Im$	- Infinite dimensional feature space
f	- Index regression number from 1 to n
H	- Height of fuzzy triangular
ith	- Sample of i
$\mathbf{J}$	- An nxn matrix
n	- Number or the respondents less than 30
N	- Number of observations 30 and above
P	- Amount of predictor variables
s	symmetric measurement
S	- Standard deviation
$\mu_{\mathrm{A}}(a)$	- Membership function of element $a$in set A
X_i	- Parameter of independent variables
$\boldsymbol{X_i}$	- Matrix parameters
$\bar{x}$	- Mean or average

$\mathbf{X}^{\mathrm{T}}$	- Transpose for matrix **X**
Y	- Dependent variables/ observations
y_i	- Observation of data
$\hat{y}$	- Output of fuzzy model
$\hat{y}_i$	-Predicted data
y^*_i	- Estimated value for y
$\boldsymbol{Y_i}$	- Vector of responses
$\sum$	- Covariance matrix

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

This chapter presents the introduction on research background of regression analysis, colorectal cancer, history of fuzzy regression, problem statement, research objectives, scope of study, research contribution and research organization. Lastly, a brief summary is included at the end of chapter.

1.2 Research background of regression analysis

Regression analysis has become one of the standard models in analysing data. It's popularity comes from several sources. The statistical equation is derived obtained from the analysis which explains the relationship of dependent and independent variables. It provides much explanatory power, especially due to its multivariate nature. It is widely available in computer packages and can be easily interpreted. It is extensively used in applied sciences, economic, engineering, computer, social sciences and other fields (Agresti, 1996).

Nonlinear modeling is of interest to many researchers in modeling, statistics, rather than linear modeling. The obtained functional form should be approximately close to the real data. If the functional form varies much from the real data, the estimation will be inconsistent, bias and so on. However, other difficulties may also arise with non-linearity approach (Rousseeuw et al., 1987). Nowadays, there are many modeling resulting from the regression analysis, such as multiple regression, quadratic regression, cubic regression, logit model, probit model, exponential model, growth model, neural network regression and fuzzy regression.

1.3 Research background of colorectal cancer

Colorectal cancer (CRC) is a term of cancer in the colon and/or rectum. The colon and rectum are two different parts in the human body which play an important role to digest food and past waste. Colorectal cancer is one of the commonest diseases malignancies in the world (Malaysian Oncological Society, 2007). According to World Health Organization (WHO) (2012), colorectal cancer was the fourth leading cause of death among cancer patients. The risk of this particular cancer is rising in many countries, especially in the Asian Region of Malaysia (Yusoff et al., 2012).

Colorectal cancer is the third leading cause of cancer deaths in Malaysia. According to the data of 2010 from WHO, lung cancer is the highest rated cause of cancer deaths by 17.93% and followed by breast cancer 15.83% and colorectal cancers 13.10%. Data from the Ministry of Health of Malaysia showed that in 1995, colorectal cancer admission rates increased from 8.1% in 1987 to 11.9%. Recent studies also showed an incremental trend in colorectal cancer among Asian population (Wendy, 2008 and Radzi, 2008). According to Second Paper of the National Cancer Registry (2003), there were 14.2% of male and 10.1% of female colorectal cancer patients were identified in Malaysia. It can be concluded that colorectal cancer is the commonest cancer among men and the second most common cancer among women.

At present, cause of CRC is not completely clear. However, it involves many factors contributing in developing the cancer such as incresing age, lack of nutritions, family history, body weight and so on. People with colorectal cancer may develop a number of non-specific symptoms. Symptoms of colorectal cancer typically include rectal bleeding and anemia which are sometimes associated with weight loss and changes in bowel habits (Cappell, 2005).

Colorectal cancer can be detected by factors of symptoms or common symptoms, namely abdominal cramps, a change in your bowels habits, blood in stools, weight loss, fatigue and loss of appetite. All these symptoms may show their activation in one month after the colorectal cancer eventuated (Vanagunas, 2011). According to American Cancer Society (2014), the cancer can be examined by physical exam test to check for early symptoms. The test includes colonoscopy, fecal

occult blood testing, sigmoidoscopy, biopsy and complete blood count. CRC includes four stages from an early stage until final stage. Stage I indicates that the cancer is confined to the inner lining of the colon or rectum. Stage II is cancer spreads through the wall of the colon or rectum while in stage III, cancer spreads to nearby lymph nodes and the final stage, IV is cancer spreads to distant parts of the body, such as liver or lungs (Malaysian Oncological Society, 2007).

