Sudan University of Science and Technology

Collage of Graduate Studies

Modelling and Forecasting Age-Specific Cancers Mortality Rate Using Lee-Carter Model

Case study: Egypt

Time period: 2001-2014

النمذجة والتنبؤ بمعدل الوفيات العمرية لسرطانات محددة بإستخدام نموذج لي-كارتر

دراسة حالة: مصر

الفترة الزمنية: 2001 - 2014

A Thesis Submitted in fulfillment of Requirement for the Ph.D. in statistics

By

Arwa Elsiddig Elfaki Elhassan

Supervisor:

Dr. Hamza Ibrahim Hamza Osman

Co-supervisor:

Dr. Manahil Sid Ahmed Mustafa

2017-1439

(واللَّهُ خَلَقَكُمْ ثُمَّ يَتَوَفَّاكُمْ وَمِنْكُمْ مَنْ يُرَدُّ إِلَى أَرْذَلِ الْعُمُرِ لِكَيْ
لَا يَعْلَمَ بَعْدَ عِلْمٍ شَيْئًا إِنَّ اللَّهَ عَلِيمٌ قَدِيرٌ)

سورة النحل الآية (70)

Dedication

I dedicated this thesis to :

My Mother, my Father's Soul , my husband, daughters and sons , all my family and all the people who has supported me throughout this times. I will always appreciate all what they have done.

Acknowledgements

I would like to express my deepest appreciation to my supervisor Dr. Hamza Ibrahim Hamza for his guidance and sharing of his opinions and experiences, which gave me the possibility to complete this thesis, and also for being devoted and patient. Also to thank my co-supervisor Dr. Manahil Sid Ahamed Mustafa.

I would like also to thank my family for their generous support , throughout my entire life and particularly during the time of the Ph.D. My great thank goes to all people who took part in making this research, especial thank goes to my mother and husband of their unconditional love. Also my grateful to Sudan Atomic Energy Commission to give me opportunity to study doctorate program and the Sudan University of Science and Technology , Collage of Science, Statistics Department and all the staff**.**

Abstract

In 1992, Lee and Carter proposed a method which combines demography and stochastic to model and forecast the mortality rates, which became the reference and a leading statistical model. In this study we identified Cancer as characterized by out-of-control cell growth and it is the second leading cause of death after Ischemic heart disease. The problem of this study is that the cancer has the highest death rate among other diseases and it's treatment required financial resources that strain the state treasury, more over the absence of data reduced the performance of the model . The importance of this study is to help the governments ,voluntary organizations and health sector to make plans and researches scientifically. According to this importance the aims of this study is to use original Lee-Carter model to model and forecast age-specific cancer mortality rate for three types of cancer (Oral, Lung and Colon) for period 2015 to 2020. The model's parameters estimated by Singular value Decomposition (SVD) and Maximum Likelihood Estimation (MLE), and used Auto Regressive Integrated Moving Average (ARIMA) Random Walk with drift (0,1,0) to forecast mortality index for Egyptian male and female based on five-year data aggregation that obtained from World Health Organization (WHO) for the period 2001-2014. The results obtained by using different statistic packages R ,ilc, Demography and forecast packages. Our findings showed that the SVD is better for male with error (ME=0.00016, MSE=25208), while for female the SVD is better with error (ME=0.02856, MSE=0.32310) for oral cancer. while the MLE is better for male with error (ME=0.00714, MSE=0.12385), and the SVD is better for female with error (ME=0.00523, MSE=0.08022) for lung cancer. while the MLE is better for male with error (ME=0.00506,

MSE=0.11065), and the SVD is better for female with error (ME=-0.00401, MSE=0.13561) for colon cancer. Also the results showed that the lung cancer has highest mortality rate and it is 76.27 per 100.000 in year 2020 in age-group (70-74) for male then colon and it is 27.91 in year 2020 in age-group (70-74) for male, after that oral cancer rate and it is 3.11 per 100.000 in year 2015 in age-group (70-74) for male. The study came out with numbers of recommendations from them the most importance are to apply Lee-Carter method to modeling and forecasting age-specific mortality rate and SVD to estimate the model's parameters, and to have care and accuracy when registering data. Health sector must make plans and programs to reduce the cancer mortality rate especially for male.

المستخـــــــلص

في سنة 1992 قدم الباحثان Lee و Carter نموذج عشوائي للتنبؤ بمعدل الوفيات وهو طريقة استقرائية لعرض الوفيات والتنبؤ بها ، واصبح النموج مرجعاً ورائداً للنماذج الإحصائية. ومن خلال هذه الدراسة نعرف مرض السرطان بوصفه بأنه نمو للخلايا وإنتشارها بشكل لا يمكن التحكم بها، وبالإضافة إنه السبب الثاني للوفاة في العالم بعد مرض القلب . وتتمثل مشكلة الدراسة في إن السرطان له معدل وفيات أعلى من بقية الأمراض وعلاجه يحتاج إلى موارد مالية ترهق خزينة الدولة ،كما أن عدم توفر البيانات يقلل من دقة النموذج للتنبؤ.كما تتمثل أهمية الدراسة بأنها تساعد الحكومات والمنظمات الطوعية والقطاع الصحي في وضع الخطط والبحوث بصورة علمية، وبناءاً على هذه الأهمية تهدف هذه الدراسة إلى التعرف على نموذج لي-كارتر الأصلي للتنبؤ بمعدل وفيات السرطان (الفم والبلعوم ، الرئة والقولون) ومن ثم التنبؤ بمعدل الوفيات للفترة الزمنية 2015-2020 . تم إستخدام المنهجية العلمية في تقدير معالم النموذج وهي طريقة تحليل القيمة المفردة و طريقة الإمكان الأعظم ومن ثم استخدام نموذج الإنحدار الذاتي التكاملي-المتوسط نموذج المشي العشوائي بإنجراف (0,1,0) للتنبؤ بدليل الوفاة إعتماداً على الفئات العمرية المتحصل عليها من منظمة الصحة العالمية في الفترة الزمنية 2001-2014 لسكان مصر (الذكور، الإناث). وتم الحصول على النتائج بإستخدام مجموعة من الحزم الإحصائية منها ilc، R demography, وforecast. ومنها توصلت الدراسة إلى عدد من النتائج أهمها أن طريقة تحليل القيمة المفردة تعطي نتائج أفضل للذكور (بمتوسط الخطأ=00016. ومتوسط مربع الخطأ=0.25208، بينما للإناث (بمتوسط الخطأ=02856. ومتوسط مربع الخطأ=0.32310) بالنسبة لسرطان الفم والبلعوم، وطريقة الإمكان الأعظم تعطي نتائج أفضل للذكور(بمتوسط

الخطأ =0.00714 ومتوسط مربع الخطأ =0.12385)، بينما طريقة القيمة المفردة تعطي نتائج أفضل للإناث (بمتوسط الخطأ =0.00523 ومتوسط مربع الخطأ =0.08022) لسرطان الرئة. وطريقة الإمكان الأعظم تعطي نتائج أفضل للذكور(بمتوسط الخطأ =0.00506 ومتوسط مربع الخطأ =0.11065))، وطريقة القيمة المفردة تعطي نتائج أفضل للإناث (بمتوسط الخطأ =0.00401- ومتوسط مربع الخطأ =0.13561) لسرطان القولون. كما أظهرت الدراسة ان هناك إرتفاع في معدل الوفيات العمرية على مر السنين لكل أنواع السرطانات وخاصة سرطان الرئة الذي لديه إعلى معدل وكان مقداره 76.27 لكل 100.000 في سنة 2020 في الفئة العمرية (70-74) للذكور ويليه القولون وكان مقداره 27.91 لكل 100.000 في سنة 2020 في الفئة العمرية (70-74) ، ثم الفم والبلعوم وكان مقداره 3.11 لكل 100.000 في سنة 2015 في الفئة العمرية(70-74). خرجت الدراسة بعدد من التوصيات أهمها إستخدام طريقة Lee-Carter في التنبؤ بمعدل الوفيات العمرية . إستخدام القيمة المفردة في تقدير المعالم لما لها من متوسط خطأ قليل كما أوضحت النتائج. الإهتمام والدقة في تسجيل البيانات . على القطاع الصحي عمل خطط وبرامج لتقليل معدل الوفيات وخاصة للذكور.

TABLE OF CONTENTS

NO.	Content	Page
	Aya	I
	Dedication	II
	Acknowledgements	III
	Abstract (English)	IV
	Abstract (Arabic)	VI
	Table of Contents	VIII
	List of tables	X
	List of figures	XIII
Chapter one (Introduction)		
1.1	Preface	1
1.2	Research Problems	2
1.3	Research Importance	3
1.4	Research Objectives	3
1.5	Research Hypothesis	3
1.6	Research Methodology	4
1.7	Data sources	4
1.8	Research Limited	5
1.9	Previous Studies	5
1.10	Research Organization	15
Chapter Two (Lee-Carter Model)		
2.1	Preface	16
2.2	Measures of Mortality Rate	17
2.3	Mortality Models Techniques	18
2.4	Criteria for Term Structure of Mortality Models	18
2.5	Mortality Forecasting Methods In The Past	18
2.6	Lee-Carter Model	20
2.7	Times Series	37
Chapter Three (Cancer)		
3.1	Preface	41
3.2	The Genetic Bases of Cancer	42
3.2.1	Oral Cancer	43
3.2.2	Lung Cancer	49
3.2.3	Colon Cancer	55
Chapter Four(Application)		
4.1	Preface	61
4.2	Results and Interpretations	61

Chapter Five (Conclusions and Recommendations)		
5.1	Conclusions	101
5.2	Recommendations	103
	References	105
	Appendix	

LIST OF TABLES

NO	Table	Page
(3.1)	Determining Tumor Characteristics for oral cancer.	46
(3.2)	Regional lymph node status for oral cancer	46
(3.3)	Determining Metastatic Status for oral cancer	47
(3.4)	The stage of oral cancer.	47
(3.5)	Describe the treatment for every stage of Oral cancer.	49
(3.6)	Determining Tumor Characteristics for lung cancer .	53
(3.7)	Regional lymph node status for lung cancer.	53
(3.8)	Determining Metastatic Status for lung cancer.	53
(3.9)	Stage of Non SCLC.	54
(3.10)	Determining Tumor Characteristics for colon cancer	58
(3.11)	Regional lymph node status for colon cancer .	58
(3.12)	Determining Metastatic Status for colon cancer .	59
(3.13)	Stage of colon cancer.	59
(3.14)	Description of Treatment for Colon cancer.	60
(4.1)	Estimation of a_x by SVD for oral cancer.	62
(4.2)	Estimation of b_x by SVD for oral cancer.	63
(4.3)	First and second estimation of k_t by SVD for oral cancer.	64
(4.4)	Estimation of a_x by MLE for oral cancer.	65
(4.5)	Estimation of b_x by MLE for oral cancer.	66
(4.6)	Estimation for k_t by MLE for oral cancer.	67
(4.7)	Comparison between SVD and MLE for estimation a_x for oral cancer.	68
(4.8)	Comparison between SVD and MLE for estimation b_x for oral cancer.	69
(4.9)	Comparison between SVD and MLE for estimation k_t for oral cancer.	70
(4.10)	Comparison between SVD and MLE for errors based on log mortality rate across ages for oral cancer.	71
(4.11)	Estimation of drift, standard error and errors of RWD(0,1,0) for oral cancer.	72

(4.12)	Forecast mortality index for period 2015–2020 for oral cancer**.**	72
(4.13)	Forecast age-specific mortality rate for period 2015–2020 for oral cancer.	73
(4.14)	Model's forecast errors based on mortality rate across ages for oral cancer .	74
(4.15)	Estimation of a_x by SVD for lung cancer.	75
(4.16)	Estimation of b_x by SVD for lung cancer.	76
(4.17)	First and second estimation of k_t by SVD for lung cancer.	77
(4.18)	Estimation of a_x by MLE for lung cancer.	78
(4.19)	Estimation of b_x by MLE for lung cancer.	79
(4.20)	Estimation for k_t by MLE for lung cancer.	80
(4.21)	Comparison between SVD and MLE for estimation a_x for lung cancer.	81
(4.22)	Comparison between SVD and MLE for estimation b_x for lung cancer.	82
(4.23)	Comparison between SVD and MLE for estimation k_t for lung cancer.	83
(4.24)	Comparison between SVD and MLE for Errors for lung cancer.	84
(4.25)	Estimation of drift, standard error and errors of RWD (0,1,0) for lung cancer.	85
(4.26)	Forecast mortality index for period 2015–2020 for lung cancer**.**	85
(4.27)	Forecast age-specific mortality rate for lung cancer per 100.000 for period 2015 – 2020.	86
(4.28)	Model's forecast errors based on mortality rate across ages for lung cancer .	87
(4.29)	Estimation of a_x by SVD for colon cancer.	88
(4.30)	Estimation of b_x by SVD for colon cancer.	89
(4.31)	First and second estimation of k_t by SVD for colon cancer.	90
(4.32)	Estimation of a_x by MLE for colon cancer.	91
(4.33)	Estimation of b_x by MLE for colon cancer.	92
(4.34)	Estimation for k_t by MLE for colon cancer.	93
(4.35)	Comparison between SVD and MLE for estimation a_x for colon cancer.	94
(4.36)	Comparison between SVD and MLE for estimation b_x for colon cancer.	95

(4.37)	Comparison between SVD and MLE for estimation k_t for colon cancer.	96
(4.38)	Comparison between SVD and MLE for Errors for colon cancer.	97
(4.39)	Estimation of drift, standard error and errors of RWD (0,1,0) for colon cancer.	98
(4.40)	Forecast mortality index for period 2015–2020 for colon cancer**.**	98
(4.41)	Forecast age-specific mortality rate for colon cancer per 100.000 for period 2015–2020.	99
(4.42)	Model's forecast errors based on mortality rate across ages for colon cancer .	100

LIST OF FIGURES

NO	Figure	Page
(4.1)	General pattern of mortality a_x by SVD for oral cancer.	62
(4.2)	General pattern of mortality b_x by SVD for oral cancer.	63
(4.3)	General pattern of k_t 2001-2014 by SVD for oral cancer.	64
(4.4)	General pattern of mortality a_x by MLE for oral cancer.	65
(4.5)	General pattern of mortality b_x by MLE for oral cancer.	66
(4.6)	General pattern of k_t 2001-2014 by MLE for oral cancer.	67
(4.7)	Comparison between SVD and MLE for estimation a_x for oral cancer.	69
(4.8)	Comparison between SVD and MLE for estimation b_x for oral cancer.	70
(4.9)	Comparison between SVD and MLE for estimation k_t 2001-2014 for oral cancer.	71
(4.10)	Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for oral cancer.	72
(4.11)	Forecast Age-specific cancer mortality rate from 2015-2020 for oral cancer .	74
(4.12)	General pattern of mortality a_x by SVD for lung cancer.	75
(4.13)	General pattern of mortality b_x by SVD for lung cancer.	76
(4.14)	General pattern of k_t2001-2014 by SVD for lung cancer.	77
(4.15)	General pattern of mortality a_x by MLE for lung cancer.	78
(4.16)	General pattern of mortality b_x by MLE for lung cancer.	79
(4.17)	General pattern of k_t 2001-2014 by MLE for lung cance r.	80
(4.18)	Comparison between SVD and MLE for estimation a_x for lung cancer.	82
(4.19)	Comparison between SVD and MLE for estimation b_x for lung cancer.	83
(4.20)	Comparison between SVD and MLE for estimation k_t 2001-2014 for lung cancer.	84
(4.21)	Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for lung cancer.	85
(4.22)	Forecast age-specific cancer mortality rate 2015-2020 for lung cancer .	87

(4.23)	General pattern of mortality a_x by SVD for colon cancer.	88
(4.24)	General pattern of mortality b_x by SVD for colon cancer.	89
(4.25)	General pattern of k_t2001-2014by SVD for colon cancer.	90
(4.26)	General pattern of mortality a_x by MLE for colon cancer.	91
(4.27)	General pattern of mortality b_x by MLE for colon cancer.	92
(4.28)	General pattern of k_t 2001-2014 by MLE for colon cancer.	93
(4.29)	Comparison between SVD and MLE for estimation a_x for colon cancer.	95
(4.30)	Comparison between SVD and MLE for estimation b_x for colon cancer.	96
(4.31)	Comparison between SVD and MLE for estimation k_t 2001-2014 for colon cancer.	97
(4.32)	Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for colon cancer.	98
(4.33)	Forecast age-specific cancer mortality rate 2015-2020 for colon cancer .	100

CHAPTER ONE

(Introduction)

1.1 Preface
1.2 Research Problems
1.3 Research Importance
1.4 Research Objectives
1.5 Research Hypothesis
1.6 Research Methodology
1.7 Data sources
1.8 Research Limited
1.9 Previous Studies
1.10 Research Organization

1.1 Preface :

The word "mortality" came from the Latin word "mors" which means (death) [1].

Mortality statistics provide a valuable measure for assessing community health status, where the importance of mortality statistics came from both the significance of death in an individual's life as well as their potential to improve the public's health, providing that it's systematically to assess and monitor the health status of a whole community. Mortality statistics are often used as a cornerstone in formulating health plans and policies to prevent or reduce premature mortality and improve our quality of life. Mortality data are some of the best sources of information about the health of living communities, they provide a snapshot of current health problems, suggest persistent patterns of risk in specific communities, and show trends in specific causes of death over time. Many causes of death are preventable or treatable, therefore, warrant the attention of public health prevention efforts, so public health administration should strongly depends on the study of mortality, specifically done for statistics on death in the population cross –classified by age, sex and the cause of death are of great value for the formulation, implementation and evaluation of public health programs.

Mortality modeling has been used for many long time, there are many models proposed since Gompertz published his law of mortality in 1825 [2] .The earliest models were simple and they were focused on producing mathematical functions to fit observed mortality rates. However, over the last 20 to 30 years the development of stochastic mortality models has been very rapid in terms of both structure and statistical techniques used to fit the models. The Lee-Carter model [3] is probably the best known

method for mortality forecasting new a days, among other models which have been proposed.

Forecasting of cancer mortality rate plays an integral role in planning and research and it can be very valuable as a tool to predict cancer burden. Considering the fact that the cancer illness brings huge expenses in health involving diagnosis, treatment, research, loss of productivity due to sick leaves ,so future information about cancer mortality is essential for Public Heaths. These information are also important to efficiently organize cancer screening programs and to prioritize prevention activities.

A wide range of methods used for forecasting cancer mortality rate has been developed . Many statistical software packages, such as Nordpred and the iterative Lee-Carter package, for forecasting age-specific cancer incidence and mortality data implicitly assume that data are aggregated to five-year intervals on the time-scale (periods) [4].

The public health in the developed countries used different models to forecast cancer mortality but they work at regional level. According to studies the performance of the model depends on the number of observed cases. Moreover, the same models can show different behavior in different countries. For example, for testis, thyroid and ovary cancers, different performance is observed with Canadian and American data [5].

Making a cancer mortality forecasting has difficulties and uncertainties, since the usual method used to construct a model which fits the historical data, so the consistency in data collection methods and definitions, within the period on which the model is based. This model is then used to extrapolate past trends to make future predictions.

