

**MODEL STATISTIK BERSEPADU
PELBAGAI ARAS BAGI TRIGLISERIDA
DAN TEKANAN DARAH MANUSIA:
KAJIAN KES DI HOSPITAL UNIVERSITI
SAINS MALAYSIA**

MOHAMAD SHAFIQ BIN MOHD IBRAHIM

**SARJANA SAINS MATEMATIK
UNIVERSITI MALAYSIA TERENGGANU**

2015

**MODEL STATISTIK BERSEPADU PELBAGAI
ARAS BAGI TRIGLISERIDA DAN TEKANAN
DARAH MANUSIA: KAJIAN KES DI HOSPITAL
UNIVERSITI SAINS MALAYSIA**

MOHAMAD SHAFIQ BIN MOHD IBRAHIM

**Tesis Ini Diserahkan Untuk Memenuhi Keperluan
Sarjana Sains Matematik
Pusat Pengajian Informatik Dan Matematik Gunaan
Universiti Malaysia Terengganu**

Ogos 2014

DEDIKASI

Tesis ini ditujukan khas untuk keluarga tersayang terutamanya kedua Ibu Bapa saya

“Mohd Ibrahim Bin Kadiasa dan Maharifa Beve”

serta ahli keluarga saya iaitu

“Seeni Nor Ashikin, Mohamad Hafizullah,

Mohd Khairul Ameen, Noorayah Beham,

Mohamad Rejabul Ameen dan Noor Shahirah”

yang banyak memberi inspirasi, kekuatan dan kasih sayang yang tidak terhingga

dalam menyiapkan tesis Sarjana Sains Matematik.

Abstrak Tesis Yang Dikemukakan Kepada Senat Universiti Malaysia Terengganu
Sebagai Memenuhi Keperluan Untuk Ijazah Sarjana Sains

**MODEL STATISTIK BERSEPADU PELBAGAI ARAS BAGI
TRIGLISERIDA DAN TEKANAN DARAH MANUSIA: KAJIAN KES DI
HOSPITAL UNIVERSITI SAINS MALAYSIA**

MOHAMAD SHAFIQ BIN MOHD IBRAHIM

Ogos 2014

Penyelia Utama : Profesor Madya Dr. Wan Muhamad Amir, Ph.D

Penyelia Bersama : Dr. Nurfadhlina Abdul Halim, Ph.D

Pusat Pengajian : Informatik dan Matematik Gunaan

Trigliserida ialah sejenis lemak yang boleh didapati di dalam saluran darah manusia. Sebahagian lemak ini dihasilkan daripada hati dan sebahagian lagi diperolehi daripada sumber makanan yang berkanji dan manis. Selain daripada itu, lemak ini juga akan terhasil apabila makanan yang telah dimakan tidak digunakan sepenuhnya oleh tubuh. Hal ini kerana makanan yang telah diambil akan ditukarkan kepada bentuk trigliserida dan kemudiannya disimpan didalam bentuk sel-sel lemak sekitar tubuh badan. Sel-sel lemak tersebut kemudiannya boleh digunakan sebagai tenaga oleh tubuh badan mengikut keperluan tubuh badan. Namun demikian, kandungan lemak yang berlebihan akan meningkatkan risiko penyakit pada arteri dan juga jantung. Menyedari tentang kepentingan aras trigliserida dalam badan, satu kajian telah dijalankan bagi