European Organisation for Research and Treatment of Cancer (EORTC) in 2014 stated that the colorectal cancer (CRC) is also increasing significantly in other countries. It is the second most common cancer in the United Kingdom after lung cancer as well as in the United States of America. More than 50,000 people from each year dies of colorectal cancer (Jemal et al., 2008). Colorectal cancer is the leading killer among smokers in the world (Frieden et al., 2002). In the Asian population, there is a rapid increase in the trend of CRC, where it becomes the most common cancer among male superseded lung cancer (National Cancer Registry, 2006).

In Asia, screening with faecal occult blood test is a national policy only in Japan, Taiwan and Korea (Hyodo et al., 2010). Currently, in Malaysia, the awareness of the importance of CRC screening is very low, especially among the general population as well as the policy makers resulting in inadequate resources allocation for faecal occult blood test and colonoscopy. The advances in the treatment of colorectal cancer and effectiveness of the screening program, especially in the developed countries resulted increasing number of colorectal cancer survivors. Not only the disease could be cured but patients can also survive longer with the disease. Therefore, the paradigm of outcome in colorectal cancer treatment in the past decade has slowly shifted towards improving the patients' quality of lives, survival and disease free survival. This shows that health is related to quality of life and it is one of the fundamental aspects in colorectal cancer management while survival and disease free survival remains (Bottomley, 2002).

In Malaysia, about 1 in 15 people develop colon cancer and the mean age for colorectal cancer is 61 years (range 15 to 95 years) and the majority patients were 50 years and above (National Cancer Patient Registry, 2010). Patients with colorectal cancer have different pathways of presentation for diagnosis; primary screening, symptomatic and incidental. Most patients presented symptomatically. The most common presenting symptom is "diarrhea, constipation, or other change in bowel

habit", followed by "weight loss" and "abdominal pain". The rectum was the most frequent site of the primary tumor followed by sigmoid colon and the rectosigmoid. Figure 1.1 and Figure 1.2 shows the tumor sites and stages.