1.2 Research Problems:

Cancer presents a global public health problem which extensively affects healthcare costs, because treatment of cancer required financial resources that strain the state treasury contributing to increase the number of

deaths.. In the absence of required data for the dead , this reduced the performance of models to forecast . This study used Lee-Carter model to forecast the mortality rate for coming years to help government , institutions and voluntary organizations to Know the mortality rate scientifically instead of prevailed. There are a few studies applying statistical models to forecast the cancer mortality rate in Arab Countries, almost all the studies on incidence of cancer were conducted by the doctors or who in the field of health.

1.3 Research Importance :

To fill the gap in recent researches on the age-specific cancer mortality rate, and to apply modern statistical models such as Lee-Carter model in our region, and to provide important information that influences practices, policies, and programs that directly affect the health sector.

1.4 Research Objectives:

The aims of this study are:

- To investigate how to apply Lee-Carter method to forecast age-specific cancer mortality rates.
- To investigate how to fit Singular Value decomposition (SVD) and Maximum Likelihood Estimation (MLE) to estimate the model's parameters.
- To Determine the appropriate method to use for estimation the parameters of the model.

1.5 Research Hypothesis:

- If a Singular Value Decomposition (SVD) fit better than Maximum likelihood Estimation (MLE) to estimate the model's parameters for male for all cancer (oral, Lung and colon).

- If a Singular Value Decomposition (SVD) fit better than Maximum likelihood Estimation (MLE) to estimate the model's parameters for female for all cancer (oral, Lung and colon).
- If forecasting age-specific mortality rate performance well for male for all cancer (oral, Lung and colon).
- If forecasting age-specific mortality rate performance well for female for all cancer (oral, Lung and colon).

1.6 Research Methodology:

In this study we used references, books ,articles ,papers and previous studies . We applied the original Lee-Carter method to model and forecast mortality rate cancer (oral, Lung and colon). To estimate the parameters of the Lee-Carter model we used two methods Singular Value Decomposition (SVD) and Maximum Likelihood Estimation (MLE). The comparison of the two methods (SVD, MLE) based on the mean error (ME) and mean square error (MSE) . Once estimated parameter we used ARIM Walk Random Drift (0,1,0) to forecast cancer mortality rate as in the original paper for both sex (male, female) for cancer (oral, lung and colon) separately and performance of forecasting based on mean percentage error (MPE) . The statistical package R, Iterative lee carter package (ilc) [6] ,forecast and demography have been used to execute modeling and forecasting .

1.7 Data Sources:

The data source for this study from the World Health Organization (WHO) [7], which contains number of deaths and population by country, year, sex, age-group and cause of death. The data have been coded appropriately using the International Classification of Diseases (ICD) are available in the database. ICD is recognized in epidemiology, health management and medicine as a benchmark tool used to keep incidence

and prevalence of diseases in population . ICD for Malignant neoplasm of lip, oral cavity and pharynx (oral cancer) is C00-C14, Malignant neoplasm of trachea, bronchus and lung (Lung cancer) is ICD-C33-C34 and Malignant neoplasm of colon (Colon cancer) is ICD-C18.

We obtained six data series deaths of cancer (oral, Lung and Colon) and population by age and year of death for Egyptian (male-female), from the period 2001-2014. The data are aggregated to five-year intervals on the time-scale and they are (5-9, 10-14,..., 70-74). We used Egyptian mortality data because we fund a few data for those diseases from 2008 - 2014 with unequal range in the Radiation & Isotopes Center – Khartoum RICK .Dental Hospital ,and we have no data in other places and this was affected f performance of the model (short period and few data).

1.8 Research Limited :

Place : Egypt.

Time period : 2001 to 2014.

1.9 Previous Studies:

1. Lee R. and carter L. (1992), Modelling and Forecasting US mortality rate. Journal of the American Statistical Association. 87:659-671[3] : They published a modern method for long-run forecasts of the level and age pattern of mortality, based on a combination of statistical time series methods and a simple approach to dealing with the age distribution of mortality. The method described the log of a time series of age-specific death rates as the sum of an age-specific component that was independent of time and another component that was the product of a time-varying parameter reflecting the general level of mortality, and an age-specific component that represented how rapidly or slowly mortality at each age varied when the general level of mortality changed. This model was fitted to historical data from the time period 1933-1987 and projections were

made up to the year 2065. The resulting estimated of the time-varying parameter was then modeled and forecasted as a stochastic time series using Random Walk with drift. From this forecast of the general level of mortality, the actual age-specific rates were derived using the estimated age effects. The forecasts of the various life table functions had probability distributions, so probability intervals can be calculated for each variable and for summary measures such as life expectancy. The projected of life expectancy from 1989 to 1997 matched the actual gain very closely and was nearly twice the gain projected by the Social Security Administration's Office of the Actuary.

2. John R. Wilmoth, (1993), Computational Method of Fitting and Extrapolating the Lee-Carter model of mortality Change. Department of Demography, University of California, Berkeley. Technical Report[8] : He purposed modern techniques Weighed Least Square and Maximum Likelihood Estimation to estimate the parameters of The Lee-Carter model and applied on Japanese women for period 1951-1990, Both techniques had the significant advantage, over the original Lee-Carter Singular Value Decomposition (SVD), that they dialed naturally with the case in which the observed number of deaths was zero, which occurred when analyzing cause specific data and/or when dealing with small countries.

3. Ronald Lee, (2000), The Lee-Carter Method For Forecasting Mortality, With Various Extensions And Applications. North American Acturial Journal. (4,1): 80-91[9]: This paper described the basic Lee-Carter method and discussed the forecasts, extensions, applications, and methodological improvements that had been made in recent years, considered shortcomings of the method, and briefly described how it had been used as a component of more general stochastic population

projections and stochastic forecasts of the finances of the U.S. Social Security system.

4. Lawrence R. Carter and Alexia Prskawetz**,** (2005), Examining Structural Shifts in Mortality Using the Lee-Carter Method. Max Planck Institute for Demographic[10] : They presented an extension of the Lee-Carter method of modeling mortality to examine structural shifts in trajectories of mortality based on Austrian data consisting of 53 years of single-age mortality rates. They used singular value decomposition to estimate parameters. They compared the observed and estimated life expectancy between original Lee-Carter and extension of the Lee-Carter and they found that the extended Lee-Carter method was better to the original Lee-Carter method, particularly for life expectancies at higher ages.

5. Steven Haberman and Maria Russolillo, (2005), Lee-Carter mortality forecasting: application to the Italian population. Actuarial Research Paper No. 167[11] : In this paper they used the Lee-Carter methodology to construct mortality forecasts for the Italian population. The model fitted to the Italian death rates for each gender from 1950 to 2000. A time-varying index of mortality is forecasted in an ARIMA framework and was used to generate projected life tables. In particular they focused on life expectancies at birth and, for the purposed of comparison, they introduced an alternative approach for forecasting life expectancies on a period basis. The resulting forecasts generated by the two methods were then compared. The results showed the life expectancies forecasted under the LC model, with the time-series-based forecast it was different.

6. Booth, Rob J. Hyndman, Leonie Tickle, Piet de Jong, (2006), Lee-Carter mortality forecasting: a multi-country comparison of variants and

extensions. Demographic Research. 15: 289-310[12] : They applied sex specific populations of 10 developed countries using data for 1986–2000 and fitted them in five variants or extensions of the Lee-Carter method original Lee-Carter, the Lee-Miller and Booth-Maindonald-Smith variants, and Hyndman-Ullah and De Jong-Tickle extensions for mortality forecasting. The finding was all variants and extensions were more accurate than the original Lee-Carter method for forecasting log death rates, by up to 61%., and there were no significant differences among the five methods in forecasted accuracy for life expectancy. The indicator is to use different statistical test include t-test they found lee-carter fit better , and MAE the LC performs least well and they used a 2-way ANOVA and they found original LC method was significantly different from all other methods, but the other four methods were not significantly different from each other . They used a 2-way ANOVA model with method and country as factors on the mean absolute errors in life expectancy to test whether the methods were significantly different. There was no significant difference between the five methods ($p = 0.21$) in the accuracy of life expectancy forecasts. The results of this comparative evaluation of forecasts showed that while each of the four variants and extensions was more accurate in forecasting log death rates than the original Lee-Carter method, none was consistently more accurate than the others. They found Hyndman-Ullah and De Jong-Tickle provided the most accurate forecasts of log death rates; however, the differences among the four methods were small and were not significant.

7. Claia Pedroza. (2006). A Bayesian forecasting model: predicting U.S. male mortality. Biostatistics (7,4): 530–550[13] : This article presented a Bayesian approach to forecast mortality rates. Markov chain Monte Carlo methods were used to fit the model and to sample from the posterior

predictive distribution. This paper also showed how to handle missing data and presented some possible extensions to the model, which applied to U.S. male mortality data based on data from 1959–1998. The age groups were 0, 1–4, 5–9, . . . , 105–109, 110+., to forecast 1990–1999. These forecasts were compared to the actual observed values. She fitted and forecasted log-mortality rates using both the original Lee–Carter method and the Bayesian model. The results showed the Bayesian prediction intervals were wider than those obtained from the Lee–Carter method, An extension to the model was also presented and the resulting forecast variability appeared better suited to the observed data.

8. Jenny Zheng Wang, (2007), Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model. Dept. of Mathematical Statistics, Stockholm University[14] : The purposed of his study showed the performance of the predictions would have changed if they had changed the length of the estimation period, and to do that he applied original Lee-Carter model to data from Sweden from 1860-2004 based on a three sub-samples of 1900-2004, 1950-2004 and 1980-2004. The Singular Value Decomposition (SVD) was used to estimate the model's parameters. Identification of a common trend of mortality change had been attempted by fitting a standard Lee-Carter model to different time series (1860-2004, 1900-2004, 1950-2004 and 1980-2204). He concluded by forecasting the mortality rates for 1901-2004 and 1951-2004 . The results indicated that the selection of an appropriate estimation period was important for forecasting mortality. The estimation periods of 1850-1900 and 1900-1950 yield the best forecasting performances for prediction series of 1901-2004 and 1951-2004, and the prediction with short estimation period like 1940-1950 did not work well.

9. Sándor Baran, József Gáll, Márton Ispány**,** Gyula Pap. (2007), Forecasting Hungarian mortality rates using the Lee-Carter method. Acta Oeconomica, (57,1):25–38[15] : A modified version of the Lee–Carter method was applied to forecast mortality rates in Hungary for the period 2004–2040 on the basis of mortality data between 1949 and 2003 both for men and women. Using singular value decomposition to estimate the parameters. The results showed increasing mortality rates for several age categories especially for men between ages 45 and 55. And the Lee–Carter method was successfully applied for Hungarian mortality rate.

10. Marie-Claire Koissi and Arnold F. Shapiro, (2008), The Lee-Carter Model Under The Condition Of Variables Age-Specific Parameters. 43rd Actuarial Research Conference, Regina, Canada[16]: In this paper, They proposed a modification of the Lee-Carter model that accommodated variations in age-specific parameters. They used the weighted least square approach to find the model parameters. They investigated the horizon beyond which forecasts conditioned on past observations were no longer relevant. The economics notion of content function was used for this purpose. In economics, the forecast content function and content horizon were used to set the horizon beyond which forecasts conditioned on past observations were no more relevant. These notions were adapted to the present model. The results of their study suggested the length of forecast period should not exceed ten years.

11. Marie Claire Koissi, Arnold Shapiro , GÄoran HÄognÄ and Ronald Lee. (2008). Fitting and Forecasting Mortality Rates for Nordic Countries Using the Lee-Carter method[17] : Presented at the 43rd Actuarial Research Conference, Regina, Canada: This paper aims to comparison between three different methods of estimating the model's parameters: the Singular Value Decomposition, the Weighted Least Square method and the

Maximum Likelihood Estimate. The LC model was applied to data from four Nordic countries: Denmark, Finland, Norway and Sweden. These approaches gave satisfactory results. The appropriate fitting period needs, however, to be well chosen. The properties of the model's parameters were studied using a bootstrap simulation. Compared the performance of the different estimation methods. The finding showed there was no variation was observed for the parameter $a_x$with the three approaches. For parameter b_x , the values obtained through WLS and the MLE were quite identical. The mortality index k_t had a common almost linear decreasing trend in the four countries with the three methods. The WLS and the MLE also gave quite identical values. The small error magnitude showed that the three approaches however gave good results. The results showed that, under an appropriately chosen estimation period, the estimated for the age parameters *a* and *b* were almost alike, while there was some variation in the estimates of the time-dependent mortality index *k*. A bootstrap simulation indicated that the used of the MLE results in smaller mean squared errors for the parameters *a* and *b* than the used of the two other methods.

12. Mariachiara Di Cesare and Mike Murphy, (2009), Forecasting Mortality, Different Approaches For Different Cause Of Deaths, The Cases Of Lung Cancer; Influenza, Pneumonia, Bronchitis; And Motor Vehicle Accidents, British Actuarial Journal British Actuarial Journal. 15:185-211[18] : The main goal of their paper to apply different models from different families of forecasting techniques. The models were Lee-Carter model, Booth-Maindonald-Simth, model, Age-period-cohort model and Bayesian models forecasting techniques to different causes of death with different underlying age and time patterns to assess which method better with the specificities of each case. This study analyzed

trends and forecasts mortality rates for three major causes of death lung cancer, influenza-pneumonia-bronchitis, and motor vehicle accidents , to assess how far different causes of death need different forecasting methods. Using data from the Twentieth and Twenty-First Century Mortality databases for England and Wales, the indicators was to use the goodness of fit and forecasting performance to assess the best model for each selected cause of death. The results showed major differences among the different forecasting techniques. In particular, when linearity was the main driver of past trends, Lee-Carter-based approaches were preferred due to their straightforward assumptions and limited need for subjective judgment. When a clear cohort pattern was detectable, such as with lung cancer, the Age-Period-Cohort model showed the best outcome. When completed and reliable historical trends were available the Bayesian model did not produce better results than the other models. The results showed major differences among the three forecasting techniques Lee-Carter and its Booth-Maindonald-Smith variant, Age-Period-Cohort model and Bayesian approach.

13. Jackie Li, (2010) Projections of New Zealand Mortality Using the Lee-Carter Model and its Augmented Common Factor Extension, Population Association of New Zealand 36:27-53[19] : This paper presented the results from an empirical study on projecting New Zealand mortality. He investigated the optimal starting year for fitting the model, and carried out residual analyses to assess model performance. He applied the Lee-Carter model and its augmented common factor extension to the mortality data and projected the death rates and life expectancy based on data by gender and single age (ages 0 to 110+.) for years 1948 to 2009. The fitted models appear to provide further insight into the underlying mortality

trends the original Lee-Carter and augmented common factor model perform similarly on the whole analysis.

14. Angela U. Chukwu and E. O. Oladipupo, (November 2012) Modeling Adult Mortality in Nigeria. Studies in Mathematical Sciences. 5:1-12[20]: An Analysis Based on the Lee-Carter Model: This study used the Lee-Carter method to model adult mortality in Nigeria. The model was applied to the age-specific mortality rates for Nigeria (for both gender) aged 15-84 years for the time periods 1990, 2000 and 2009, and forecasted from 2010-2019 was made. The model's parameters were estimated using the singular value decomposition technique, while the mortality index was predicted using the approach developed by Nan Li et al. (2002) for period 2010-2019.The results showed the model followed the mortality pattern very well for most of the ages.

15. Wasana Aberathna ، Lakshman Alles ، W. N.Wickremasinghe and Isuru Hewapathirana, (2014), Modeling and Forecasting Mortality in Sri Lanka. Sri Lankan Journal of Applied Statistics. (15-3)141-170[21] : This study was focused on modeling and forecasting mortality rates using Sri Lankan data and generating sex-specific life tables and to project future sex-specific and age-specific mortality for males and females, using the Lee-Carter approach. the mortality index forecast using several alternative univariate time series models, and the vector autoregressive (VAR) model performed better than the univariate models. From the estimated VAR model, mortality forecasts were generated for the period up to 2030 and life tables were generated for the selected periods of 2006-2008.The results showed the life expectancy at birth for males was 70.3 years, and 76.8 for females.

16. Farid Flici, (April 2015), Mortality forecasting for the Algerian population with considering cohort effect[22] : The aim of this paper to choose the best model to use for mortality forecasting, he applied the Lee-Carter method , RH model ,Age-Period-Cohort model and simpler APC for data from 1977 - 2011 for males and females age group 0-1, 1-5, and after by 5-age groups until 80, to forecast the period 2011-2013 using different time series models ARIMA(0,1,0), ARIMA(1,0,0) and ARIMA (2,0,0) . He estimated the parameters by the classical way then applied Weighed Least Squared then re-estimate $a_x$by including it solving the optimization problem . He used the parameters estimated in LC model as a starting values to estimate the APC model, then he used the parameters of APC model as a starting values for simpler APC. He compared between models, and finding the models lead approximately to the same results with some differences in the age specific mortality schemes.

17.Wan Zakiyatussariroh Wan Husin, Mohammad Said Zainol and Norazan Mohamed Ramli, (2015) Performance of the Lee-Carter State Space Model in Forecasting Mortality[23] : In their paper they used original Lee-Carter model and Lee-Carter incorporated State Space (LC-SS) Formulation on data from Peninsular Malaysia for period 1980-2009 to forecast mortality rate and the comparison between the two models based on Mean Square Error(MSR) and Mean Absolute Percentage Error (MAPE). The results indicate that LC-SS model performance better than the LC model.

18. Wouter van Wel (2015),Mortality Modeling and Forecasting using Cross-Validation Techniques. marble. (1.92)[24] : In this paper, the Heligman-Pollard model and the Lee-Carter model had been applied to modeling and forecasting. Cross-validation techniques were used to measure how accurately these two models performed in practice. The

main analysis based on data from the Netherlands (total population) where the data set was divided into the ”training set” (1850-1979) and the ”testing set” (1980-2009). The results, based on the MAPE, showed that the Heligman-Pollard model seemed to fit better to the Dutch data than the Lee-Carter model.

19. Lucia Andreozzi, Maria Teresa Blacona ،Nora Arnesi, The Lee Carter Method For Estimating And Forecasting Mortality[25] : An Application For Argentina. National University of Rosario, Argentina: They applied Lee-Carter model to age-specific death rates by gender in Argentina from 1979 to 2006 to forecast period from 2007-2011. The general index of mortality was forecasted using ARIMA(0,1,2) and Space State Model SSM models. Forecasts models such as ARIMA(0,1,2) with constant and SSM that were used to project the k index present an adequate fit. The results showed the SSM models present wider intervals than the ARIMA models, and the estimations of death rates and life expectancy were similar for both forecast models. The Lee Carter method in combination with ARIMA and Space-State models successfully predicted future death rates. However, long term forecast were necessary.

