MODEL OF RISK PROBABILITY OF CHRONIC KIDNEY DISEASE IN NORTH CENTRAL REGION OF SRI LANKA

M.D.Nandana Gunaratne

(8207)

Degree of Masters of Science

Department of Mathematics

University of Moratuwa
Sri Lanka

February 2015

Declaration page of the candidate

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

Declaration of the Supervisor

The above candidate has carried out research for the Masters Dissertation under my supervision.

Name of the Supervisor : Mr. T. M. J. A. Cooray
Faculty of Engineering
Department of Mathematics
University of Moratuwa

Signature of the supervisor: Date

ABSTRACT

The chronic kidney disease of unknown etiology in farmers in North Central region has become an emerging health problem in Sri Lanka. Lack of scientifically proven experiments of epidemiological studies on the etiology is one of the major problems in solving this issue in Sri Lanka. Therefore, identifying the independent preventable risk factors mainly related to the occupation may helps in decreasing the number of patients suffering from CKD-U and slowing its progression. The objective of the study is to identify the epidemiological characteristics and potential risk factors related to agricultural activities in the development of CKD-U by using a case control study and modeling the risk probability of the disease occurrence.

The study was undertaken in Madawachchiya, Padaviya and Kebithigollawa area and two hundred and seventy four (274) patients with CKD-U (cases) and two hundred and seventy four (274) healthy individuals were selected as controls for the age and sex 1: 1 matched case control study. The relative risk of each factor was compared in terms of odds ratios (OR) and 95% confidence intervals (CI) by applying the conditional logistic regression model. The risk probability was calculated with the above model in cases and control separately and compared.

Involvement in agricultural activities, low protective measures against agrochemicals, and cultivating lands without labour exchange were identified as significant risk factors for the disease. Smoking and family history of CKD-U, drinking water source (shallow wells) and history of snake bite were identified as other life style related risk factors for the disease occurrence. The risk probability can be used as an index of the disease diagnosis and the receivers operating characteristic (ROC) curve plots the true positive fraction (TPF) against the false positive fraction (FPF) for different choice of cut off ranges. Therefore involvement in agricultural activities is significantly related for the disease. Low levels of protective measures in the application of agrochemicals indicate the importance to educate the farmers on protective measures of agrochemical applications.

ACKNOWLEDGEMENT

The contributions of the following are gratefully acknowledged:

At the time of putting demise to my thesis dissertation, on top of every one it is with heartfelt gratitude that I would like to extend my oozing out pleasures to my supervisor and teacher Mr. T.M.J.A. Cooray, Senior Lecturer, Department of Mathematics, University of Moratuwa

I am very gratitude to Mr. K.B. Jayasekara, Faculty of Medicine, University of Peradeniya who made his effort to collect data and technical assistance.

My sincere thanks goes to the Director and medical officers of the Department of Health Services North Central Provincial for providing us details of patients and medical advises.

Last but not least, I desire my deepest gratitude to my wife son and parents for their support and encouragement, to get through this agonizing period in the most positive way.

I also place on record, my sense of gratitude to one and all who directly or indirectly have lent their helping and hand in this venture.

CONTENT

Declaration of the candidate & supervisor i
Abstract ii
Acknowledgement iii
Content iv
List of figures v
List of tables vi
Abbreviations vii

1.0 INTRODUCTION 1

1.1 Chronic kidney disease 1

1.2 The chronic Kidney disease in Sri Lanka 1

1.2.1 Etiology of the chronic Kidney disease in Sri Lanka 2

1.3 Problem statement 5

1.4 Objective of the study 5

2.0 LITERATURE REVIEW 7

2.1 Chronic kidney disease 7

2.2 The Chronic kidney disease in Sri Lanka 7

2.3 The Case- Control Study 13

2.3.1 Matching 13

2.4 The logistic regression model 15

2.4.1 Conditional Logistic Regression 16

2.5 Estimation 17

2.6 ROC curves 17

3.0 METHODOLOGY 19

3.1 Study design 19

3.2 Sample size 19

3.3 Data and data collection 20

3.4 Data analysis 22

3.4.1 The Linear logistic model 23

3.4.1.1 Estimation.....24

3.4.2 Measures of association between diseased and exposure 25

3.4.3 Confidence Interval 26

3.4.4 Testing for the significance of coefficients.....27

3.4.4.1 The Likelihood Ratio Test 27

3.4.4.2 Wald Test 27

3.4.5 Computer software 28

4.0 RESULTS 29

4.1 Univariate analysis 29

4.2 Multivariate analysis 31

5.0 DISCUSSION AND CONCLUSION 37

REFERANCES ..... Error! Bookmark not defined.

Annex 01……45

Annex 02……...46

Annex 03……49

LIST OF FIGURES

	Page
Figure 2.1: Prevalence of CKD-U in the North Central Area of Sri Lanka	8
Figure 4.1 : The "Model Fit Statistics" table of the SAS output	32
Figure 4.2: Box plot diagram of risk probabilities of the disease occurrence in case and control	35
Figure 4.3: The distribution of risk probability for diseased and non-diseased samples	35
Figure 4.4: The ROC curve for the diagnosis of the disease with risk probability	36

LIST OF TABLES

	Page
Table 4.1: Odds ratios of agricultural and life style related risk factors	31
Table 4.2 : Analysis of Maximum Likelihood Estimates	33
Table 4.3: Odds ratios of agricultural and life style related risk factors in conditional logistic regression	34
Table 4.4: TPF and FPF at each cut off level	36

ABBREVIATIONS

AIC	Akaike Information Criterion
CI	Confidence Interval
CKD	Chronic Kidney Disease
CKD-U	Chronic Kidney Disease of Unknown Etiology
CRF	Chronic Renal Failure
CVD	Cardio Vascular Disease
FNF	False Negative Fraction
FNR	False Negative Rate
FPF	False Positive Fraction
FPR	False Positive Rate
GDP	Gross Domestic Production
ha	Hectare
ML	Maximum Likelihood
MRFIT	Multiple Risk Factor Intervention Trial
N	Sample size
NCR	North Central Region
OR	Odds Ratio
P	Probability
ROC	Receivers Operating Characteristic Curve
PREVEND	Prevention of Renal and Vascular End Stage Disease
TNF	True Negative Fraction
TNR	True Negative Rate
TPF	True Positive Fraction
TPR	True Positive Rate
SAS	Statistical Analysis System
SC	Schwarz Criterion