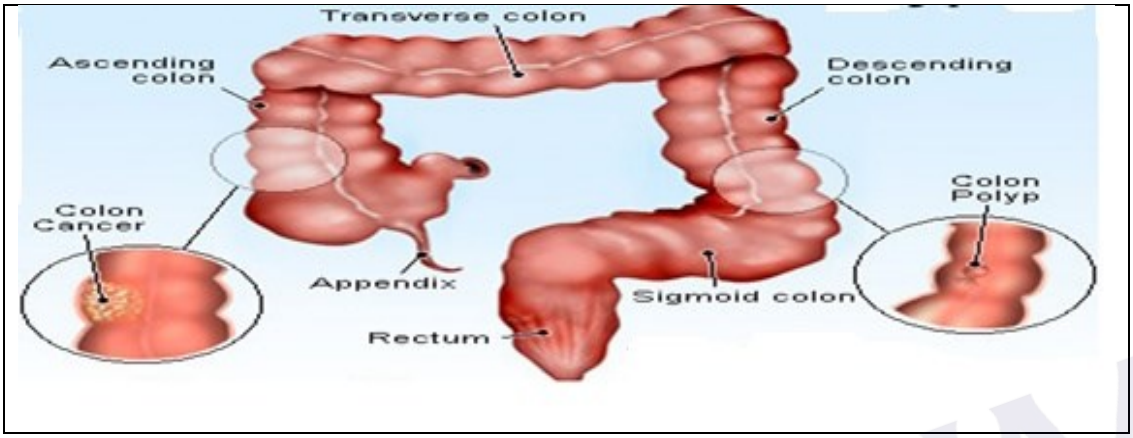

Figure 1.1: Colon cancer and polyp

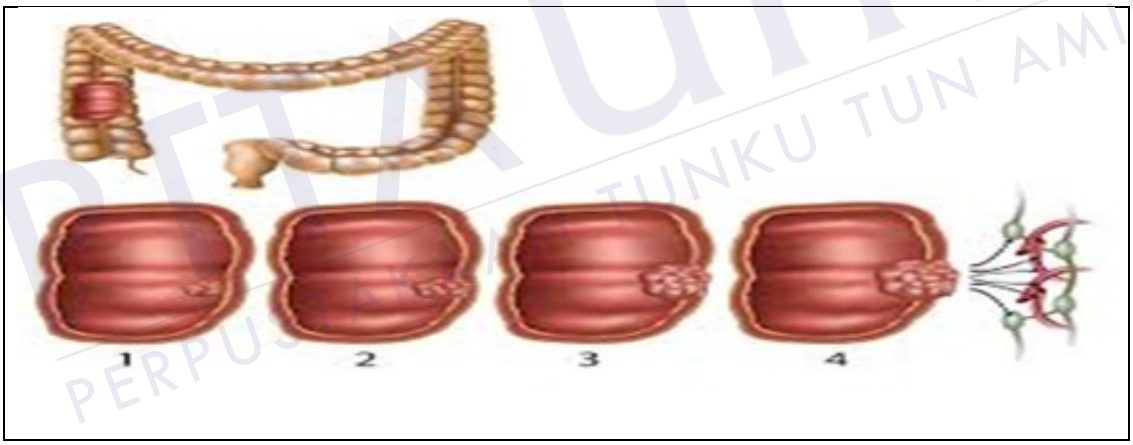

Figure 1.2: Stages of colorectal cancer

1.4 Research background on fuzzy regression

A fuzzy regression model is used in evaluating the functional relationship between the dependent and independent variables in a fuzzy environment. In many cases of fuzzy regression, the linear regression is recommended for practical situations when decisions often have to be made on the basis of imprecise and/or partially available data. Several methods have been presented to estimate fuzzy regression models. At first, fuzzy regression was developed by Tanaka et al. in year 1982 for linear case by focusing on extension principle (Taheri, 2003).

Making sense of data is an ongoing task for researchers and professionals in almost every practical endeavor. The age of information technology, characterized by a vast array of data, has enormously amplified this quest and made it even more challenging. Data collection has become the reality of our lives at any time and from everywhere. It is reported that understanding the data, revealing underlying phenomena, and visualizing major tendencies are the major undertakings to pursue in intelligent data analysis (IDA), data mining (DM), and system modeling (Pedryoz, 2005).

Fuzzy regression used in complex systems such as in industry, economy, finance, marketing, and ecology function in the real world and it is more imprecision. Such systems require decisions based on human thinking and judgmental and involve human–machine interactions. In such environments, human often not able to obtain exact numerical data about the system. The nature of information about the complex systems with vagueness is frequently fuzzy. In general, fuzzy regression seems to be intuitively more adequate for real life problems. Therefore, fuzzy regression analysis is more effective for modeling of complex systems. The pioneering work in this field reported that the authors used Zadeh's extension principle, a-level procedure, interval arithmetic, and linear programming techniques to develop a fuzzy linear regression analysis. Minimization of these distances in the fuzzy number space with respect to the unknown parameters of regression models leads to solving systems of equations (Aliev et al., 2002).

1.5 Problem statement

This research is carried out to detect size of the tumor of colorectal cancer which are in stage I and stage II in Malaysia. There are several studies about colorectal cancer (Varela, 2010). However, solutions for early stage detection is still not found.

In Malaysia, only a few studies were done to access the quality of life among CRC patients. Natrah (2012) conducted a study to determine the quality of life in CRC patients, especially those under treatment. The respondents in this study were mainly from stage III and IV, which are considered the tumor size in final stages of the disease and no respondents were from stage I and II from the public tertiary level general hospitals. This finding shows that the awareness of tumor size of colorectal cancer screening in Malaysia to detect the early stage of CRC is still very poor and not implemented as a nation's wide program. This correlates with findings from a study made by Kong et al. (2010), where it claimed that it is close to zero awareness of CRC screening tumor size among general population.