20. Rosella Giacomettia, Marida Bertocchib, Svetlozar T. Rachevc, Frank J. Fabozzid, A comparison of the Lee-Carter model and AR-ARCH model for forecasting mortality rates[26] : In their paper they compared performance of two models AR(1)-ARCH(1) model with Lee-Carter model. They fitted the models, with Gaussian and *t*-student innovations, for Italian death rates from 1960 to 2003 taken from “Human Mortality Database” . They compared the forecast ability of the two models for the period 2004-2006 and find that the AR(1)-ARCH(1) model with *t*-student innovations provides were best fitted than the Lee-Carter models

1.10 Research Organization:

The study is organized as follows: Chapter one (Introduction) contains a preface , research problems, research importance, research objectives, research hypothesis, research methodology, data sources and previous studies. Chapter two (Lee-Carter Model) contains Preface, Measures of Mortality Rate, Mortality Models Techniques, Criteria for Term Structure of Mortality Models, Mortality Forecasting Methods In The Past, Lee-Carter Model and Times Series. Chapter three (Cancer) contains Preface and The Genetic Bases of cancer .Chapter four (Application) contains Preface and Results and Interpretations. Chapter five (Conclusions and Recommendations) contains conclusions and recommendations .

CHAPTER TWO

(Lee-Carter Model)

2.1 Preface
2.2 Measures of Mortality Rate
2.3 Mortality Models Techniques
2.4 Criteria for Term Structure of Mortality Models
2.5 Mortality Forecasting Methods In The Past
2.6 Lee-Carter Model
2.7 Times Series

2.1 Preface [27]:

A rate is the number of events divided by the amount of exposure time that yielded the events (the speed with which the events took place) .

Mortality rate or death rate is the number of people who die in a year and area, divided by the population in the region or period of time.

2.2 Measures of Mortality Rate :

There are several different *mortality rates* used to monitor the level of mortality in populations , the following are most commonly used [28]:

2.2.1 Crude mortality rate: is the all deaths divided by population. It used to compare mortality rate among countries and regions.

2.2.2 Age specific mortality rate: Death occurs at all ages and the risk of mortality varies with age. It would therefore be necessary to analysis death rates for populations at different ages or age groups .

2.2.3 Cancer mortality rate: Is the number of death with cancer as the underlying cause of death occurring in specified population during a year per 100,000. It given by:

$$\text{Cancer mortality rate} = \frac{\textbf{Cancer Deaths}}{\textbf{Population}} \textbf{X100,000} \quad \text{.....(2.1)}$$

2.2.4 Cause-Specific Death Rates: Is the number of death with cause per a year per 1000 people of given age.

2.2.5 Infant mortality rate: Is the number of deaths among children under one year of age divided by the number of live births.

2.2.6 Maternal mortality rate: Is the number of mothers who die in incident related to child bearing. It divided by the number of live births.

2.3 Mortality Models Techniques:

There are many types of techniques used when the model the mortality rate and these are some of them[29] **:**

2.3.1 Extrapolative : Are based on projecting historical trends in mortality into the future. Simple extrapolative methods depend on the change in mortality rates in the past will continue to have a similar impact in the future.

2.3.2 Explanatory: Explanatory-based models use regression to predict mortality based on economic or environmental factors This type of model requires a determination of explanatory variables, and is not commonly used .

2.4 Criteria for Term Structure of Mortality Models:

To model mortality as a stochastic process, it is a reasonable requirement is that any mortality model would meet the following criteria [30]:

2.4.1 The model should keep the force of mortality positive.

2.4.2 The model should be consistent with historical data.

2.5 Mortality Forecasting Methods In The Past :

Many of these method are very simple and they are not used technical methods for mortality forecasting, and there are [31]:

2.5.1 Graphical Period Forecast:

This method for forecasting is simplest and it used for every age and does not flow any technical method for forecast and the error is a large and it is not objective because it depend on the person. The technique is plotted the various values of μ (x, t) where is mortality rate and t for a constant value of x (x1 say), drawing a smooth curve through the points and extending the curve to give values of μ (x1, t) for future values of t.

2.5.2 Graphical Generation Forecast:

As above, this method is simple and used for every year and it does not follow any technical method for forecast and the error is a large and it not objective because it depend on the person. The technique is plotted the mortality rate $\mu_{\theta,x}$ and θ for various values f x, where θ = t - x is the year of birth and they are joined by a curve and are extrapolated at the same direction. This method affected by temporary phenomena such as an epidemic.

2.5.3 Rhodes's Method:

In (1943) Kermack, Mckendrik and Mckinlay Makeham Period showed $\mu_{\theta,x}$ depend on two factors the age x and the year of birth θ. It is given as:

$$\mu_{\theta,x} = Q(x)R(\theta)$$

.....(2.2)

where

$Q(x)$ is a function of age .

$R(\theta)$ is a function of the birth.

This method is a simple formula and it faced problems for ages over 30.

2.5.4 Makeham Period Method:

This method was first used by R. Blaschke in 1923.This method required extensive data and is limited to ages over 30. In this method, for each calendar year t1, for which μ (x, t) is available, we graduate μ (x,t1) by the Makeham curve and consider the constants so obtained as functions of time, It's given by :

$$\mu\,(x,t) = A + BC^{x-30}$$

.....(2.3)

where

A, B and C are functions of x .

Nowadays, many methods have been proposed to model and forecast mortality rates, there is an extensive list of mortality forecasting models,

In the following section, we describe the most popular model, which we used in this study .

2.6 Lee-Carter Model [3]:

Lee-Carter (LC) model introduced by Lee and carter (1992) with article "Modeling and Forecasting the U.S. mortality in journal of American Statistical Association". The method describes the log of a time series of age-specific death rates as the sum of an age-specific component that is independent of time and another component that is the product of a time-varying parameter reflecting the general level of mortality, and an age-specific component that represents how rapidly or slowly mortality at each age varies when the general level of mortality changes. Lee-Carter model is one of the most popular methods for modeling mortality rates for all ages, because it is easily applied and provides fairly accurate mortality estimations and population projections. It became reference and leading statistical model for forecasting mortality. The model combines a demographic model with statistical model time series to forecast mortality rate [31].

2.6.1 Area Applied the Model:

The model applied in many countries as U.S. data from 1933 to 1987 (Lee and Carter, 1992), Canada data from 1922 to 1995 (Lee and Nault, 1993), Chile data from 1952 to 1987 (Lee and Rofman, 1994), China (Lin, 1995), Japan (Wilmoth, 1996), Finland (Alho, 1998), Brazil (Fígoli,1998),the seven most economically developed nations (G7) (Tuljapurkar et al., 2000), Belgium (Brouhns et al., 2002) (Brouhns and Denuit, 2001), Austria (Carter and Prskawetz, 2001), Portuguese mortality 1942 –1999 (Coelho, 2001),Australia (Booth et al., 2002, De Jong and Tickle, 2006) , Norway (Keilman et al., 2002), U.K. (Renshaw and Haberman, 2003b), Sweden (Lundström and Qvist, 2004,

Tuljapurkar, 2005), Italy 1950 – 2000 (Haberman and Russolillo, 2005), Spain (Felipe et al., 2002, Debón et al., 2006), the China and South Korea with limited data (Li, Lee and Tuljapurkar, 2004), the Nordic countries (Koissi et al., 2006), U.S. male mortality data: mortality rate forecasts are formed for the period 1990 – 1999 based on data from 1959 – 1989 (Pedroza, 2006Sweden 1860 – 2004) (Wang, 2007), Canada and the United States (Li and Chan, 2007,Taiwan (Wang and Liu, 2010), the Romanian female population, during 1970 – 2002 (Lazar),) The Continuous Mortality Investigation Bureau (CMIB, 2006) in Britain, U.S. Social Security Technical Advisory Panels and US Census Bureau (2000). The model used to all causes and cause specific mortality rate [32].

2.6.2 Advantages of The Model:

The strength of the model are simplicity and negative mortality rate cannot occur in forecasting. Lee-Carter model reduces the role of subjective judgment. There are no more decisions must be made about how far the historical data and what the model must be used, but the new studies showed that the period of historical data and the starting and long have effect on performance of forecasting,[12,17,33]. The model has a few parameters and easy to estimate and interpretable . It represent a large proportion of variability in mortality rate and produce stochastic forecast with probabilistic prediction intervals [29]. Finally only K_t need to predict the mortality rate.

2.6.3 Disadvantages of The Model:

The disadvantage of the LC model is the constant assumption for the parameters and the limiting mortality of 0. These issue has prompted lots of discussions and many proposed modifications. And these assumptions

of the parameters to be constant over time, whereas empirical studies in various countries do not support this assumption time [34,35] .

Nevertheless, there are still some limitations to the standard LC model. The model does not work well to forecast mortality for a group of populations [36], and it cannot deal with limited data, also it cannot include external factors which have impacts on mortality and it has narrow prediction intervals [13].

2.6.4 Development of The Model:

The model has undergone and different extensions and modifications added to improve the performance of the model by adding additional statistical features as non parametric smoothing, kalman filtering, Possion-gamma setting by Delwarde et al. (2007), and Li et al. (2009). and multiple principle component ,The extensions by Lee and Tuljapurkar 1994; wilmoth (1993); Carter (1995), Lee and Miller (2001). Second by Booth et al.(2002) and (2005), and Dejong and Tickle (2006). The other two extension by Hyndman and Ullah (2007) . Lee-carter has been applied to cause of death data (Wilmoth, 1998) to sex separately and by age (Carter 1996a; Carter and Lee 1992). A state space model is used (Carter, 1996b). Lee (2000) summarized the model's development, extensions and applications as stochastic forecast of Social Security system finances.

In the following section, we describe the extension and development of the model [37].

2.6.4.1 The Lee-Miller Variant (LM) [34]:

In 2001 Lee and Miller noted that for US data the forecast was biased when using the fitting period 1900–1989 to forecast the period 1990–1997. The source of error was the mismatch between fitted rates for 1998

and actual rates in this year in life expectancy for males and females . Jump-off bias was avoided by constraining the model such that k_t passes through zero in the jump-off year. also noted that the pattern of change in mortality was not fixed over time, as the LC model assumes, for 1900–1950 and 1950–1995. They adopted 1950 as the first year of the fitting period. The adjustment of k_t by fitting to e(0) was adopted to avoid the use of population data as required for fitting to D_t.

2.6.4.2 The Booth-Maindonald-Smith Variant (BMS) [38]:

In 2002 Booth, Maindonald and Smith modified the LC model by choosing optimally time period over which to fit the model They are choose the fitting period based on the statistical goodness-fit criteria. The procedure for the adjustment of k_t was modified, instead fitted to total deaths D_t they fitted to the age distribution of deaths $D_{x,t}$, using the Poisson distribution to model the death process. The jump-off rates are taken to be the fitted rates.

2.6.4.3 Lee-Carter Age Period Cohort (APC) [36]:

In 2003 Renshaw and Haberman proposed an extension to the LC model intended to capture age, period and cohort effects depends on the specific age of birth t-x. The change from the original LC model was the addition of a variable to capture the change in mortality between successive cohorts . A cohort effect, which the year of birth into the model. They extended the Lee-Carter model to include the second SVD term to allow for age-specific enhancement and compared its forecast with similarly-enhanced GLM and Poisson log-bilinear forecast. The k_t and γ_{t-x} parameters are forecasted using univariate time series models and also a multivariate time series could be used.

$$lnm_{x,t} = a_x + b_x^{(1)} k_t + b_x^{(2)} \gamma_{t-x} + \epsilon_{x,t} \text{.....(2.4)}$$

where

$lnm_{x,t}$ is the mortality rate at age *x* in year *t*.

a_x is the average of the mortality rate over time,

$b_x^{(1)}$ and $b_x^{(2)}$ measure the response at age *x* to changes in k_t and γ_{t-x} respectively.

k_t represents the overall level of mortality in year *t*.

γ_{t-x} represents the overall level of mortality for the cohort born in year t-x .

$\epsilon_{x,t}$ is the residual.

2.6.4.4 Augmented Common Factor Lee-Carter Model (ACFLC):

In 2005 Li and Lee to avoid the problem of fitting one population they suggested to extended the Lee-Carter model into Augmented Common Factor Lee-Carter Model (ACFLC). And identified by:

$$lnm_{x,t,i} = a_{x,i} + b_{x,i} k_{t,i} + B_x K_t + \varepsilon_{x,t,i} \text{.....(2.5)}$$

where

$lnm_{x,t,i}$ represent mortality rate in age x, time t and sex i.

$a_{x,i}$ represent the general shape of mortality rate in age x and sex i.

$b_{x,i} k_{t,i}$ is specific for sex i and allows for a short-term or medium-term difference between the rate of change in sex i mortality rates and that rate of change implied by the common factor .

$B_x K_t$ is a common factor, represent a main trend in mortality change of the whole population.

$\varepsilon_{x,t,i}$ are homoskedastic normally distributed random with mean zero and variance σ_ε^2.

2.6.4.5 De Jong and Tickle Model (DJ) [12]:

In 2006 De Jong and Tickle reduce the number of parameters in LC model to model mortality rates as a smoothed state space model**.** They have been used MLE to estimate b and k_t are derived by kalman filtering and smoothing and random walk with drift (0,1,0) for forecasted. The fitting period is restricted to 1950 on to avoid outliers. They termed model LC(smooth) and it is given as **:**

$$y_t = Xa + Xbk_t + \in_t \quad \text{.....(2.6)}$$

where

y_t is log mortality rate at each age in year t.

X is a known design matrix where the rows is more than columns.

a and b are parameter of age.

k_t is a mortality index.

$\in_t$ is residual has mean zero and variance $\sigma_{\in}^2$.

2.6.4.6 The Hyndman-Ullah Functional Data Method (HU) [12]:

In 2007 Hyndman and Ullah extents LC model by assumed mortality rate is a function of age with error and estimating death rate by using nonparametric smoothing methods and more than one set of (b_x, k_t) components is used. They used state space models for exponential smoothing are used to forecast mortality rather than random walk with drift and used robust estimation for unusual years due to wars or epidemics, and it does not adjust kt. It given by:

$$lnm_{x,t} = a(x) + \sum_j k_{t,j} b_j(x) + e_t(x) + \sigma_t(x)\varepsilon_{x,t} \quad \text{.....(2.7)}$$

where

$a(x)$ is the average pattern of mortality by age across years.

b_j (x) is a "basis function" and kt,j is a time series coefficient.

$e_t(x)$ is modeling error.

$\sigma_t(x)\varepsilon_{x,t}$ is accounts for observational error across age x.

2.6.5 The LC Model:

The $m_{x,t}$ denotes the central death rate experienced with aged x and year t. The following is a definition of central death rate :

$$m_{x,t}=\frac{d_{x,t}}{e_{x,t}}$$

.....(2.8)

where $d_{x,t}$ and $e_{x,t}$ for the number of death and the number of people in aged x year t.

$$lnm_{x,t} = a_x + b_x k_t + \epsilon_{x,t}$$

.....(2.9)

where the a_x coefficient describe the overall level of mortality corresponding with a_x age-specific pattern of mortality, The b_x coefficients reflect the age specific sensitivity to changes in the mortality index and it be is invariant over time for all age. The k_t coefficient represent the time trend reflecting general level of mortality in time and the model includes no assumption about the nature of the trend in k_t, the product of $b_x k_t$ reflect the age specific development of the mortality level in time. The $\epsilon_{x,t}$is an error term at age x and time t assumed to follow independent $N(0,\sigma^2)$.

2.6.5.1 Estimating The Model's Parameters:

In their original paper to estimate a_x, b_x and k_t they applied two stages estimation procedure, the first is singular value decomposition (SVD), which applied to the log of mortality

$$\ln m_{x,t} - a_x = b_x k_t$$

.....(2.10)

They add these constraint $\sum b_x = 1$ and $\sum k_t = 0$ to find unique solution of b_x and k_t.

a_x is computed as average of mortality rate over time t.

Let: $Q(a, b, k) = \sum_{x,t}(\ln m_{x,t} - a_x - b_x k_t)^2$.....

(2.11)

Let $\frac{\partial Q}{\partial a_x} = \frac{\partial Q}{\partial b_x} = \frac{\partial Q}{\partial k_{t_t}} =$

0 .....(2.12)

$$\frac{\partial Q}{\partial a_x} = 2\sum_t (\ln m_{x,t} - a_x - b_x k_t) \quad = 0$$

$$\frac{\partial Q}{\partial b_x} = 2\sum_t \left(\ln m_{x,t} - a_x - b_x k_t\right) k_t = 0$$

$$\frac{\partial Q}{\partial k_t} = 2\sum_x \left(\ln m_{x,t} - a_x - b_x k_t\right) b_x = 0$$

Then we can get directly,

$$\sum_t a_x = \sum_t \ln m_{x,t} - \sum_t b_x k_t = \sum_t \ln m_{x,t}$$

$$\hat{a}_x =$$

$$\frac{1}{T}\sum_t \ln m_{x,t}$$.....(2.13)

To estimate b_x, k_t by SVD. Wilmosth (1993) has improved method based on SVD is called weighted SVD that collapse two stage in one, and Maximum Likelihood Estimate (MLE) to estimate k_t in one stage.

2.6.5.1.1 Singular Value Decomposition (SVD):

Principle Component Analysis made its first appearance in demography with Ledermann and Breas (1959), who used factor analysis to analyze life table data from different countries. Then it presented by Bozik and Bell (1987) for projecting age-specific fertility rate ,and it extend by Bell and Monsell (1991) to forecast age-specific mortality rate[23].SVD is a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. It is based on a theorem from linear algebra which says that a matrix A can be broken down into the product of three matrices - an orthogonal matrix U, a diagonal matrix D, and the transpose of an orthogonal matrix V[40] .

$A_{m,n}$ can be decomposed uniquely as

$$Z = UDV^T \quad \text{.....} \quad (2.14)$$

U is mxn and orthogonal (its columns are eigenvectors of ZZ^T) and $U^TU = I$.

V is nxn and orthogonal (Its columns are eigenvectors of Z^TZ) and $VV^T = I$.

D is diagonal (is real values called singular values).

D=diag$(p_1, p_2, \dots, p_n)$ ordered so that $p_1 \geq p_2 \geq \cdots \geq p_n$

Transform the forecasting an age-specific vector $\ln m_{x,t}$ into forecasting as scalar k_t with small error.

SVD Method:

Obtain the logarithm of $m_{x,t}$ of the mortality rate.

obtain the $\hat{a}_x = \frac{1}{n}\sum_t \ln m_{x,t}$. itis a column vector of average of mortality rate.

Create matrix $Z_{x,t}$ for estimating b_x and k_t

where $Z_{x,t} = \ln m_{x,t} - \hat{a}_x$

.....(2.15)

Apply singular value decomposition to $Z_{x,t}$ to decompose the matrix $Z_{x,t}$ into the product three matrix

$$SVD(Z_{x,t}) = ULV$$

.....(2.16)

U : represent the age component.

L : represent the singular values.

V : represent the time component.