mengenal pasti tahap trigliserida dalam tiga fasa utama tekanan darah iaitu normal, pra-hipertensi, hipertensi dan keseluruhan tekanan darah pesakit dengan menggunakan pendekatan ANOVA, model regresi linear berganda dan model regresi logistik berganda. Seterusnya, kajian simulasi Monte Carlo dilaksanakan bagi menentukan pengesahan model yang terbaik antara model regresi linear berganda, model regresi logistik berganda dan model gabungan. Kaedah penganggaran berangka digunakan bagi membuat perbandingan di antara kedua-dua model iaitu ralat min kuasa dua (MSE), ralat punca min kuasa dua (RMSEA), ralat min mutlak (MAE) dan peratus ralat min mutlak (MAPE). Keputusan ANOVA mendapati bahawa tekanan sistolik darah menunjukkan bahawa terdapat perbezaan yang signifikan antara ketiga-tiga fasa utama tekanan darah adalah signifikan ($F(2,997) = 3.595; p = 0.028$). Min ralat bagi peratus mutlak bagi model regresi linear berganda ialah $MAPE_1 = 0.001204$, $MAPE_2 = 0.000718$, $MAPE_3 = 0.000615$ dan $MAPE_4 = 0.000118$. Manakala, min ralat bagi peratus mutlak bagi model regresi logistik berganda ialah $MAPE_1 = 0.004237$, $MAPE_2 = 0.002706$, $MAPE_3 = 0.002517$ dan $MAPE_4 = 0.001017$. Seterusnya, min ralat bagi peratus mutlak bagi model gabungan ialah $MAPE_1 = 0.001236$, $MAPE_2 = 0.000724$, $MAPE_3 = 0.000621$ dan $MAPE_4 = 0.000141$. Berdasarkan nilai ralat bagi min peratus mutlak mendapati model regresi linear berganda mempunyai ralat yang lebih kecil berbanding model regresi logistik berganda dan model gabungan. Oleh itu, model regresi linear berganda merupakan model yang lebih sesuai untuk digunakan bagi mendapatkan keputusan yang lebih tepat.

Abstract of Thesis Presented To The Senate of Universiti Malaysia Terengganu In
Fulfillment of The Requirement for The Degree of Master of Science

**MULTILEVEL INTEGRATED STATISTICAL MODEL FOR
TRIGLYCERIDES AND HUMAN BLOOD PRESSURE: A CASE STUDY IN
HOSPITAL UNIVERSITY SAINS MALAYSIA**

MOHAMAD SHAFIQ BIN MOHD IBRAHIM

August 2014

Main Supervisor : Associate Professor Dr. Wan Muhamad Amir, Ph.D

Co-Supervisor : Dr. Nurfadhline Abdul Halim, Ph.D

School : Informatic and Applied Mathematics

Triglycerides is a type of fat which is found in the bloodstream .The fat is partly derived from the liver and partly derived from starchy and sugary foods. In addition, this fat will be produced when the food that is consumed is not fully utilized by the body. The food then will be converted into triglycerides and stored in the form of fat cells around the body. Fat cells can then be used as energy by the body according to its needs. However, the excessive fat content in the body will increase the risk of disease in the arteries and heart. Recognizing the importance of the level of triglycerides in the body, a study was conducted to identify the level of triglycerides in the three main phases of blood pressure, that are normal blood pressure, pre-hypertension, hypertension and overall blood pressure of patients by using ANOVA approach, multiple linear regression model and multiple logistic regression model. Next, Monte

Carlo simulation study was conducted to determine the best model verification between the multiple linear regression model, the multiple logistic regression model and combination model. Numerical estimation methods was used to make the comparison between the three models namely the Mean Squared Error (MSE), Root Mean Square Error (RMSEA), the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). ANOVA results found that systolic blood pressure were showing highly significant differences ($F(2, 997) = 3.595, p = 0.028$) across the three distinct phases of blood pressure. The mean absolute percentage error for the multiple linear regression model was $MAPE_1 = 0.001204$, $MAPE_2 = 0.000718$, $MAPE_3 = 0.000615$ and $MAPE_4 = 0.000118$. Whereas, the mean absolute percentage error for the multiple logistic regression model was $MAPE_1 = 0.004237$, $MAPE_2 = 0.002706$, $MAPE_3 = 0.002517$ and $MAPE_4 = 0.001017$. Next, the mean absolute percent error for the combined model is $MAPE_1 = 0.001236$, $MAPE_2 = 0.000724$, $MAPE_3 = 0.000621$ and $MAPE_4 = 0.000141$. Based by the mean absolute percentage error it was found that multiple linear regression model reported a smaller error than the multiple logistic regression model and combination model. Therefore, multiple linear regression model was the most appropriate model to use to obtain more precise results.