People, whose family members have colorectal cancer, especially at an old age, have a higher risk of developing this disease. A woman who has had cancer of the ovary, uterus or breast is more likely to develop colorectal cancer. Many people do not know about the CRC program and also there are people who know that they are having the symptoms of CRC but refuse to go for screening. For the screening program to be effective, an "at-risk" population has to be identified. The incidence of colorectal cancer increases exponentially with age and current figures from the developed world show that 90% of cases arise from those aged over 50 years old. In developed nations, the condition is thus too rare below the age of 50 years old to make screening test (Bennet, 1996).

In fuzzy linear regression, there are important models such as fuzzy regression and multiple linear regressions. Both models are usually used by researchers with continuous dependent data. Fuzzy linear regression models were focused to get the factors and symptoms that give influence to tumor size and to get the less error of model. Recently, researchers are interested to get small value of error in analysis by using fuzzy models (Saifullah et al., 2012). Hence, fuzzy linear regression is the effective model to adopt in this study based on previous research.

1.6 Research objectives

This research is an attempt to find the most appropriate colorectal cancer model in order to predict the size of tumor among CRC patients. The specific objectives of this study are detailed as below:

(i) To apply the existing latest model of fuzzy linear regression with symmetric parameter in order to identify the factors and symptoms of CRC patients.

(ii) To apply the existing models which are multiple linear regression, fuzzy linear regression (Tanaka), fuzzy linear regression (Ni), and extended fuzzy linear regression by benchmarking models under fuzziness (Chung) in order to identify the factors and symptoms of CRC patients.

(iii) To make comparison among the linear regression models as mention above in order to find the best model for predicting tumor size in colorectal cancer.

(iv) To identify the factors and symptoms that give most significant influence to the colorectal cancer.

1.7 The scope of the study

The scope of the study will discuss about two sections which are the scope of the data and followed by the scope for the model.

1.7.1 Data scope

This study used secondary data obtained from the general hospital in Kuala Lumpur. This data consists of 180 patients as respondents for colon cancer. It was recorded by nurses and doctors using cluster sampling. Dependent variable or outcome is the size of tumor. Actually, there were twenty five independent variables which are gender, ethnic, age, classification of diseases and related health problems (icd10), tumor nodes metastases (TNM) staging, family history of colon cancer, diabetes mellitus, crohn's disease, ulcerative colitis, polyp, history of cancer, endometrial, gastric, small bowel, hepatobiliary, urinary tract, ovarian, other cancer, intestinal obstruction, colorectal, weight loss, diarrhea, anaemia, blood stool and abdominal pain.

The range value size of the investigated tumor is 20mm till 100mm. The patient's feedback on colorectal cancer was interviewed face to face by doctors and the answers were obtained immediately. The questions covered twenty four variables except for TNM staging.

Stage of colorectal cancer focuses on how big the cancer is and whether it is spread in the bowel of human body from polyp. Doctors and nurses used test and scanned to determine the stage of colon cancer faced by the patients. TNM stands for tumor, node and metastases. This way is widely used by medical cancer fields in the world. TNM describes the size of a primary tumor (T), any lymph contain cancer cells (N) and the spread of cancer to another part of body (M).

There are four stages in a primary tumor (T) which classified as,T1: tumor is only in the inner layer of the bowel,T2: the tumor is grown into the muscle layer of the bowel wall, T3: the tumor is grown into the outer lining of the bowel wall, T4: the tumor is grown through the outer lining of the bowel wall. While, three possible stages for lymph contain cancer cells identified in the lymph nodes (N). They are categorized as N0: there are no lymph nodes containing cancer cells, N1:1 to 3 lymph nodes close to the bowel contains cancer cells, N2: there are cancer cells in 4 or more nearby lymph nodes. Furthermore, there are two stages of cancer spread (M)

REFERENCES

Agresti, A. (1996). *An Introduction to Categorical Data Analysis.New York : John Wiley & Sons, Inc.*

Akbari Zahra, Safari-Alighiarloo Safari and Montazer Mahdi.(2014). Lack of Influence of the SMAD7 Gene rs2337107 Polymorphism on Risk of Colorectal Cancer in an Iranian Population.*Asian Pacific Journal of Cancer Prevention*, (15), 4437-4441.