$\hat{b}_x$: is derived from the first vector of U. And $\hat{k}_t$ is derived from the first vector of V.

$$\hat{b}_x = \frac{1}{\sum_x u_{x,i}^2} . (u_{1,1} \quad u_{2,1} \quad \dots \quad u_{x,1})$$

.....(2.17)

$$\hat{k}_t = \sum_t u_{x,1}^2 . l_1 . (v_{1,1} \quad v_{2,1} \quad \dots \quad v_{t,1})$$

.....(3.18)

They made second stage to estimate k_t to find value that makes the observed number of death equal to the predicted number of death, because they noticed the observed number total of death is not equal to the fitted number of death and this is called jump-of bias , so uses $\hat{a}_x$, $\hat{b}_x$ values from the first stage to obtain new estimation of $k_t$which comply with the following :

$$D_t = \sum_x N_{x,t} \; e^{\hat{a}_x + \hat{b}_x \hat{k}_t}$$

.....(3.19)

where D_t is the total of deaths in year t. $N_{x,t}$ the population of age x in year t (exposure to risk), and this difference has occurred because estimate k_t by minimizing least square error over log mortality not mortality itself said wilmoth. He proposed two methods to avoid this issue and avoid zero cells death in cause-death and these are [8]:

2.6.5.1.2 Weighted Least Squares (WLS):

The WLS technique is based on the recognition that one could weight the first stage of Lee-Carter in such a way that observed and predicted deaths are closer to each other. To be specific, Wilmoth suggests finding the parameters a_x, b_x and k_t of the LC model of (3.1) as the solution of the weighted least squares (WLS) problem.

$$f_{x,t} = \ln m_{x,t} = a_x + b_x k_t$$

$$\sum_{x,t} d_{x,t}(f_{x,t} - a_x - b_x k_t)^2$$

.....(2.20)

where

$d_{x,t}$ is the observed number of deaths in age group x at time t.

The equation(2.20) above gives more weight to those age groups and years with large numbers of deaths, and the resulting estimates are more likely to fit the total number of deaths in each year. To solve the equation(2.20) we must compute its first derivation with respect to a_x, b_x and k_t and to set these equal to zero, then solving for required parameter which:

WLS used the same constraint $\sum b_x = 1$ and $\sum k_t = 0$ to find unique solution of b_x and k_t.

Let:$Q(a,b,k) = \sum_{x,t} d_{x,t}(f_{x,t} - a_x - b_x k_t)^2$.....(2.21)

Let $\frac{\partial Q}{\partial a_x} = \frac{\partial Q}{\partial b_x} = \frac{\partial Q}{\partial k_t} 0$.....(2.22)

$$\frac{\partial Q}{\partial a_x} = 2\sum_{t} d_{x,t}(f_{x,t} - a_x - b_x k_t) \ = 0$$

$$\frac{\partial Q}{\partial b_x} = 2\sum_{t} d_{x,t}\left(f_{x,t} - a_x - b_x k_t\right)k_t = 0$$

$$\frac{\partial Q}{\partial k_t} = 2\sum_{x} d_{x,t}\left(f_{x,t} - a_x - b_x k_t\right)b_x = 0$$

Then we can get directly,

$$\sum_{t} d_{x,t} a_x = \sum_{t} d_{x,t}(f_{x,t} - \hat{b}_x k_t)$$

$$\hat{a}_x = \frac{\sum_t d_{x,t}(f_{x,t} - \hat{b}_x \hat{k}_t)}{\sum_t d_{x,t}} \text{.....(2.23)}$$

$$\hat{b}_x = \frac{\sum_t d_{x,t}\hat{k}_t(f_{x,t} - a_x)}{\sum_t d_{x,t}\widehat{k^2}_t} \text{.....(2.24)}$$

$$\hat{k}_t = \frac{\sum_x d_{x,t}\hat{b}_x(f_{x,t} - a_x)}{\sum_x d_{x,t}\widehat{b^2}_x} \text{.....(2.25)}$$

The advantage of WLS are the first is that it eliminates the problem that log-mortality is not defined when the number of deaths is zero, and the predicted values are closest to observed deaths rates for those ages and years when the raw number of deaths was highest .A third appealing feature of Equation (2.20) is that it is easy to write down the corresponding first order conditions . but the procedure is not statistically

sound and the estimates resulting from this minimization problem have no known statistical properties**.**

2.6. 5.1.3 Maximum Likelihood Estimation (MLE):

MLE is a statistical technique for estimating model parameters, proceeds to maximize a likelihood function, which in turn maximizes the agreement between the model and the data. The MLE referred as Poisson log bilinear model. It gives optimal solution of the LC model under a Poisson model .Let $D_{x,t}$ denote a random variable represented the death count at age x and time t, let $d_{x,t}$ be corresponding number of deaths observed at age x and time t. $D_{x,t}$ can be satisfactorily approximated by a Poisson distribution with mean $\Lambda_{x,t}$ where $\Lambda_{x,t} = m_{x,t}E_{x,t}$. is population at age x and time t. Then can be written as

$$L(d, \Lambda) = \frac{\Lambda^{d} e^{-\Lambda}}{d!} \quad \text{.....(2.26)}$$

$$= d ln\Lambda - \Lambda - \ln d!$$

Sum over all cells to obtain the full log likelihood

$$L = \sum_{x,t}(d_{x,t} ln\Lambda_{x,t} - \Lambda_{x,t} - \ln d_{x,t}!)$$

.....(2.27)

The third term does not depend on $\Lambda_{x,t}$ so

$$L = \sum_{x,t}(d_{x,t} ln\Lambda_{x,t} - \Lambda_{x,t}) \quad \text{.....(2.28)}$$

If there are no restriction on the form of then the equation(2.28) has a maximum value when $\Lambda_{x,t} = d_{x,t}$ so

$$\Lambda_{x,t} = m_{x,t}E_{x,t} = e^{a_x + b_x k_t} E_{x,t} \quad \text{.....(2.29)}$$

The ML estimates of the parameters of the LC model are found by substituting in equation(2.28), and maximizing the equation with a_x, b_x and k_t.

$$L = \sum_{x,t}((d_{x,t}(a_x + b_x k_t + \ln E_{x,t}) - (e^{a_x + b_x k_t} E_{x,t}))...(2.30)$$

2.6.6 Forecasting Kt and Age-Specific Mortality Rate:

After k_t index is obtained it possible to forecast mortality index, many techniques have been used and specified . Lee and Carter observed that in most cases a random walk with drift (0,1,0) can be appropriate for modelling the mortality index. It is given as:

$$k_t = k_{t-1} + d + \epsilon_t$$

.....(2.31)

where

d : is drift parameter.

ϵ_t: is an error term with zero mean and constant variance.

$$\hat{d} = \frac{\hat{k}_T - \hat{k}_1}{T-1}$$

.....(2.32)

It depend on the first and last of k_t estimation.

The drift d estimated with uncertainty and standard error of it estimated and it used to form more complete measure of uncertainty in forecasting $\hat{k}_t$.

Finally to obtain forecast of the mortality rates, the forecasted values $\hat{k}_t$ of are implemented along with estimate values of $\hat{a}_x$ and $\hat{b}_x$. The forecast of the mortality rate for year t+1 is:

$$\hat{m}_{x,t+1} \approx e^{\hat{a}_x + \hat{b}_x \hat{k}_t}$$

.....(2.33)

2.6.7 Dealing with Uncertainties:

The original LC model incorporates uncertainty arising from the LC model $\epsilon_{x,t}$ and time series model ϵ_t. But LC model's error is not important and they ignored it .

2.6.7.1 The Residual from LC Model:

The error from the model calculated directly and is given as

$$\hat{\epsilon}_{x,t} = \ln m_{x,t} - \hat{a}_x - \hat{b}_x \hat{k}_t \quad \text{.....(2.34)}$$

but this error is very small, so the researcher ignored it in their studies as the Lee and Carter in their original paper. In original LC model the variance of $\hat{a}_x$ and $\hat{b}_x$ has been ignored because it was very small and the variance of $\hat{k}_t$ is derived from time series.

2.6.7.2 The Residual From RWD Model:

The error from the model called the associated forecast error variance and included the error from estimate parameter and the error from the model and the called respectfully the standard error constant sec and the standard error equation see.

$$\hat{\sigma}_u^2 = \Delta t \times (see)^2 + (\Delta t \times sec)^2 \quad \text{.....(2.35)}$$

$$\hat{\sigma}_u = \sqrt{\Delta t \times (see)^2 + (\Delta t \times sec)^2} \quad \text{.....(2.36)}$$

where

$$see = \hat{\sigma}_{rw} = \sqrt{\frac{1}{T-1}\sum_t \left(\hat{k}_t - \hat{k}_{t-1} - \hat{d}\right)^2} \quad \text{.....(2.37)}$$

$$sec = \hat{\sigma}_d = \frac{see}{\sqrt{T-1}} \quad \text{.....(2.38)}$$

2.6.7.3 Various Forecast Performance Measures:

Here discuss about the commonly used performance measures and their important properties. Each of these measures has some unique properties, different from others. It is better to consider more than one performance criteria. This will help to obtain a reasonable knowledge about the amount, magnitude and direction of overall forecast error.

2.6.7.3.1 Mean Error:

$$\mathrm{ME} = \frac{1}{\mathrm{A}}\sum_{\mathrm{x,t}} \hat{\epsilon}_{\mathrm{x,t}} \quad \text{.....(2.39)}$$

The properties of ME are:

- It is a measure of the average deviation of forecasted values from actual ones.
- It shows the direction of error and thus also termed as the Forecast Bias.
- There is no way to know their exact amount of ME, because the effects of positive and negative errors cancel out .
- A zero ME does not mean that forecasts are perfect, it indicates that forecasts are on proper target.
- For a good forecast, to have a minimum bias, it is better that the MFE is as close to zero as possible.

2.6.7.3.2 Mean Square Error:

$$\mathrm{MSE} = \frac{1}{\mathrm{A}}\sum_{\mathrm{x,t}} \hat{\epsilon}_{\mathrm{x,t}}{}^{2}. \quad \text{.....(2.40)}$$

The properties are:

- It is a measure of average squared deviation of forecasted values.
- The opposite signed errors do not offset one another, MSE gives an overall idea of the error occurred during forecasting.

• MSE emphasizes the fact that the total forecast error is in fact much affected by large individual errors.

• MSE does not provide any idea about the direction of overall error.

• MSE is sensitive to the change of scale and data transformations.

2.6.7.3.3 Sum of Squared Error:

$$\text{SSE} = \sum_{x,t} \hat{\epsilon}_{x,t}^{\,2} \quad \text{.....(2.41)}$$

The properties of SSE are:

• It measures the total squared deviation of forecasted observations, from the actual values.

• The properties of SSE are same as those of MSE.

2.6.7.3.4 Mean absolute Error:

$$\text{MAE} = \frac{1}{A}\sum_{x,t}\left|\hat{\epsilon}_{x,t}\right| \quad \text{.....(2.42)}$$

The properties are:

• It measures the average absolute deviation of forecasted values from original ones.

• It is also termed as the Mean Absolute Deviation (MAD).

• It shows the magnitude of overall error, occurred due to forecasting.

• The effects of positive and negative errors do not cancel out.

• For a good forecast, the obtained MAE should be as small as possible.

2.6.7.3.5 Mean Absolute Percentage Error:

$$\text{MAPE} = \frac{1}{A}\sum_{x,t}\left|\frac{\hat{\epsilon}_{x,t}}{m_{x,t}}\right| \quad \text{.....(2.43)}$$

The properties of MAPE are:

• This measure represents the percentage of average absolute error occurred.

• It is independent of the scale of measurement, but affected by data transformation.

• It does not show the direction of error.

• In this measure, opposite signed errors do not offset each other.

2.6.7.3.6 Mean Percentage Error:

$$\mathrm{MPE} = \frac{1}{A}\sum_{x,t}\frac{\hat{\epsilon}_{x,t}}{m_{x,t}} \quad (2.44)$$

The properties of MPE are:

• MPE represents the percentage of average error occurred, while forecasting.

• It has similar properties as MAPE.

• It shows the direction of error occurred.

• Opposite signed errors affect each other and cancel out.

• Thus like MFE, by obtaining a value of MPE close to zero, we cannot conclude that the corresponding model performed very well.

• It is better that for a good forecast the obtained MPE should be small.

where

$\hat{\epsilon}_{x,t} = m_{x,t} - \hat{m}_{x,t}$.

x 1,2,...,A.

t 1,2,...,T.

2.7 Times Series [41]:

Mathematical modeling plays an important role in forecasting, and it's a simple and useful model . Time series is a set of observation x_t, each one being recorded at specific time t. To answer what will the rate of

mortality of cancer would be next year ? by use time series data to develop forecasting models .

There are two types of time series. A discrete time series ,the observations are taken in discrete set, the other is continues time series that are obtained when observations are taken over time interval $T_0 = [0,1]$.

The data's patterns are very important to understand the behavior of time series in the past.

2.7.1 The Common Type of Time Series:

These are common type of series :

2.7.1.1 Horizontal pattern: When the data located around the constant mean.

2.7.1.2 Trend pattern: When it shows gradual movement to high or low value over a long time.

2.7.1.3 Seasonal pattern: When the same pattern repeating over period of time.

2.7.1.4 Trend and Seasonal pattern: It combination of two patterns.

2.7.2 Auto Regressive Integrated Moving Average ARIMA Models:

The ARIMA is fitted to data to understand and predict future data of time series and response time series as a linear combination of its own past values and past errors.

ARIMA(p,d,q) where p order of regressive ,d order of differencing and q order of moving average, and popularized by Box and Jenkizes (1975).

The common ARIMA models:

- ARIMA(0,0,0)+c.
- ARIMA(0,1,0) random walk model.
- ARIMA(0,1,0)+c random walk with drift.

2.7.2.1 Random Walk:

Is defined as a process where the current value of a variable is composed of the past value plus an error term that defined as white noise and is given as

$$x_t = x_{t-1} + \epsilon_t \quad \text{.....(2.45)}$$

random walk is nonstationary

2.7.2.2 Random Walk with Drift (RWD):

It is an one of the simplest and important models, it presented as a current observation equal to previous observation with a random step up or down, which are independently and identically distributed (I.I.d.) , and is given as

$$x_t = x_0 + d + \epsilon_t \quad \text{.....(2.46)}$$

where

$x_0 = 0.$

d is a drift.

ϵ_t is IIdN(0,σ_ϵ^2).

It can be rewrite .

$$x_t = x_0 + dt + \sum_1^n \epsilon_t \quad \text{.....(2.47)}$$

$$E(x_t) = dt + \sum_1^n E(\epsilon_t) = td \quad \text{.....(2.48)}$$

$$V(x_t) = 0 + \sum_1^n V(\epsilon_t) = t\,\sigma_\epsilon^2 \quad \text{.....(2.49)}$$

so the random walk drift is non stationary because the mean and variance are depend on time.

The autocovariance for random walk drift is:

$$\gamma(s,t) = \text{cov}(x_s, s_t) = \text{cov}\left(\sum_{j=1}^{s} \epsilon_j \sum_{k=1}^{t} \epsilon_k\right) = \min(s,t)\sigma_\epsilon^2$$

.....(2.50)

2.7.3 Fitting the ARIMA Model:

The first step is identify the model : It includes to specify the model (AR,MA and ARMA) and the order of it by Plot Autocorrelation Function (acf) and Partial Autocorrelation Function (pacf) and use stationary tests to determine if differencing is necessary to eliminate nonstationary in time series or fitting many models and then choose the best of them by using a goodness-of-fit. The second steps is Estimate the coefficients : estimate the them and test of significant if some of coefficients are be unnecessary . Least Square Method to estimate coefficients of AR model, estimation coefficients of MA and ARMA models are too complicated and it accomplished by computer programs. test of residuals indicate weather are contain addition information that may be suggest more complex models to use, the last step is to check the model : Must check to element that are :The residual of the model are random and parameters are statistically significant .The last stage is Forecasting stage: Forecast future values of time series and generate confidence intervals for this forecasts.

2.7.3.1 Test of Residuals :

The residual of the model must be random and acf must be zero at all lags except lag zero if there is no dependence between residuals only need to estimate the mean and variance and if there are dependence between them then we need to look for more complex models.

To examine if residuals are uncorrelated by scanned acf to see if coefficients fall outside of prediction intervals (PI)around zero. The autocorrelation coefficient r_k at lag k is normally distributed with :

$$E(r_k) = 0$$

.....(2.51)

$$Var(r_k) = \frac{1}{N}$$

.....(2.52)

$$PI = \pm \frac{1.96}{\sqrt{N}}$$

.....(2.53)

PI for 95%. When r_k outside the PI this is an evidence the residuals were not random.

2.7.3.2 Test of Coefficients :

The estimated coefficients should computed with their standard deviation to test the significantly different from zero. for estimated coefficient which had normally distributed with

$$Var(\hat{\alpha}_1) = \frac{1-\hat{\alpha}_1}{N}$$.....

(2.54)

The approximate 95% P.I for $\hat{\alpha}_1$ is $\hat{\alpha}_1 \pm 1.96\sqrt{Var(\hat{\alpha}_1)}$. If CI include zero reject hypothesis that coefficients were different from zero.

If we have correctly identified the model the coefficients should be significantly different from zero, and residuals acf and pacf look good then the model was good.

CHAPTER THREE
(Cancer)

3.1 Preface
3.2 The Genetic Bases of cancer
3.2.1 Oral Cancer
3.2.2 Lung Cancer
3.2.3 Colon Cancer

3.1 Preface [42]:

Cancer has afflicted humanity from prehistoric times. In the mammals the oldest evidence of cancer was found in fossilized dinosaurs and human bones from prehistoric times.

In nineteen century manuscript which written record about cancer in ancient Egyptian. Edwin Smith and George Ebers described surgery pharmacological and magical treatment between 1500- 1600 BC. Imhotep written the first reference to breast cancer.

In Rome and Greek the father of medicines is Hippocratis, he has written about diseases produced masses and recognized the progress of Krakinomas.

Aulus Cornelius Celsus is a Roman physician (25 BC -50 AD) evaluation of tumors from cacoethes later called carcinomas .

Archigenes of Apamea , Syria (75 - 129 AD) , he believed that the successful of remedies in the early stage of cancer is surgery for advanced cancer but only for strong patient. Galen classified tumors into types and origin and graded . He has written document about cancerours and non cancerours.

By the end of fourth century Oribasius described the cancer of face, breast and genitalia and cancer's painful. Paulus Eginate wrote four books about cancer.

Cancer knowledge of Greek spread into Arabs and the most famous Ibn Al Nafis who described pulmonary circulation blood in detail . Avenzoar has described symptoms of esophageal and stomach cancer in his book Kitab al taysir.

Cancer is the second leading cause of death globally after Ischaemic heart disease. It was responsible for 8.8 million deaths in 2015. In low- and middle-income countries approximately 70% of deaths came from cancer occur. In 2012, there were an estimated 8.2 million deaths from

cancer in the world 4.7 million (57%) in males and 3.5 million (43%) in females [43]. The most common causes of cancer death are cancers of: Lung (1.69 million deaths), Liver (788 000 deaths), Colorectal (774 000 deaths), Stomach (754 000 deaths) and Breast (571 000 deaths)[44]. Oral cancer is reported to be the eighth most commonly diagnosed cancer. WHO reported an Oral cancer mortality rate of approximately 2 per 100,000 in the Middle East[45] . Lung cancer is one of the top three cancers that caused the most economic impact globally and highest number of death rate among other diseases [46]. In the Arab world the studies show that (68.1%) of the Arab countries have lung cancer as one of the most common cancer [43]. WHO reported lung cancer deaths in Egypt reached 0.96% of total deaths. In the Arab countries have colon cancer as one of the most frequent five types of cancer , they are gradually increasing in the region.

3.2 The Genetic Bases of Cancer :

Genome determine the structure and function of organs ,it contains genes that packaged in 46 chromosomes .Genes are Deoxyribonucleic acid DNA contains the code for cells to produce protein that are the signals to control the structural and function of the cell . In cell cycle the genome duplicated and passed from cell to cell and from parent to offspring the error may occur.

The genetic material of cell contains 23 pair of chromosomes which are made up of DNA. DNA inside each cell which contains unique genetic blue print which has specific segment genes .

The chemical like tobacco ,air pollution, radiation from sun, viruses, and even chemical from bodies can damage genetic .The abnormality of genes lead to divide uncontrollably and not to die in the timeframe ,cancerous cell accumulate in the body forming tumors as result .

In this study we concern about oral, lung, and colon cancer which they developed, risk factors, growth and spread, symptoms, screen, staging and treatment.

3.2.1 Oral Cancer:

Oral cancer develop at any part of oral cavity including tongue ,gums, tonsils lining of the mouth, lips and upper part of the throat. There are two kind of oral cancer: oral cavity cancer and oropharyngeal cancer [47].

3.2.1.1 Risk Factor:

A risk factor is anything that changes a person's chance of getting a disease such as cancer [48] **:**

- Tobacco and Alcohol: People who used they have increase risk.
- Betel quid and Gutka: It is made up of areca nut and lime wrapped in a betel leaf.
- Genetic Factors: including:
    - epidermal growth factor receptor EGFR.
    - P53: It is suppressor gene.
- HPV Infection Human Papilloma Virus: People who have HPV increase risk.
- Infection Factor: It can be induced by bacteria, fungus and virus.
- Gender: It is common in men than women.
- Age: It is common in the older because cancer develop in many years.
- Ultraviolet light UV: People who have outdoor jobs they increase risk.
- Poor Nutrition : low of fruits and vegetables. Poor of vitamins A, C and E and iron trace elements such as selenium and zinc. and high animal products.
- Weakened Immune System.
- Graft Versus Host Diseases GVHD.

- Lichen Planus: It skin disease but sometimes it leads to cancer in lining of the mouth and throat.
- Air Pollution.

3.2.1.2 Growth and Spread :

Oral cancer can be spread to lymph nodes, it depend to tumor size, and it can spread to distant site in late stage of cancer.

3.2.1.3 Symptoms:

The most common symptoms of oral cancer are [48]**:**

- A sore and Pain in the mouth for long time.
- Lump in the cheek.
- White or red patch on gums, tongue, tonsil or lining of the mouth.
- Difficulty chewing or swallowing.
- Difficulty moving the jaw or tongue.
- Losing of the teeth or pain around the teeth.
- Voice changes.
- Weight loss.

3.2.1.4 Screen and Diagnosis:

A doctor make test to check the signs or symptoms of oral cancer , and include [48] :

- Medical history and physical exam.
- Complete head and neck exam: It used to detect lymph nodes of the neck.
- Indirect pharyngoscopy and larynoscopy : They used mirror to look inside the throat to detect any tumor here.
- Direct pharyngoscopy : It used endoscope to look into throat by inserted it in mouth or nose.

- Panendoscopy: It uses different types of endoscope. They insert into mouth or nose to detect oral cavity, oropharynx and other parts of mouth.
- Biopsy: Take sample of tissue and seen it under the microscope.
- Exfoliative cytology: This procedure is easy but it does not detect all cancer.
- Incisional biopsy: Cut a small tissue that is affected.
- Fine Needle Aspiration biopsy FNA: For this technique use thin hollow needle attached to syringe to take cells from the tumor lump. It is important because it use for:
    - Finding a new neck mass.
    - Learning the stage of cancer.
    - Seeing if cancer come back after treatment.
- Blood test.
- Dental exam: It requires when will use radiation therapy.
- Imaging test:
    - Computerized Tomographic CT Scan.
    - Magnetic Resonance Imaging MRI Scan.

3.2.1.5 Staging:

Staging is process to determine whether cancer has spread from original tumor [48], the size and characteristics determine the stage of cancer.

TNM it is an One of the most common methods used for cancer staging , which assigns a degree of severity based on the size, location, and spread of cancer in the body[49] .

Table(3.1) Determining Tumor Characteristics for oral cancer .

Tx:	Tumor can't be assessed.
T0:	No evidence of tumor
Tis:	Carcinoma in situ .
T1:	Tumor is 2 cm.

T2:	Tumor is larger than 2 cm but smaller than 4 cm.
T3:	Tumor is larger than 4 cm.
T4a:	Tumor is growing into near structures.
	Oral cancer tumor has grown in jawbones, face, tongue, face's skin or maxillary sinus.
	Oropharyngeal tumor has grown in larynx, tongue, hard palate and jaw.
T4b:	Tumor has grown nearby structures and deeper area or tissues..

T: It tells which tissues the primary tumor has grown.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.2) Regional lymph node status for oral cancer .

Nx:		Tumor nearby lymph nodes cannot be assessed.
No:		No evidence of cancer near lymph nodes.
N1:		Cancer has spread into lymph nodes and larger than 3cm.
N2:	**N2a:**	Cancer has spread into one lymph node and larger than 3cm and less than 6cm.
	N2b:	Cancer has spread into two lymph nodes.
	N2c:	Cancer has spread into lymph nodes and not larger than 6cm.
N3:		Cancer has spread into lymph nodes and larger than 6cm.

N: Tells where cancer has spread with near lymph node.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.3) Determining Metastatic Status for oral cancer .

M0:	No evidence of cancer has spread to distant sites.
M1:	Cancer has spread to distant sites.

M: Tells if cancer spread distant sites.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.4) The stage of oral cancer.

Stage	**T**	**N**	**M**
Stage0	Tis	N0	M0
Stage1	T1	N0	M0
Stage11	T2	N0	M0
Stage111	T3	N0	M0
	T1-T3	N1	M0

Stage1V	**Stage1VA**	T4a	N0 or N1	M0
		T1-T4a	N2	M0
	Stage1VB	T4b	Any N	M0
	Stage1VC	Any T	N3	M0
		Any T	Any N	M1

Source: UICC TNM classification of malignant tumours (UICC, 2002).

3.2.1.6 Treatment:

The treatment depend on the stage and location of the cancer. the main treatments are [48]:

- Surgery: It is an operation used to remove cancer cells and it has several type:
  - Tumor restriction: It is an operation used to remove tumor and normal cells that surrounding the affected area.
  - Mohsmicrographic surgery: It is an operation used to remove slice of tumor and test it under microscope and repeat this procedure until whole tumor removed.
  - Glossectomy: It is an operation used to treat tongue cancer by removing small part or whole part of tongue.
  - Mandibulectomy: It is an operation used to remove jaw bone that is affected.
  - Maxillectomy: It is an operation used to remove tumor where in hard palate.
  - Trans-oral-robotic surgery: It used to reseed cancer of the oropharynx and throat.
  - Laryngectomy: It is an operation used to remove larynx that has affected.
  - Neck dissection: It is an operation used to remove lymph nodes that have effected and it has several types :
    - Partial or selective neck dissection: Remove few lymph nodes.

- Modified neck dissection: Remove most of lymph nodes between jaw bone and collarbone.
- Radical neck dissection: Remove most of lymph nodes on one side ,muscles, nerves and veins.

- Radiation Therapy: In oral cancer it used:
  - As main treatment for small cancer.
  - After surgery sometimes.
  - Before surgery to shrink cancer cells.
  - To relieve symptoms of advanced cancer.

  It has two types:
  - External beam radiation therapy.
  - Brachy therapy.
- Chemotherapy: It used for:
  - Instead surgery as main treatment.
  - After surgery to kill small cancer cells and it called adjuvant chemotherapy.
  - Before surgery to shrink large cancer and it called induction chemotherapy.
  - Treat advanced cancer that cannot remove by surgery.
- Target therapy: It has less side effect .

Table(3.5) Describe the treatment for every stage of Oral cancer [50].

Stage	**Treatment**
Stage0	Surgery or radiation therapy
Stage1+ Stage11	Surgery, radiation therapy or chemotherapy +radiation therapy
	Lip cancer radiation and surgery
	Oral cancer surgery, radiation+ chemotherapy
	Orapharynx radiation, surgery, radiation therapy+ chemotherapy
Stage111	Oral cancer Surgery+ Radiation therapy
Stage1V	Orapharynx Radiation therapy+ Chemotherapy, chemotherapy
Stage1VB	Chemotherapy

Source: UICC TNM classification of malignant tumours (UICC, 2002).

3.2.2 Lung Cancer:

Lungs are important part of respiratory system. when we understand how lungs and respiratory system work then we understand how cancer's effected the body .

On each side of the chest there is one lung, the right lung is larger than the left .And divided into lobes (upper, middle and lower) , it has elastic fiber which it help it to expand and contract and covered by visceral pleura and parietal pleura . The main bronchi branch into lobar then divide into segmental then into bronchioles , the final branchi is atria and in alavoli which surrounding by capillaries. Lung contain also lymphatic vessels .

The ability of respiratory system to do its works depend on the health tissues of lungs.

Lung cancer sometimes referred as bronchio genic cancer or bronchio genic carcinoma [50]. Lung cancer is a malignant tumor in the tissue of one or

both of the lungs [51].

3.2.2.1 Risk Factor:

A risk factor is anything that changes a person's chance of getting a disease such as [52]:

- Smoking: People who are either current or former tobacco smoker.
- Second-hand smoke: People who breathe in air that contains tobacco smoke are exposed to its carcinogens.
- Environmental carcinogens : These are substances in the environment capable of producing genetic damage.
    - Asbestos: Is a fibrous mineral .
    - Radon: Is radioactive gas.

- Chromium: It has several forms not all of them increase risk of cancer but only chromium(VI).
- Nickel: It is hard, slivery and white metal.
- Polycyclic Aromatic Hydrocarbons (PAHs): Formed from several chemicals during incomplete burning .

- Genetic factor: Genetic controls how handle the exposure carcinogens. If some relative have cancer it increase a risk.
- Age: It contribute to increase risk because genetic damage tend to accumulate over time.

3.2.2.2 Growth and Spread:

Lung cancer is slow growing and it has ability to spread to other part of the body. Lung has many blood vessels and lymphatic vessels which are became a router for cells cancer to travel through them and spread to other part or to lymph nodes.

3.2.2.3 Type of Lung Cancer:

There are two main types of lung cancer according to World Health Organization (WHO) and International Association for study Lung Cancer (IASLC) which were updating classification of lung cancer in 1990 into Small Cell Lung Cancer (SCLC) and Non Small Cell Lung Cancer (NSCLC) which are different characteristics and treatment [55] .

3.2.2.3.1 Small Cell Lung Cancer (SCLC):

It start from large airway and it grows and spread quickly to lymph nodes and other organs, There is relationship between SCLC and tobacco.

3.2.2.3.2 Non Small Cell Lung Cancer (NSCLC):

It divide into three types adenocarcinoma, squamous cell carcinoma and large cell carcinoma all of them have same characteristics so they grouped together.

There are other cancer arise in lung as carcinoid tumor, malignant pleural mesothelioma.

3.2.2.4 Symptoms:

They are not specific but is common [53]: cough, change in preexisting cough, cough with blood, loss of weight, difficult and or painful breathing, chest pain and wheezing, difficulty swallowing, shoulder pain with or without arm and hand. numbness , weakness of extremities and facial swelling.

3.2.2.5 Screen and Diagnosis:

The most common methods to screening and diagnosis include [52]:

- Sputum cytology: A sample of sputum took on slide to examine malignant cells.
- Tumor Marker: Is substance released in the blood when cancer is found.
- Imaging Test : To determine if lung tumor is benign or malignant.
    - Chest X-ray: To test and study metastatic lung cancer. There are new technique digital chest x-ray and Computer Assisted Diagnosis(CAD).
    - Computerized Tomographic CT scan: It has ability to take picture in 3 dimension, It used to detect small tumor and determine the size and shape and where exact location , they invented new technique called spiral belical CT scan.
    - Magnetic Resonance Imaging MRI Scan: It produces image in 3 dimension by using magnet, and it used to detect specific area.
    - Pssitirom Emission Tomographic PET Scan: for this procedure it used amount of radiation to show brighter area (cells cancer).

- Tissue Diagnosis: It used to determine the type of cancer, it takes sample (biopsy) from tumor and examine it. It has different types
    - Bronchoscopy.
    - Mediastinoscopy.
    - Throracoscopy.
    - Transthoracic Needle Bibopsy.

3.2.2.6 Staging:

TNM classification system is used to determine the stage of lung cancer. Tables below show the process.

Table(3.6) Determining Tumor Characteristics for lung cancer :

T0:	No evidence of primary tumor.
Tis:	Carcinoma in situ
T1:	Tumor that is less than 3 cm in size and surrounding by lung tissue.
T2:	Tumor that is larger than 3 cm and surrounding by lung tissue and not invading chest wall.
T3:	Tumor of any size invades the chest wall, diaphragm, or the pleura of the mediastinum or heart.
T4:	A tumor of any size that invades the mediastinum or a vertebral body

T: It tells which tissues the primary tumor has grown.
Source: UICC TNM classification of malignant tumours (UICC, 2002)

Table(3.7) Regional lymph node status for lung cancer:

N0:	No evidence of cancer in lymph nodes.
N1:	Cancer in ipsilateral hilar lymph nodes.
N2:	Cancer in ipsilateral mediastinal lymph nodes.
N3:	Cancer in contralater lymph nodes or supraclavicular area.

N: Tells where cancer has spread with near lymph node.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.8) Determining Metastatic Status for lung cancer:

M0:	No distant metastasis found.
M1:	distant metastasis found.

M: Tells if cancer spread distant sites.

Source: UICC TNM classification of malignant tumours (UICC, 2002).

3.2.2.6.1 SCLC Stages:

limited stage affected one lung, the mediastinum, and original lymph nodes. It equality to stage 1 through 111B.

Extensive stage : It spread to contralateral lung associated with malignant pleura it equality to stage1v.

3.2.2.6.2 Non SCLC Stages:

It has four stages, and the table below describe them.

Table(3.9) The Stage of Non SCLC .

Stage		**T**	**N**	**M**
Stage1	**1A**	T1	N0	0
	1B	T2	N0	M0
Stage2	**11A**	T1	N	M0
	11B	T2	N1	M0
		T3	N0	M0
Stage3	**111A**	T3	N1	Mo
		T3	N2	M0
		T1	N2	M0
		T3	N2	M0
	111B	T4	N0	M0
		T4	N1	M0
		T4	N2	M0
		T1	N3	M0
		T2	N3	M0
		T4	N3	M0
Stage4		Any T	Any N	Any M

Source: UICC TNM classification of malignant tumours (UICC, 2002).

3.2.2.7 Treatment:

Lung cancer treatment depending on the type of cancer and its stage. by using surgery, radiotherapy, and chemotherapy, they used alone or in combination . All these treatments have side effects on the body [52].

- Surgery: Local treatment it used to remove tumor , and it has several procedures.

- Radiotherapy: It uses ionizing radiation to stop division of cells. There are two types:
  - Adjuvant Radiotherapy.
  - Palliative Radiotherapy.
- Chemotherapy: It used when cancer is spread out of original location.

3.2.3 Colon Cancer:

The digestive system start at the mouth and end at the anus ,and also consists of small intestine and large intestine .Digestive system breaks down food and turn it to an energy ,and also gets rid that body doesn't use it known as fasces or stool. The colon is a part of the large intestine. It is almost about 5 feet along, it removes water and nutrients from digested food and sent the remain material (stool) to rectum and it leaves the body through the anus.

Colon divided into four parts are ascending, transverse, descending, and sigmoid colon.

The colon's wall has four main layers mucosa it is inner layer, it made of epithelium that absorbs water from stool and makes mucus. Mucus helps stool to move through the colon. The second layer is submucosa that consists tissues, blood, lymph nodes, and nerve cells. The third layer is muscularis propria made of muscle fibers that helps stool to move through colon. The fourth layer is outer that consists of adventitia or serosa . Adventitia is connective tissue that binds the colon to other structures. Serosa is called visceral periloneum , it has a layer of connective tissue called subserosa. Subserosa covered by cells make lubricating fluid that allows colon move smoothly against other organs. Almost all colon cancer are adenocarcinomas, that start in the cells that line glands [54].

There are several types of colon polyps:

- Adenomatous or Adenomas polyps : They are most commonly polyps, most of them don't become cancer, but polyps that cause cancer start here.
- Hyperplastic polyps: Their cells grow fast and they found in the last part of colon, They were rare to become cancer.
- Inflammatory polyps: They occur after inflammatory bowel disease, and they rarely become cancer.
- Sessile polyps: They occur above colon wall.
- Serrated polyps: They have associated with cancer, but they rare.
- Pedunculated polyps.

3.2.3.1 Risk Factor:

The common risk factor include [54] :

- Hereditary Non Polyps Colon Cancer HNPCC: It is called lynch syndrome, it causes cancer.
- Polyps: People who have polyps in abdomen.
- Familial Adenomatous Polyps FAP: It often to leads cancer.
- Age: It is commonly in an older people.
- Diet: Diet that contains fat especially fat of animal ,low in calcium, foliate, fiber, fruits and vegetables.
- Life Style Factors: Overweight, smoking, and drinking alcohol.

3.2.3.2 Growth and Spread:

Colon cancer spread slower than others cancer. It can spread firstly to lymph nodes then to other distant organs. Overtime the benign growth and become malignant cells and they can reach blood and travel to other part of body like lymph nodes, lung, and liver.

3.2.3.3 Symptoms:

The most common symptoms of [54]**:**

- A change in bowel habit.
- A change in appearance of stool.
- Blood in the stool.
- weight loss.
- Weakness.
- Having nausea and or vomiting and feeling bloated.

3.2.3.4 Screen and diagnosis:

To look for polyps or cancer in people who don't have any symptoms [55]:

- Faecal Occult Blood Test FOBT: This is first procedure done for looking blood in stool to detect any polyps exist.
- Colonscopy: It process allows to look for polyps or diseases inside large intestine.
- Imaging Test:
  - Computed Tomograph CT Scan: It takes pictures for abdomen to see how cancer spread on second layer of colon wall.
  - Positron Emission Tomograph PET: It process used when the pictures from CT scan aren't clear.
  - MRI Scan.
- Blood Test:
- Complete Blood Count CBC: It measures number of white and red blood cells and platelets.
- Chemistry Profile: It measures Carcino Embyonic Antigen CEA that occurs when cancer spread.
- Molecular Testing: Takes sample tissue to examine a genes that have affect on treatment.
- RAS Mutation: It is protein exist in cells. Cancer cells have a control of this protein so treatment can't work.

- BRAF: It helps to determine prognosis, and it done after RAS mutation if the result is normal.
- Sigmoidoscopy: It uses to detect polyps by using lighted tube.
- Double Contrast Barium Enema: It used enema with barium solution and air, and they help to show polyps in colon and rectum also.
- Ultrasound: It take pictures by using sound waves, and it has two types, but use one type to detect colon cancer, it is Abdominal Ultrasound : It uses to check cancer spread to liver .