Aliev, R. A., Fazlollahi, B. and Vahidov,R. (2002). Genetic algorithms-based fuzzy regression analysis. *Journal of Soft Computing,* (6), 470 – 475.

American Cancer Society. (2012). *How is colorectal cancer staged?* American Cancer Society.

American Cancer Society. Cancer Facts & Figures 2014 Atlanta.(2014*). American Cancer Society*.

Baig Faran, Khan M. Saleeem, Noor Yasir and Imran M.. (2011). Design Model of Fuzzy Logic Medical Diagnosis Control System. *International Journal of Computer Science and Engineering* (IJCSE),(3), No.5, 2093-2108.

Bennet, D.H. and Hardcastle JD.(1996). Early Diagnosis and Screening.*Colorectal Cancer.Ed Williams NS. Churchill Livingstone*, 21-37.

Bezdek, J. C. (1974). Cluster validity with fuzzy set. *Journal Cybernetic*, (3), 58–72.

Bezdek, J. C. (1981). Pattern recognition with fuzzy objective function algorithms. USA,*Kluwer Academic Publishers.*

Bottomley A. (2002). The cancer patient and quality of life. *The Oncologist Cancer Research UK. Cancer Stats Key Facts*,(7), 120-5.

Brand M. Rhonda, David D Jones, Henry T Lynch, Randali E Brand, Patrice Watson, Ramesh Ashwathnayaran and Hmeant K Roy. (2006). Risk of Colon Cancer in Hereditary Non-Polypsis Colorectal Cancer Patients as Predicted by Fuzzy Modeling: Influence of Smoking. *World J Gastroenterol* , 12(28), 4485-4491.

Cappell S. Mitchell. (2005). The pathophysiology, clinical presentation and diagnosis of coloncancer and adenomatous polyps. *The Medical Clinics of North America,* (89), 1-42

Center MM, Jemal A and Ward E. (2009). International trends in colorectal cancer incidence rates. *Cancer Epidemiology Biomarkers*, (19), 1688-94.

ChungWilliam.(2012). Using the Fuzzy Linear Regression Method to Benchmark the Energy Efficiency of Commercial Buildings. *Applied Energy*, 45-49.

Chung William. (2012). Construction of Benchmarking Models Using Fuzzy Linear Regression Techniques. *International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012) IEEE*. 555-559.

Dubois, D. J. and Henry. (1980). Fuzzy set and systems: *Theory and applications*. Academic Press Inc.

Dunn, J. (1974). A fuzzy relative of the ISODATA process and its use in detecting compact well separated cluster. *J. Cybernetics,* 3(3), 32-57.

Elizabeth S. and Sujathan L. (2013). Apllication of Fuzzy Membership Matrix in Medical Diagnosis and Decision Making. *Applied Mathematical Sciences*,(7), No. 127, 6297-6307.

El-Melegy Moumen T. and Mokhtar Hashim M. (2014). Tumor Segmentation in Brain MRI Using A Fuzzy Approach with Class Center Priors. *Journal on Image and Video Processing*. 1-14.

European Organisation for Research and Treatment of Cancer (EORTC). (2014). *Colorectal Cancer Facts and Figures In support of the Spectacolor Biobank Project in United Kingdom.*

Frieden. (2002). Physical Activity, Dietary Fat and Colorectal Cancer. *PhD Thesis Martina Perše University of Ljubljana, Faculty of Medicine, Institute of Pathology, MEC,Slovenia.*

Hata. (1998). Medical Image Segmentation by Fuzzy Logic Techniques. *IEEE 1998 Department of Computer Engineering, Himeji Institute of Technology, Japan,* 4098-4103.

Hyodo I, Suzuki H and Takahashi K. 2010. Present status and perspective of colorectal cancer in Asia.Colorectal Cancer Working Group Report in 30th Asia-Pacific Cancer Conference.*Jpn J ClinOncol*, (40), 38-43.

Ibrahim B.S.K.K, Tokhi M.O., Huq M.S and Gharooni S.C.. (2011). Optimized Fuzzy Control for Natural Trajectory Based Fes-Swinging Motion. *International Journal of Integrated Engineering*, (3), No. 2, 17-23.

Jemal A, Siegel R and Ward E. (2008). Cancer Statistics. C.A. *A Cancer J Clin*, (58), 71-96.