3.2.3.5 Staging :

TNM classification system is used to determine the stage of colon cancer. And those tables below show the process [54] .

Table(3.10) Determining Tumor Characteristics for colon cancer .

Tis:	No tumor in the mucosa.
T1:	Tumor in the submucosa.
T2:	Tumor in muscularia.
T3:	Tumor in the serosa or adventitia.
T4a:	Tumor through the serosa.
T4b:	Tumor next to or into organs and structures.

T: It tells which tissues the primary tumor has grown.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.11) Regional lymph node status for colon cancer .

No:		No evidence of cancer near lymph nodes.
N1:		Cancer has spread to 1-3 near lymph nodes.
	N1a:	Cancer has spread to 1 near lymph nodes.
	N1b:	Cancer has spread to 2-3 near lymph nodes.
	N1c:	Cancer deposits inside or outside the colon wall.
N2:		Cancer has spread to 4 or more near lymph nodes.
	N2a:	Cancer has spread to 4-6 near lymph nodes.
	N2b:	Cancer has spread to 7 or more near lymph nodes.

N: Tells where cancer has spread with near lymph node.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(3.12) Determining Metastatic Status for colon cancer .

M0:	No evidence of cancer has spread to distant sites.

M1:		Cancer has spread to distant sites.
	M1a:	Cancer has spread to one distant sites.
	M1b:	Cancer has spread to two or more distant sites.

M: Tells if cancer spread distant sites.
Source: UICC TNM classification of malignant tumours (UICC, 2002).

Table(2.13) The Stage of colon cancer [56] .

Stage		**T**	**N**	**M**
Stage0		Tis	N0	M0
Stage1		T1	N0	M0
		T2	N0	M0
Stage11	**Stage11A**	T3	N0	M0
	Stage11B	T4a	N0	M0
	Stage11C	T4b	N0	M0
Stage111	**Stage111A**	T1-T2	N1/N1c	M0
	Stage111B	T3-T4a	N1/N1c	M0
		T2-T3	N2a	M0
		T1-T2	N2b	M0
	Stage111C	T4a	N2a	M0
		T3-T4a	N2a	M0
		T4b	N1-N2	M0
Stage1V	**Stage1VA**	Any T	Any N	M1a
	Stage1VA	Any T	Any N	M1b

Source: UICC TNM classification of malignant tumours (UICC, 2002).

3.2.3.6 Treatment:

colon cancer treatment depending on the type of cancer and its stage [54].

- Surgery**:** There are two types of surgery:
  - Colectomy: It removes a part of colon that affected by cancer cells and it has two method:
    - Open: Removes tissues by cutting a large part of an abdominal.
    - Laparoscopy: Cuts small part of colon that affected with cancer.
  - Lymph Adenectomy: It removes affected lymph nodes.

- Metastasectomy: Surgery to remove metastases and it depend where cancer spread.

- Ablation : It to destroy small tumors and it has three types:
  - Cryoablation: Freeze and kill cancer cells by nitrogen.
  - Radio Frequency: Use high energy wave to kill cancer cells.
  - Micro wave ablation: Use high energy wave to kill cancer cells.
- Radiation Therapy: Use high energy rays, the rays damage DNA and kill the cancer cells or to stop new cancer cells, it has two type:
  - External Radiation: It is machine called Linear Accelerator.
  - Internal Radiation: It is a tube insert into near tube.
- Chemotherapy: It has many drugs, most of them are a liquid and a dose depend of the stage, and it uses to kill and or slow growth of cancer cells.

Table (3.14) Description of Treatment for Colon cancer.

Stage	Treatment
Stage0	Surgery (polypectomy)
Stage1	Surgery
Stage11	Surgery+ Chemotherapy
Stage111	Surgery+ Chemotherapy
Stage1V	Surgery+ Radiation therapy+ Chemotherapy +Targeted therapy

Source: UICC TNM classification of malignant tumours (UICC, 2002).

CHAPTER FOUR
(Application)

3.1 Preface
3.2 Results and Interpretations

4.1 Preface:

In this chapter we fitted the LC model to Egyptian data. The SVD and MLE were used to estimate the model's parameters, which were presented in Chapter two and comparison between two methods based on ME and MSE from equations(2.39) and (2.40) respectively, after the method has been chosen, RWD (0,1,0) was used to forecast the mortality index as in original LC model . While performance of forecasting based on MPE from equations (2.43) .the data from 2001-2004, the aged-group began fr (5-9) to (70-74) because in raw data of population has missing observations for age 0 to 4 and over age 74 for some years. The aim was forecasting age-specific cancer mortality rate for years from 2015-2020 for both sex (male, female) for all types of cancer. The results were obtained via ilc, demographic, forecast and R packages.

4.2 Results and Interpretations:

This section showed the results which were obtained after fitting LC model and represented in tables and figures for oral , lung and colon cancer for male and female.

4.2.1 Oral Cancer:

4.2.1.1 The Singular Value Decomposition (SVD):

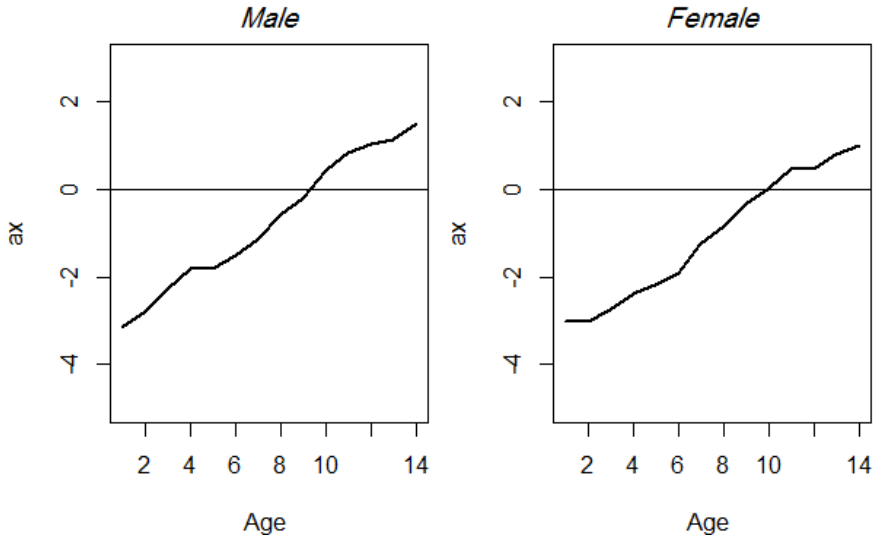
According to LC model we obtained the parameter a_x first from equation (2.13). We had the following table and figure as a result.

Table (4.1) Estimation of a_x by SVD for oral cancer.

Age	Male	Female
5-9	-3.1510724	-3.01883095
10-14	-2.7893000	-3.00847642
15-19	-2.2734324	-2.75196732
20-24	-1.8151230	-2.39459497
25-29	-1.7998020	-2.18171296
30-34	-1.4929181	-1.90279304
35-39	-1.1527998	-1.25411272
40-44	-0.5822332	-0.83834565
45-49	-0.1890201	-0.33331201
50-54	0.4469852	0.03577805
55-59	0.8441306	0.46944624
60-64	1.0256584	0.45567385
65-69	1.1504057	0.81195364
70-74	1.5065264	1.01323082

Source: Author calculation by ilc and demography and R.

Figure (4.1) General pattern of mortality *ax* by SVD for oral cancer..



Source: Author plotted by ilc and demography and R.

Table(4.1) shows the values of a_x ,which represents the general pattern (age shape) of mortality by age for both sex (male-female), and Figure(4.1) shows the pattern of a_x and values of a_x is increasing overtime for both sex (male, female), and this indicates that they have up trend in mortality and the younger ages have lower mortality than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate.

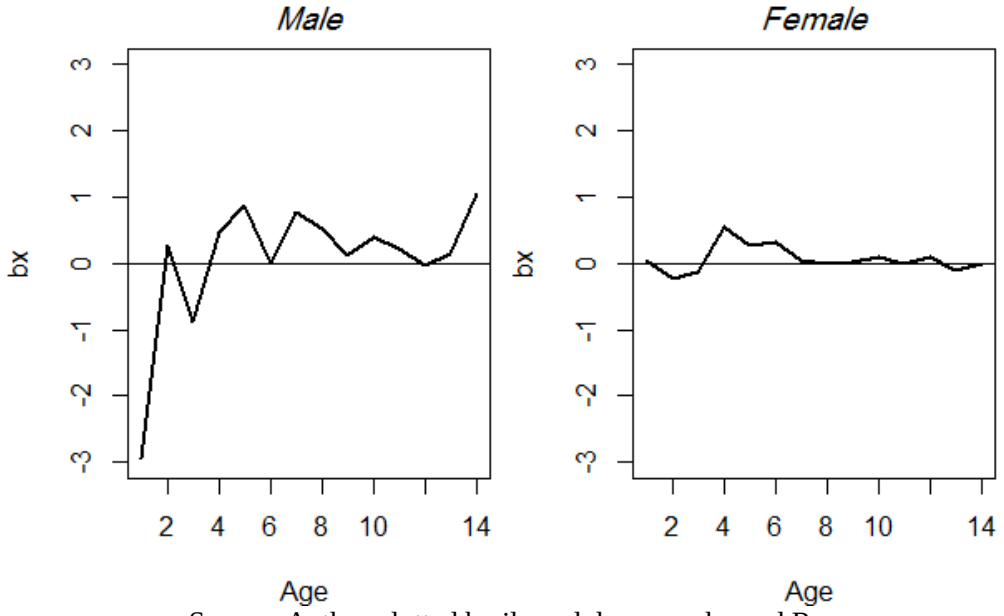
The second step is estimated the parameter b_x from the equation (2.17).

Table (4.2) Estimation of b_x by SVD for oral cancer.

Age	Male	Female
5-9	-2.942069258	0.044085855
10-14	0.281733207	-0.235875239
15-19	-0.896142231	-0.128941587
20-24	0.467421347	0.546850357
25-29	0.873450655	0.282011410
30-34	-0.006388266	0.333863726
35-39	0.782910137	0.055829625
40-44	0.534280613	-0.005215547
45-49	0.132409604	0.028318662
50-54	0.395859207	0.099419405
55-59	0.208887508	-0.002093927
60-64	-0.023123039	0.097542945
65-69	0.138784924	-0.104899960
70-74	1.051985593	-0.010895728

Source: Author calculation by ilc and demography and R.

Figure(4.2) General pattern of mortality b_x by SVD for oral cancer..



Source: Author plotted by ilc and demography and R.

Table (4.2) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality rate changes. The figure (4.2) shows the cancer mortality change for younger ages for male, and the cancer mortality among younger ages have highest values. For female the values of b_x are invariant for age-group (35-39) to (70-74). The high values of b_x indicate improvement in mortality rate at these

ages, while the negative values at some ages indicate that mortality rate is increasing.

The first estimated of the parameter k_t from the equation(2.18) and re-estimated of k_t from equation(2.19).

Table(4.3) First and second estimation of k_t by SVD for oral cancer..

Year	Male		Female	
	1st estimation	2nd estimation	1st estimation	2nd estimation
2001	-0.18627398	-0.300521276	0.48253481	-1.4341430
2002	0.06209605	0.104478547	0.83134614	-0.5754702
2003	-0.23507346	0.165772443	1.12271595	2.5655038
2004	0.79914156	0.896409136	-1.15508876	3.1201887
2005	-0.06076030	0.343221516	0.90232747	3.0711050
2006	0.18283313	0.447161745	-0.69555331	0.6263540
2007	-0.21302982	0.364179609	-0.01488561	2.4640554
2008	-0.14352936	-0.835571537	1.85151387	1.2093693
2009	0.02829768	-0.489343914	0.66933300	-0.2348930
2010	-0.21099560	-0.006261097	0.54407939	1.9223505
2011	0.31412619	0.329309700	0.61531928	1.1233423
2012	-0.10032421	-0.381246106	-2.42249259	-3.5903471
2013	-0.17292320	-0.257155280	-0.30412515	-1.0497980
2014	-0.06358468	-0.348792267	-2.42702448	-3.6195502

Source: Author calculation by ilc and demography and R.

Figure (4.3) General pattern for *kt* 2001–2014 by SVD for oral cancer.

Source: Author plotted by ilc and demography and R.

Table (4.3) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure (4.3) shows the

mortality index k_t has non-linear trend overtime for male and female. The high values of k_t indicate there is no improvement of cancer mortality rate. For male highest values of $k_t$in 2005 to 2014 except 2007, while for female in 2004.

4.2.1.2 Maximum Likelihood Estimation (MLE):

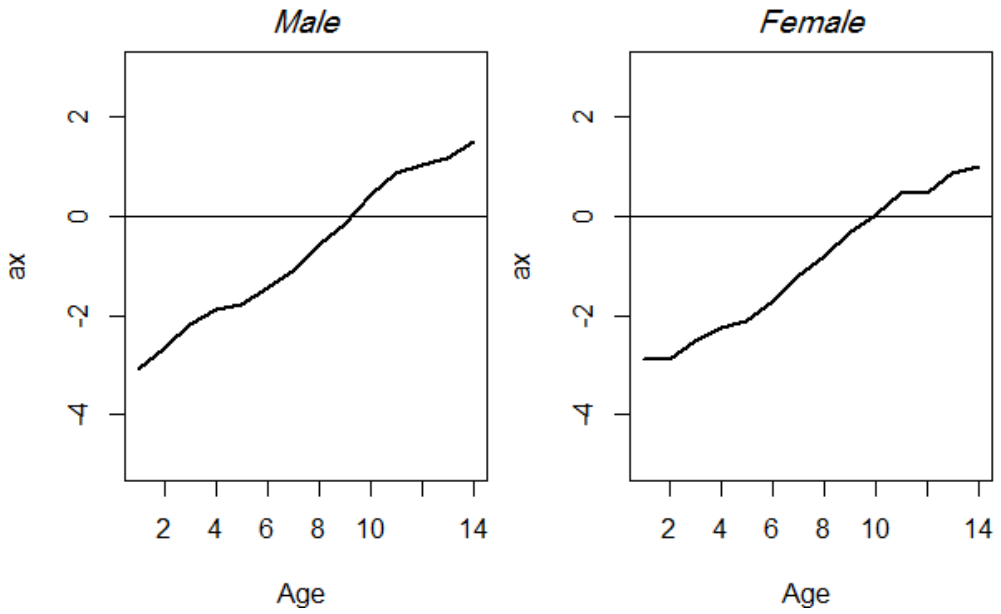
After fitting the technique of (MLE), we obtained at the following results. We obtained the parameter a_x first from equation (2.30).

Table (4.4) Estimation of a_x by MLE for oral cancer.

Age	Male	Female
5-9	-3.0705606	-2.89066781
10-14	-2.6407957	-2.87050367
15-19	-2.1869090	-2.52602793
20-24	-1.8805788	-2.23086308
25-29	-1.7761237	-2.11437286
30-34	-1.4549623	-1.72548413
35-39	-1.0939142	-1.21396365
40-44	-0.5623586	-0.78953446
45-49	-0.1649595	-0.32640467
50-54	0.4334790	0.04276143
55-59	0.8563559	0.47494393
60-64	1.0382069	0.48399334
65-69	1.1742353	0.85666320
70-74	1.4995503	1.00656041

Source : Author calculation by ilc and demography and R.

Figure(4.4) General pattern of ax by MLE for oral cancer.



Source: Author plotted by ilc and demography and R.

Table(4.4) shows the values of a_x ,which represents the general pattern (age shape) of mortality by age for both sex (male-female), and Figure(4.4) shows the pattern of a_x and the values of a_x is increasing overtime for both sex (male, female), and this indicates that they have up trend in mortality and the younger ages have lower mortality rate than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate.

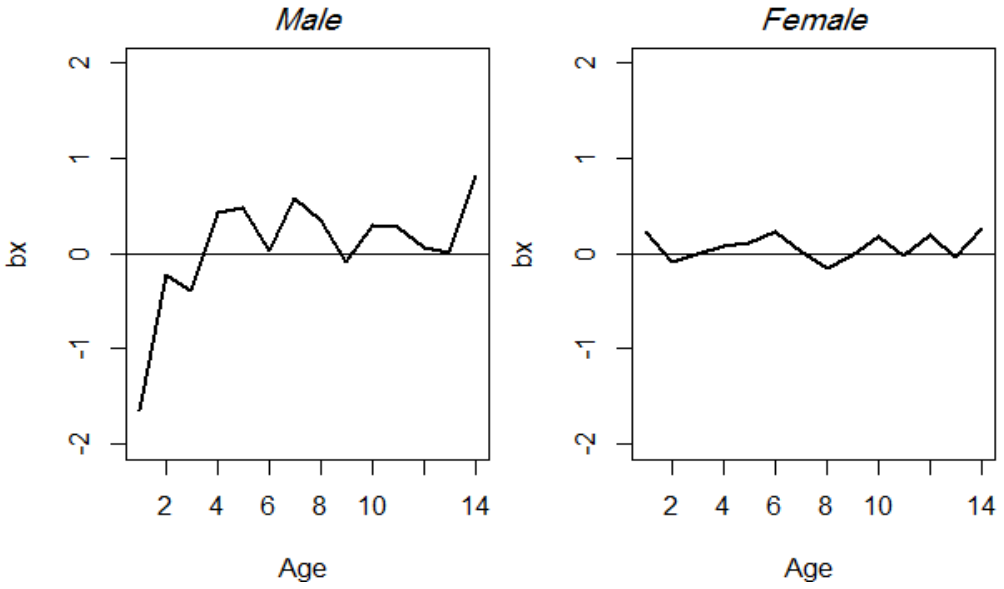
The second step is estimated the parameter b_x from the equation(2.30).

Table(4.5) Estimation of b_x by MLE for oral cancer.

Age	Male	Female
5-9	-1.649627248	0.227661450
10-14	-0.223681685	-0.093650929
15-19	-0.389310324	0.001912058
20-24	0.439891540	0.077170236
25-29	0.483421717	0.107314822
30-34	0.024558749	0.228254773
35-39	0.586616390	0.028336567
40-44	0.345512053	-0.148137905
45-49	-0.091753624	-0.018331061
50-54	0.306361254	0.176181377
55-59	0.273155000	-0.023376550
60-64	0.069284626	0.203461461
65-69	0.007110983	-0.037895593
70-74	0.818460569	0.271099295

Source : Author calculation by ilc and demography and R.

Figure(4.5) General pattern of b_x by MLE for oral cancer.



Source: Author plotted by ilc and demography and R.

Table (4.5) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality changes. The figure (4.5) shows the cancer mortality has a negative values for younger ages for male, and positive values for older ages while the middle ages have invariant mortality. For female the values of b_x are closet for all ages. The high values of b_x indicate improvement in mortality rate at these ages, while the negative values at some ages indicate that mortality rate is increasing.

The parameter k_t estimated from the equation(2.30)

Table(4.6) Estimation for k_t by MLE for oral cancer.

Year	Male	Female
2001	-0.4004024	-1.8120754
2002	-0.1055330	-1.6504141
2003	-0.2934653	0.2681107
2004	1.1910454	0.9915315
2005	0.1075799	1.5327918
2006	0.2990911	-0.9234152
2007	0.1715210	0.6919523
2008	-0.3695789	1.2313918
2009	-0.2834992	-0.6546531
2010	-0.3496329	1.3747088
2011	0.5096010	2.0000210
2012	-0.2277883	-0.6955708
2013	-0.1063316	-0.2939659
2014	-0.1426069	-2.0604135

Source: Author calculation by ilc and demography and R.

Figure(4.6) General pattern of k_t 2001-2014 by MLE for oral cancer.

Source: Author plotted by ilc and demography and R.

Table (4.6) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure (4.6) shows the mortality index k_t has non-linear trend for 2001-2006 for male and nonlinear trend for female. The high values of k_t indicate there is no improvement of cancer mortality rate, and year 2004 has highest mortality rate for male, while for female have high mortality in 2012.

4.2.1.3 Comparison between SVD and MLE:

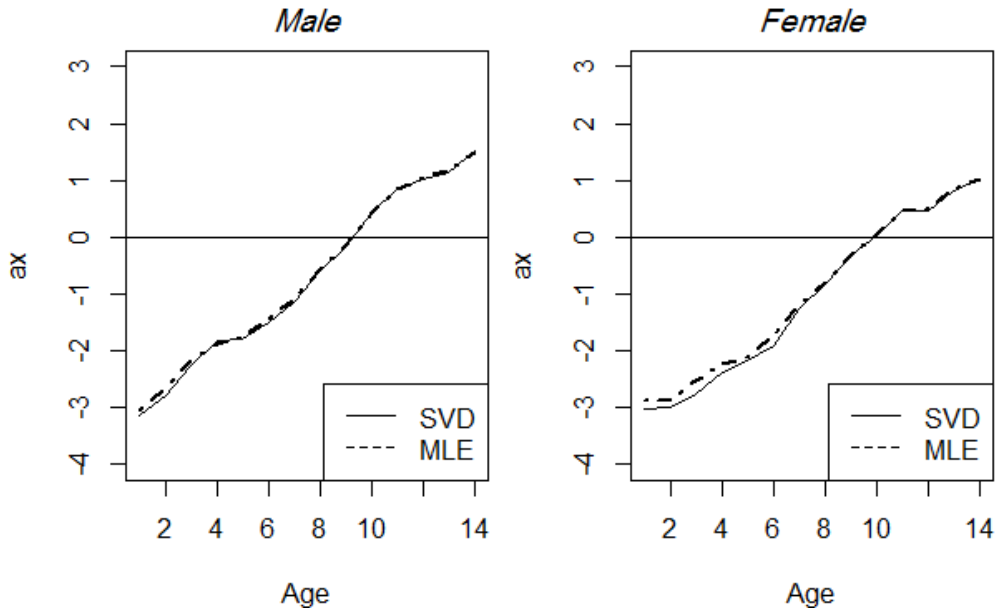
The ME and MSE are obtained from equations (2.39) and (2.40) respectively from chapter two .

Table(4.7) Comparison between SVD and MLE for estimation a_x for oral cancer.

Age	Male		Female	
	SVD	MLE	SVD	MLE
5-9	-3.1510724	-3.0705606	-3.01883095	-2.89066781
10-14	-2.7893000	-2.6407957	-3.00847642	-2.87050367
15-19	-2.2734324	-2.1869090	-2.75196732	-2.52602793
20-24	-1.8151230	-1.8805788	-2.39459497	-2.23086308
25-29	-1.7998020	-1.7761237	-2.18171296	-2.11437286
30-34	-1.4929181	-1.4549623	-1.90279304	-1.72548413
35-39	-1.1527998	-1.0939142	-1.25411272	-1.21396365
40-44	-0.5822332	-0.5623586	-0.83834565	-0.78953446
45-49	-0.1890201	-0.1649595	-0.33331201	-0.32640467
50-54	0.4469852	0.4334790	0.03577805	0.04276143
55-59	0.8441306	0.8563559	0.46944624	0.47494393
60-64	1.0256584	1.0382069	0.45567385	0.48399334
65-69	1.1504057	1.1742353	0.81195364	0.85666320
70-74	1.5065264	1.4995503	1.01323082	1.00656041

Source: Author calculation by ilc and demography and R.

Figure(4.7) Comparison between SVD and MLE for estimation a_x for ora l cancer.



Source: Author plotted by ilc and demography and R.

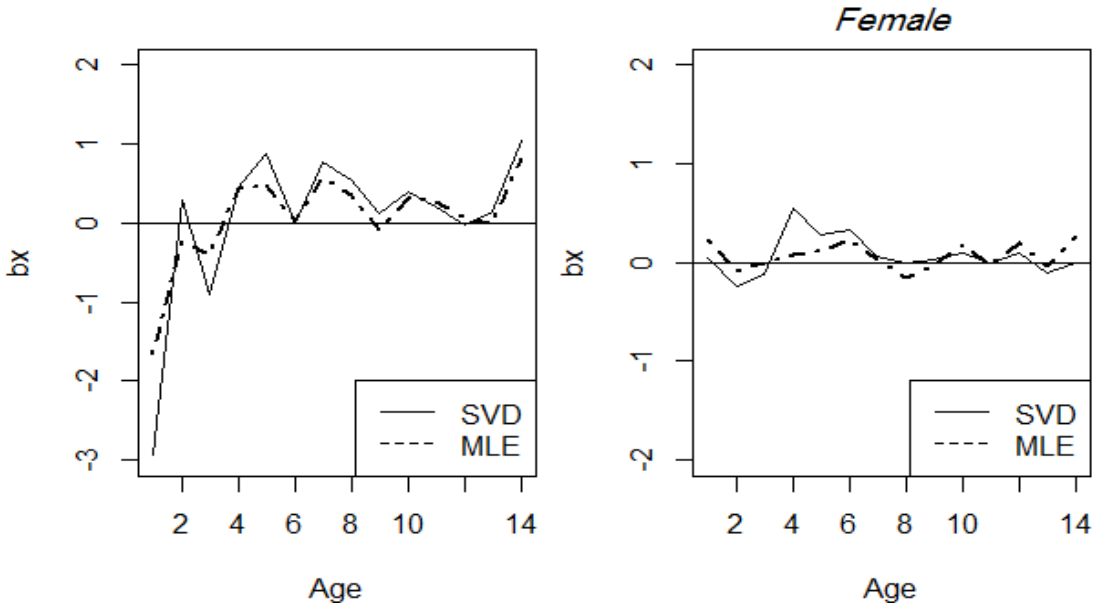
If we take a look to table (4.7) and figure (4.7), we will notice that the esti mation of parameter a_x from SVD and MLE is slight difference and this is very clear in the figure (4.7) for both sex (male, female).The maximum difference value of estimation a_x is 0.1485043 in age-group (10-14) for male, while for female is 0.2259394 in age-group (15-19).

Table(4.8) Comparison between SVD and MLE for estimation b_x for oral cancer.

Age	Male		Female	
	SVD	MLE	SVD	MLE
5-9	-2.942069258	-1.649627248	0.044085855	0.227661450
10-14	0.281733207	-0.223681685	-0.235875239	-0.093650929
15-19	-0.896142231	-0.389310324	-0.128941587	0.001912058
20-24	0.467421347	0.439891540	0.546850357	0.077170236
25-29	0.873450655	0.483421717	0.282011410	0.107314822
30-34	-0.006388266	0.024558749	0.333863726	0.228254773
35-39	0.782910137	0.586616390	0.055829625	0.028336567
40-44	0.534280613	0.345512053	-0.005215547	-0.14813791
45-49	0.132409604	-0.091753624	0.028318662	-0.018331061
50-54	0.395859207	0.306361254	0.099419405	0.176181377
55-59	0.208887508	0.273155000	-0.002093927	-0.02337655
60-64	-0.023123039	0.069284626	0.097542945	0.203461461
65-69	0.138784924	0.007110983	-0.104899960	-0.037895593
70-74	1.051985593	0.818460569	-0.010895728	0.271099295

Source: Author calculation by ilc and demography and R.

Figure(4.8) comparison between SVD and MLE for estimation b_x for oral cancer.



Source: Author plotted by ilc and demography and R.

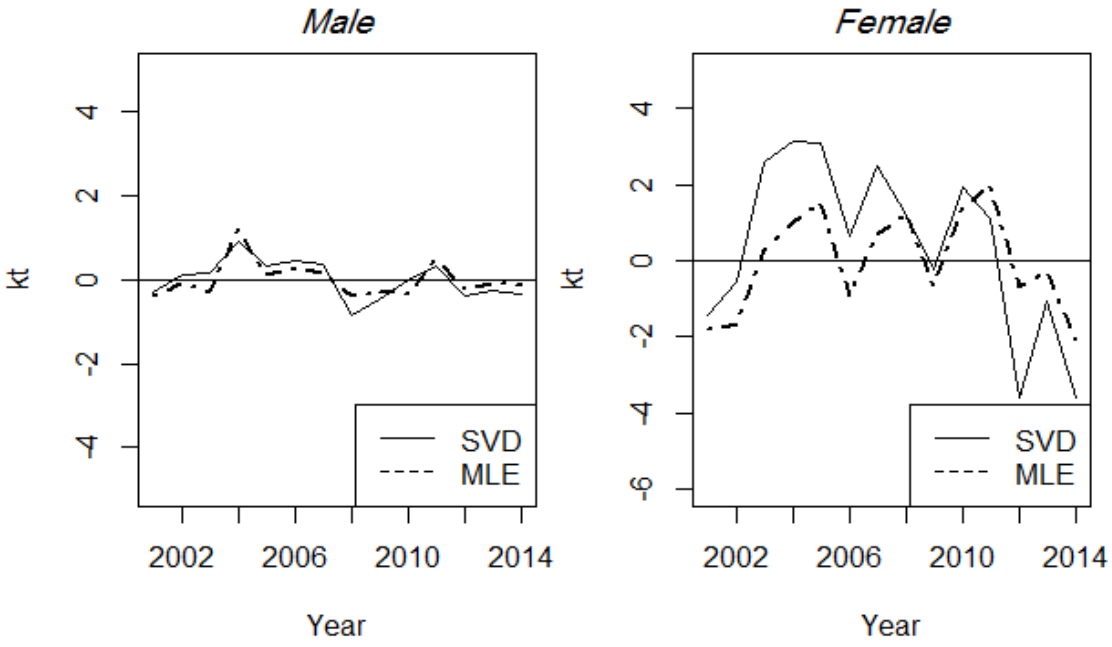
If we take a look to table (4.8) and figure (4.8), we will notice that the esti mation of parameter b_x from SVD and MLE is slight difference for both s ex (male, female), and this is very clear in the figure (4.8). The maximum difference value of estimation b_x is 1.292442 in age-group (5-9) for male , while for female is 0.4696801 in age-group (20-24).

Table(4.9) Comparison between SVD and MLE for estimation $k_t$for oral cancer.

Year	Male		Female	
	SVD	MLE	SVD	MLE
2001	-0.300521276	-0.4004024 0.1	-1.4341430	-1.8120754
2002	04478547	-0.1055330	-0.5754702	-1.6504141
2003	0.165772443	-0.2934653	2.5655038	0.2681107
2004	0.896409136	1.1910454	3.1201887	0.9915315
2005	0.343221516	0.1075799	3.0711050	1.5327918
2006	0.447161745	0.2990911	0.6263540	-0.9234152
2007	0.364179609	0.1715210	2.4640554	0.6919523
2008	-0.835571537	-0.3695789	1.2093693	1.2313918
2009	-0.489343914	-0.2834992	-0.2348930	-0.6546531
2010	-0.006261097	-0.3496329	1.9223505	1.3747088
2011	0.329309700	0.5096010	1.1233423	2.0000210
2012	-0.381246106	-0.2277883	-3.5903471	-0.6955708
2013	-0.257155280	-0.1063316	-1.0497980	-0.2939659
2014	-0.348792267	-0.1426069	-3.6195502	-2.0604135

Source: Author calculation by ilc and demography and R.

Figure(4.9) Comparison between SVD and MLE for estimation of k_t 200 1-2014 for oral cancer.



Source: Author plotted by ilc and demography and R.

If we take a look to table (4.9) and figure (4.9), we will notice that the esti mation of parameter k_t from SVD and MLE is difference for both sex (m ale, female), and this is very clear in the figure (4.9) for both sex (male, fe male). The maximum difference value of estimation k_t is 0.4659926 in y ear 2008 for male, while for female is 2.894776 in year 2012.

Table (4.10) Comparison between SVD and MLE for Errors based on log mortality rate across ages for oral cancer.

Sex	Method	ME	MSE
Male	SVD	0.00016	0.25208
	MLE	0.03162	0.15594
Female	SVD	0.02856	0.32310
	MLE	0.07680	0.22285

Source: Author calculation by ilc and demography and R.

From table(4.10) the SVD is better Than MLE for both sex (male, female) with errors (ME=0.00016, 0.02856) respectively.

4.2.1.4 Forecast k_t and Age-specific Cancer Mortality Rate:

After obtained the values of k_t from SVD for male and female. We obtai ned forecast the mortality index from equations(2.31) , drift ($\hat{d}$) from equ ation(2.32), standard error ($\widehat{s.e}$) from equation(2.54) and error ($\hat{\sigma}^2$)from e quation(2.37) .

Table(4.11) Estimation of drift, standard error and errors of RWD (0,1,0) for oral cancer.

Sex	**Male**	**Female**
Method	**SVD**	**SVD**
$\widehat{d}$	-0.00371	-0.16811
$\widehat{s.e}$	0.1431	0.3795
$\widehat{\sigma}^2$	0.2883	5.3441

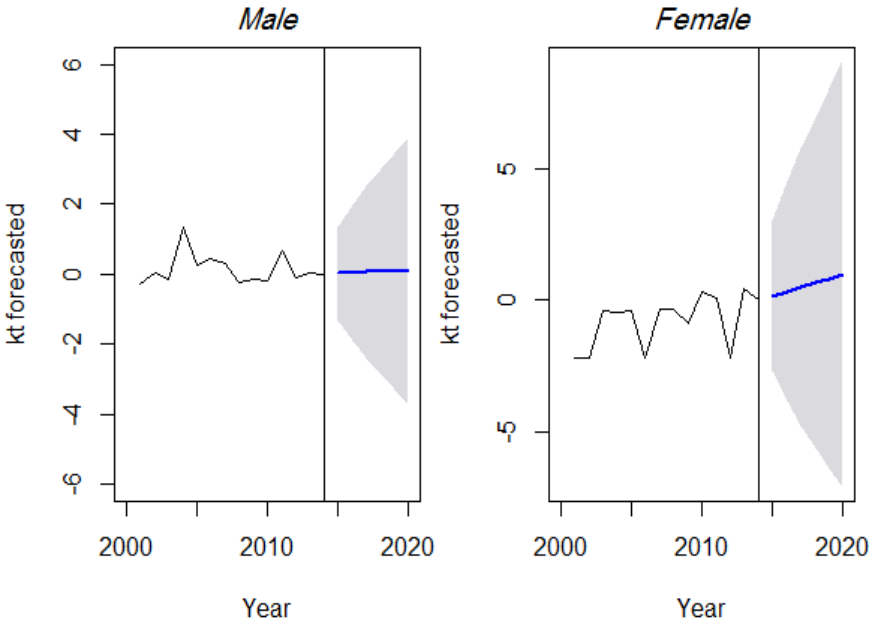
Source: Author calculation by ilc and demography, forecast and R.

Table (4.12) Forecast mortality index for 2015–2020 for oral cancer**.**

Year	**Male**			**Female**		
	k_t**forecast**	**lower**	**upper**	k_t**forecast**	**lower**	**Upper**
2015	-0.00371315	-1.095732	1.08831	-0.1681082	-4.870041	4.533825 6
2016	-0.00742631	-1.605978	1.59113	-0.3362165	-7.219142	.546709 8.
2017	-0.01113946	-2.033165	2.01089	-0.5043247	-9.210613	201963 9.6
2018	-0.01485261	-2.421544	2.39184	-0.6724330	-11.034989	90123 11.0
2019	-0.01856577	-2.787338	2.75021	-0.8405412	-12.762117	81035 12.4
2020	-0.02227892	-3.138429	3.09387	-1.0086495	-14.425941	08642

Source: Authors calculation by ilc and demography, forecast and R.

Figure(4.10) Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for oral cancer.



Source: Author plotted by ilc and demography and R.

The table (4.11) and (4.12) show the values of drift, standard error, errors and k_t for both sex (male, female). figures(4.10) shows that the k_t is incr easing for both sex (male, female) .

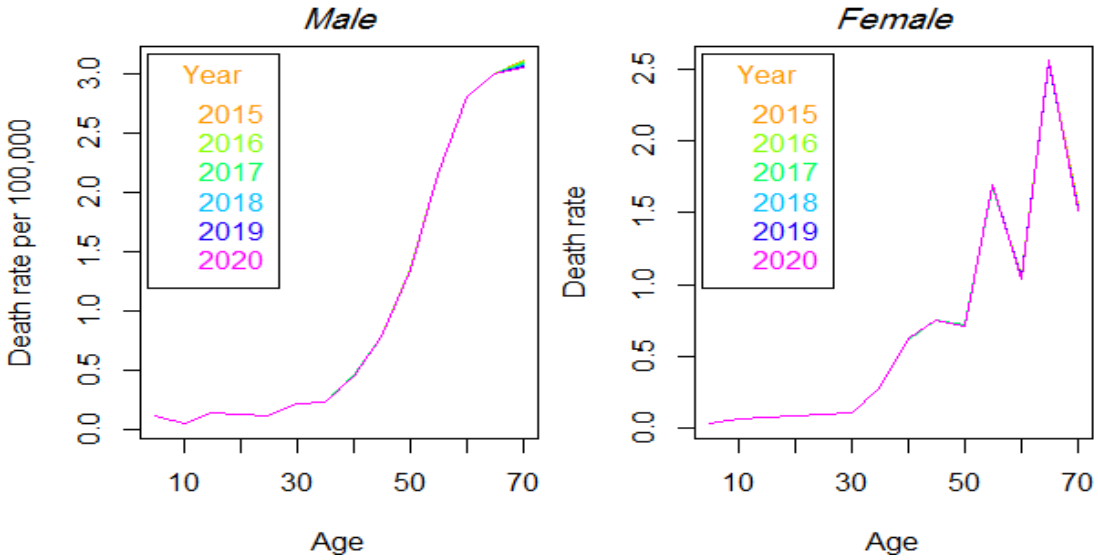
The age-specific cancer mortality rate, $m_{x,t}$ is now forecasting for years 2015-2020 from the equation(2.30).

Table (4.13) Forecast age-specific mortality rate for period 2015–2020 for oral cancer per 100.000.