Kaneko Rena, Nakazaki Natsuko, Tagawa Teppei and Ohishi Chitose. (2014). A New Index of Abdominal Obesity which Effectively Predicts Risk of Colon Tumor Development in Female Japanese. *Asian Pasific Journal of Cancer Prevention*, (5), 1005-1010.

Kong CK, Roslani AC and Law CW. (2010). Impact of Socioeconomic Class on Colorectal Cancer Patient Outcomes in Kuala Lumpur and Kuching, Malaysia.*APJCP*, (11), 969-74.

Kutner H. Michael, Nachtsheim, Neter John and Li William.(2004).Applied Linear Statistical Models.Fifth Edition. *Applied Linear Statistical Models Mc Graw Hill,* 197-209.

Lazim and Nadia.(2012). Matrix Driven Multivariate Fuzzy Linear Regression Model in Car Sales.*Journal of Applied Sciences*, 1-8.

Lavdaniti Maria, Barbas George, Fratzana Aikaterini and Zyga Sofia.(2012). Evaluation of Depression in Colon Cancer Patients. *Health Science Journal*, (6), Issue 4, 681-692.

Liu, Xiaofeng, Ma, Lin, Zhang, Sheng, & Mathew, Joseph. (2005). Using fuzzy c-means and fuzzy integrals for machinery fault diagnosis. *In International Conference on Condition Monitoring*, Cambridge, England.1-9.

Malaysian Oncological Society Novartis Corporation (Malaysia). (2007). *The Lancet Oncology*,(8), 773-783.

Ministry Of Health, Malaysia. (1995). *Information and Documentation System Unit. Planning & Development Division.*

Natrah MS, SharifaEzat WP, Syed MA, Mohd Rizal AM and Saperi S. (2012). Quality of Life in Malaysian Colorectal Cancer Patients: A Preliminary Result. *APJCP*, (13), 1-6.

National Cancer Registry, Ministry of Health Malaysia. (2006). *Malaysian Cancer Statistics – Data and Figure Peninsular Malaysia*.

National Cancer Registry, Ministry of Health Malaysia. (2010). *Malaysian Cancer Statistics – Data and Figure Peninsular Malaysia.*

Ni Yongshen. (2005). Fuzzy Correlation and Regression Analysis. *PhD Thesis of University of Oklahoma Graduate College*, UMI number:3163014.

Nishihara Reiko, Wu Kana and Lochhead Paul. (2013). Long Term Colorectal Cancer Incidence and Mortality After Lower Endoscopy. *The New England Journal of Medicine*, (369), No. 32, 1095-1105.

Nwoye, E., Khor L.C., Dlay S.S. and Woo W.L. (2008). *Spectral and Statistical Features in Fuzzy Neural Expert Machine for Colorectal Adenomas and Adenocarcinoma Classification.*

Obi J. C. and Imianvan A. A.. (2012). Fuzzy Neural Approach for Colon Cancer Prediction. *Scienta Africana*, (11), No.1, 65-76.

Pedryoz, W. (2005). Knowledge-based clustering: From data to information granules. *John Wiley & Sons*.

Pereira M. Graca, Ana Paula Figuiredo and Frank D. Fincham. (2011). Anxiety, Depression, Traumatic Stress and Quality of Life in Colorectal Cancer after Different Treatments: A Study With Portuguese Patients and Their Partners. *European Journal of Oncology Nursing,* 1-6.

Phillips L. Karon and Smith Lee Matthew. (2013). Correlates of Initiating Colorectal Cancer Screening Beginning at Age 50. *J Community Health*,(38), 23-30.

Qureshi Mohammed Akhtar, Mahendra Raj and Jayaram Menon. (2012). Screening for Colorectal Cancer in Malaysia Consensus/ Clinical Practice Guidelines. *Academy of Medicine, Malaysia,* 2-12.

Rajeswari K. and Vaithiyanathan V.(2011). Fuzzy Based Modeling for Diabetic Diagnosis Decision Support using Artificial Neural Network. *International Journal of Computer Science and Network Society,* (11), No. 4, 126-130.

Roslani April Camilla, Taufiq Abdullah and Kulenthran Arumugam. (2012). Screening for Colorectal Neoplasias with Fecal Occult Blood Tests: False-positive Impact of Non-Dietary Restriction. *Asian Pacific Journal of Cancer Prevention*, (13), 237-241.

Rousseeuw, Peter J., Annick M. Leroy. (1987). Robust regression and Outlier Detection. *John Wiley & Sons.*

Rusiman Saifullah, Robiah Adnan, Efendi Nasibov and Kavikumar Jacob.(2012). Adjustment of an Intensive Care Unit (ICU) Data in Fuzzy C-Regression Models. *Journal of Science and Technology*, 4 (2), 99-108.

Second Paper of the National Cancer Registry.(2003*). Case Studies on Decision for Cervical Cancer Screening among Working Women.*

Shchneider, J., Bitterlich, N. and Schulze.(2005). Improved Sensitivity In The Diagnosis Of Gastro-Intestinal Tumours By Fuzzy Logic-Based Tumour Marker Profiles Including The Tumour M2-PK. *Anticancer Research,*(25), 1507-1516.

Stefani De Eduardo, Hugo Deneo-Pellegrini and Alvaro L Ronco. (2011). Dietary Patterns and Risk of Colorectal Cancer: a Factor Analysis in Uruguay. *Asian Pacific J Cancer Prev,*(12), 753-759.

Taheri S. Mahmoud. (2003). Trends in Fuzzy Statistics.*Austrian Journal Of Statistics,*(32), Number 3, 239-257.

Tanaka, H. (1987). Fuzzy Data Analysis by Possibilities Linear Models. *Fuzzy Set Syst*, (24), 363-375.

Tanaka Hideo, Uejima Satoru and Asai Kiyoji. (1982). Linear Regression Analysis with Fuzzy Model. *IEEE Transactions On Systems, Man and Cybernetics*, SMC-12, 903-907.

Udupa Jayaram K., Odhner Dewey and Zhao Liming. (2014). Body-Wide Hierarchical Fuzzy Modeling, Recognition and Delineation of Anatomy in Medical Images. *Medical Image Analysis*, (18), 752-771.

Vanagunas. (2011). Role of Endoscopy in the Staging and Management of Colorectal Cancer. *Gastrointestinal Endoscopy*,(78), No. 1, 8-12.

Varela Alejandro, Lina Jandrof and Katherine DuHamel. (2010). Understanding Factors Related to Colorectal Cancer (CRC) Screening Among Urban Hipanics: Use of Focus Groups Methodology. *J Cane Edu*, (25), 70-5.

Wendy and Radzi. (2008). Editorial Cell Therapy Centre, Universiti Kebangsaan Malaysia Medical Centre.*Med J Malaysia*, (63), No 4, 57-58.

World Health Organization. (2012).*Publications of the World Health Organization are available on the WHO web site (www.who.int).*

World Health Organization Data. (2010).*Publications of the World Health Organization are available on the WHO web site (www.who.int).*

Xu Ye, Wen Jingran and Liu Lei.(2009). Comparison of AdaBoost and Logistic Regression for Detecting Colorectal Cancer Patients with Synchronous Liver Metastasis.*IEEE Advancing Technology and Human Huminity,* 4244-4764.

Yilmaz Atinc and Ayan Kirsat. (2013). Cancer risk analysis by fuzzy logic approach and performance status of the model. *Turkish Journal of Engineering & Computer Science*, (21), 897-912.

Yusoff Harmy Mohamed, Norwati Daud, Norhayati Mohd Noor and Amry Abdul Rahim. (2012). Participants and barriers to colorectal cancer screening in Malaysia. *Asian Pacific J Cancer Prev*, (13), 3983-3987.

Zadeh, L.A.. (1965). Fuzzy sets. *Inform Control*,(8), 338-358.

Zhang Wei. (2004). Hospital Quality and Patient Trust in Care for Colorectal Cancer. *Certificate Degree of Doctor Philosophy Harvard University, Cambridge, Massachusetts, United States.*

Zolfaghari Zahra Sadat, Mohebbi Mohebbeat and Najariyan Marzieh. (2014). Application of fuzzy linear regression method for sensory evaluation of fried donut. *Applied Soft Computing,* (22), 417-423.