Sex	Age	2015	2016	2017	2018	2019	2020
Male	5-9	0.12	0.12	0.12	0.12	0.13	0.13
	10-14	0.06	0.06	0.06	0.06	0.06	0.06
	15-19	0.14	0.14	0.14	0.14	0.14	0.14
	20-24	0.14	0.14	0.14	0.14	0.14	0.14
	25-29	0.12	0.12	0.12	0.12	0.12	0.12
	30-34	0.23	0.23	0.23	0.23	0.23	0.23
	35-39	0.24	0.24	0.24	0.24	0.24	0.24
	40-44	0.46	0.46	0.46	0.46	0.46	0.46
	45-49	0.79	0.79	0.79	0.79	0.79	0.79
	50-54	1.36	1.36	1.36	1.35	1.35	1.35
	55-59	2.16	2.16	2.16	2.16	2.15	2.15
	60-64	2.81	2.81	2.81	2.81	2.81	2.81
	65-69	3.01	3.01	3.01	3.00	3.00	3.00
	70-74	3.11	3.10	3.09	3.08	3.07	3.05
Female	5-9	0.04	0.04	0.04	0.04	0.04	0.04
	10-14	0.12	0.13	0.13	0.14	0.14	0.15
	15-19	0.10	0.11	0.11	0.11	0.11	0.12
	20-24	0.01	0.01	0.01	0.01	0.01	0.01
	25-29	0.04	0.04	0.04	0.03	0.03	0.03
	30-34	0.04	0.04	0.04	0.04	0.03	0.03
	35-39	0.23	0.23	0.23	0.22	0.22	0.22
	40-44	0.44	0.44	0.44	0.44	0.44	0.44
	45-49	0.64	0.64	0.64	0.63	0.63	0.63
	50-54	0.71	0.70	0.69	0.68	0.67	0.65
	55-59	1.61	1.61	1.61	1.61	1.61	1.61
	60-64	1.09	1.07	1.05	1.04	1.02	1.00
	65-69	3.35	3.41	3.47	3.53	3.60	3.66
	70-74	2.87	2.88	2.88	2.89	2.89	2.90

Source: Author calculation by ilc and demography, forecast and R.

Figure (4.11) Forecast age-specific cancer mortality rate 2015-2020 for oral cancer.



Source: Author plotted by ilc and demography and R.

Table (4.13) and finger (4.11) show the values and pattern of age-specific cancer mortality rate for both sex (male, female). The mortality rate is increasing for all ages for male. While for female the mortality rate is increasing and decreasing on ages.

Table (4.14) Model's forecast errors based on mortality rate across ages for oral cancer.

Sex	Method	MPE
Male	SVD	0.14960
Female	SVD	0.25731

Source: Author calculation by ilc and demography, forecast and R.

The table (4.14) shows the errors for the forecasting age-specific cancer mortality rate for both sex (male, female), and they are satisfactory well for both sex (male, female).

4.2.2 Lung Cancer:

4.2.2.1 The Singular Value Decomposition (SVD):

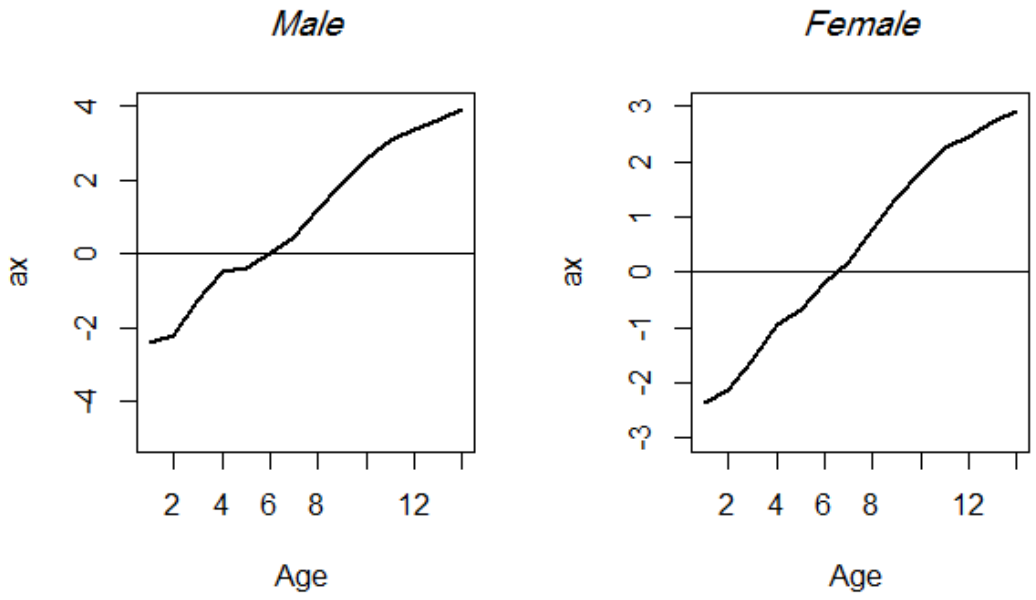
According to LC model we obtained the parameter a_x first from equation (2.13). We had the following table and figure as a result.

Table(4.15) Estimation of a_x by SVD for lung cancer.

Age	Male	Female
5-9	-2.3936008	-2.3535010
10-14	-2.2321204	-2.1380211
15-19	-1.2540668	-1.6145002
20-24	-0.4781030	-0.9532129
25-29	-0.3663091 0.016268	-0.6822896
30-34	9 0.4718716	-0.1978954
35-39	1.1973920 1.904633	0.1714027
40-44	7 2.5889595	0.7748322
45-49	3.0610082 3.357705	1.3440209
50-54	7 3.6424854	1.8404448
55-59	3.9286976	2.2675470
60-64		2.4544046
65-69		2.7047827
70-74		2.9080115

Source: Author calculation ilc and demography and R.

Figure (4.12) General pattern of mortality *ax* by SVD for lung cancer**.**



Source: Author plotted by ilc and demography and R.

Table (4.15) shows the values of a_x, which represents the general pattern (age shape) of mortality by age x for both sex (male-female), and

Figure(4.12) shows the pattern of a_x and the values of a_x is increasing over time for both sex (male, female), and this indicate that they have up trend in mortality and the younger ages have lower mortality rate than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate.

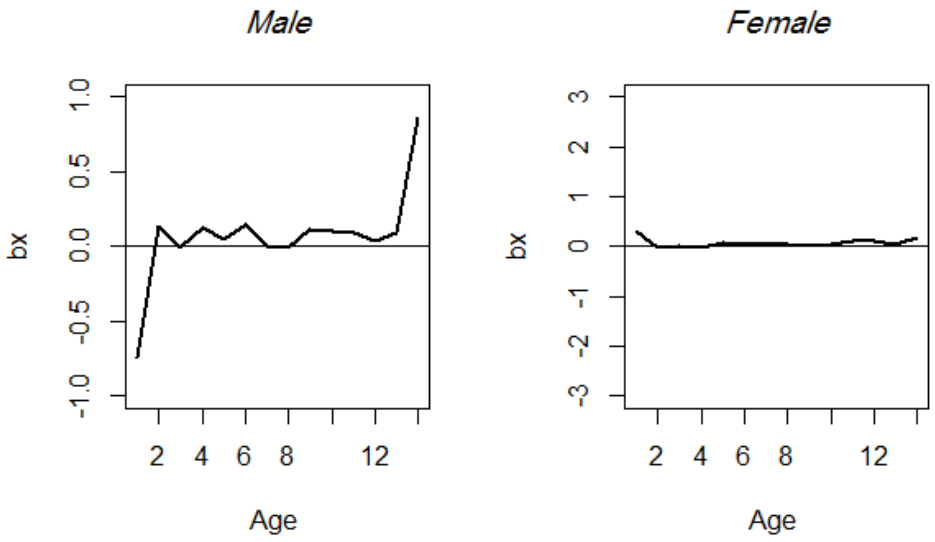
The second step is estimated the parameter b_x from the equation(2.17).

Table (4.16) Estimation of b_x for lung cancer.

Age	Male	Female
5-9	1.337894858	0.311528114
10-14	0.203900633	-0.020155534
15-19	-0.179014610	0.009752804
20-24	-0.595744360	-0.020111052
25-29	0.048773895	0.079049840
30-34	-0.008071041	0.057697675
35-39	-0.100369745	0.051034575
40-44	0.056773560	0.057117687
45-49	0.008372227	0.027837393
50-54	-0.048284805	0.030817762
55-59	0.114600679	0.099173565
60-64	0.156950197	0.096185995
65-69	0.006384318	0.050966660
70-74	-0.002165808	0.169104514

Source: Author calculation ilc and demography and R.

Figure(4.13) General pattern of b_x by SVD for lung cancer.



Source: Author plotted by ilc and demography and R.

Table (4.16) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality changes. The

figure (4.13) shows the mortality change for younger ages for male, and the mortality among younger ages have higher values. For female the mortality is constant and the values of b_x are invariant for all ages. The high values of b_x indicate improvement in mortality at these ages, while the negative values at some ages indicate that mortality rate is increasing. The parameter k_t first estimated from the equation (2.18) and re-estimated from equation(2.19).

Table (4.17) First and second estimation of k_t for lung cancer.

Year	Male		Female	
	1st estimation	2nd estimation	1st estimation	2nd estimation
2001	-0.935592329	-4.52625305	-4.27136967	-4.1957395
2002	0.018149564	-1.93422071	-2.22840608	-4.5891530
2003	-0.360310556	-1.933934z05	-1.18389375	-3.3751181
2004	-1.545523361	0.46087834	1.36216373	1.1693789
2005	-0.003563399	1.71249586	0.36074288	2.1449693
2006	0.248684239	0.74305146	-0.33700795	0.5684068
2007	0.184202959	0.70544229	1.93415254	0.4722695
2008	0.729029402	0.01372249	-2.19514057	0.9711164
2009	0.341494954	0.34186040	2.23905117	1.7808479
2010	0.241479144	1.61211845	1.38218424	1.9337381
2011	0.275597844	1.66845237	0.94080801	1.3501663
2012	0.156866767	0.95335131	1.78545130	2.4316543
2013	0.561147926	1.41541856	-0.06460958	-0.6172925
2014	0.088336844	1.01183937	0.27587374	0.9801094

Source: Author calculation ilc and demography and R.

Figure (4.14) General pattern of *kt* 2001–2014 by SVD for lung cancer.

Source: Author plotted by ilc and demography and R.

Table (4.17) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure (4.14) shows the mortality index k_t has non-linear trend overtime for male and female. The high values of k_t indicate there is no improvement of cancer mortality rate. For male high values of k_t in 2005-2014 except 2007, while for female the highest value in years 2012.

4.2.2.2 Maximum Likelihood Estimation (MLE):

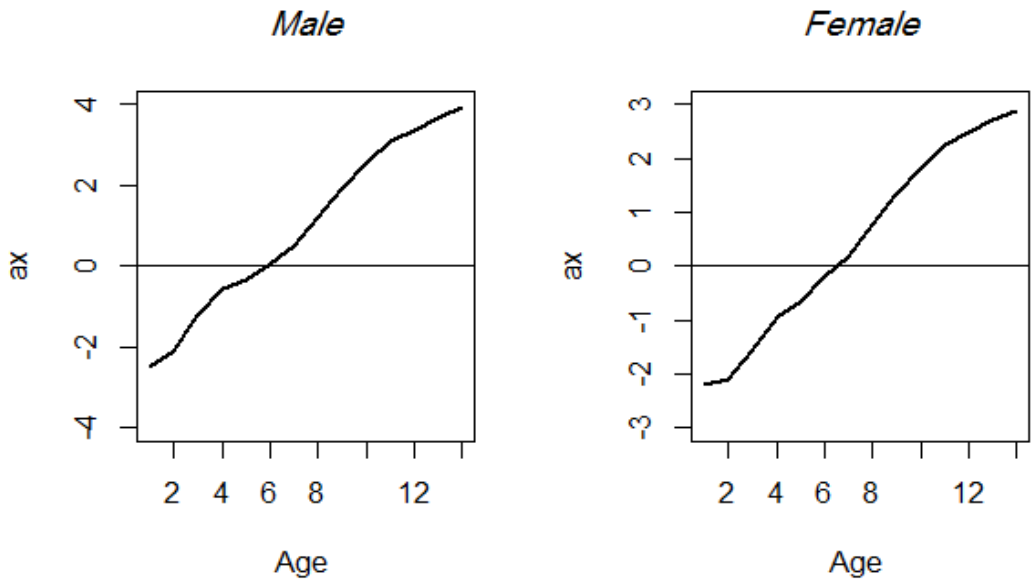
After fitting the technique of (MLE), we obtained at the following results. We obtained the parameter a_x first from equation (2.30).

Table(4.18) Estimation of a_x by MLE for lung cancer.

Age	Male	Female
5-9	-2.48703008	-2.1966573
10-14	-2.09432884	-2.0855666
15-19	-1.22450790	-1.5685070
20-24	-0.57469015	-0.9451742
25-29	-0.35102896	-0.6588752
30-34	0.04328598	-0.1695632
35-39	0.49722615	0.1921928
40-44	1.20430203	0.7754019
45-49	1.90416469	1.3456710
50-54	2.58533534	1.8371858
55-59	3.07773243	2.2595131
60-64	3.37980414	2.4729280
65-69	3.65489335	2.7176495
70-74	3.92969038	2.8867585

Source: Author calculation ilc and demography and R.

Figure(4.15) General pattern of ax by MLE for lung cancer.



Source: Author plotted by ilc and demography and R.

Table(4.18) shows the values of a_x, which represents the general pattern (age shape) of mortality by age for both sex (male-female), and Figure(4.15) shows the pattern of a_x and the values of a_x is increasing over time for both sex (male, female), and this indicates that they have up trend in mortality and the younger ages have lower mortality rate than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate.

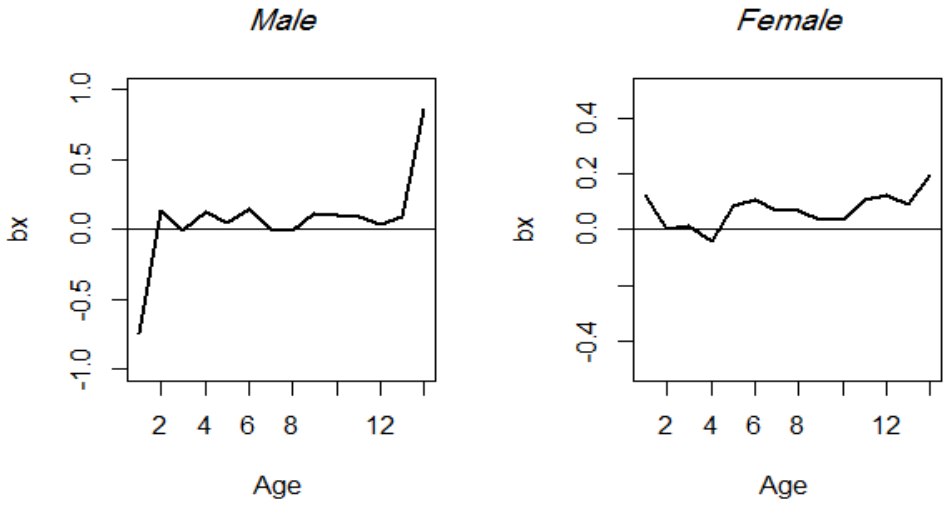
The second step is estimated the parameter b_x from the equation(2.30).

Table(4.19) Estimation of b_x by MLE for lung cancer.

Age	Male	Female
5-9	-0.739474294	0.1249235639
10-14	0.133800117	0.0004487961
15-19	-0.008422668	0.0111812489
20-24	0.129146698	-0.0436408872
25-29	0.052706689	0.0825098864
30-34	0.143112927	0.1080296991
35-39	0.005607565	0.0660035497
40-44	-0.008251806	0.0655079933
45-49	0.111165902	0.0344939491
50-54	0.104612823	0.0378819634
55-59	0.087534014	0.1054686606
60-64	0.040522396	0.1247160782
65-69	0.087139744	0.0896117518
70-74	0.860799893	0.1928637467

Source: Author calculation ilc and demography and R.

Figure(4.16) General pattern of b_x by MLE for lung cancer.



Source: Author plotted by ilc and demography and R.

Table (4.19) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality changes. the figure (4.16) shows the b_x has a negative value for younger ages for male, and positive value for older ages while the middle ages have invariant values. For female the values of b_x are closet for all ages. The high values of b_x indicate improvement in mortality at all ages, while the negative values at some ages indicate that mortality is increasing.

The parameter k_t estimated from the equation(2.30).

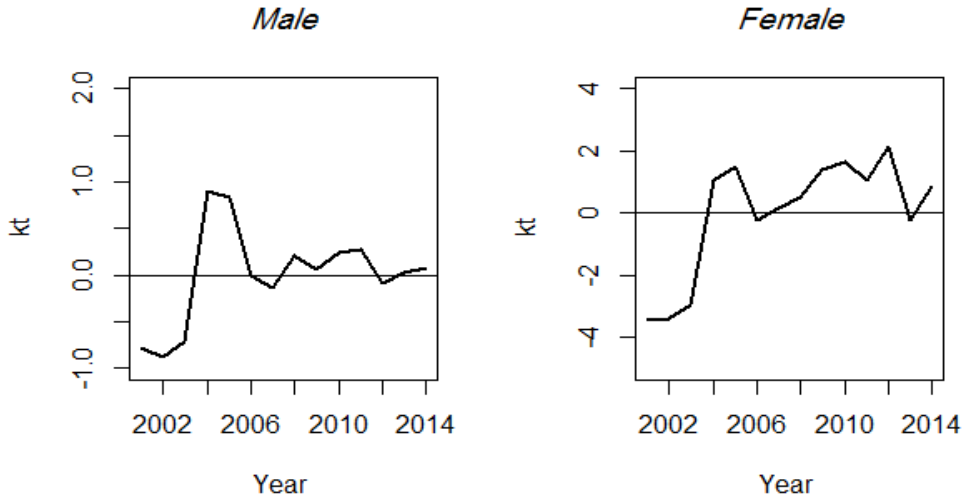
Table (4.20) Estimation of k_t by MLE for lung cancer.

Year	Male	Female
2001	-0.78370685	-3.4550937
2002	-0.88547995	-3.4175455
2003	-0.72001820	-2.9335145
2004	0.89754008	1.0571542
2005	0.83409272	1.5076962
2006	-0.01138519	-0.2544051
2007	-0.13434391	0.1393346
2008	0.20286007	0.4870224
2009	0.05924104	1.3801507
2010	0.24886670	1.6361492
2011	0.27205262	1.0673409

2012	-0.08322350	2.1486644
2013	0.02935940	-0.2163707
2014	0.07414499	0.8534167

Source: Author calculation ilc and demography and R.

Figure(4.17) General pattern of k_t 2001-2014 by MLE for lung cancer.



Source: Author plotted by ilc and demography and R.

Table (4.20) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure (4.17) shows the mortality index k_t has non-linear trend for male and female. The high values of k_t indicate there is no improvement of cancer mortality rate. For male have highest mortality rate in 2004, while for female have high mortality in years 2008 to 2014.

4.2.2.3 Comparison between SVD and MLE:

We obtained forecast the mortality index from equations (2.31) , drift ($\hat{d}$) from equation(2.32), standard error ($\widehat{s.e}$) from equation(2.54) and erro r ($\hat{\sigma}^2$) from equation(2.37) respectively.

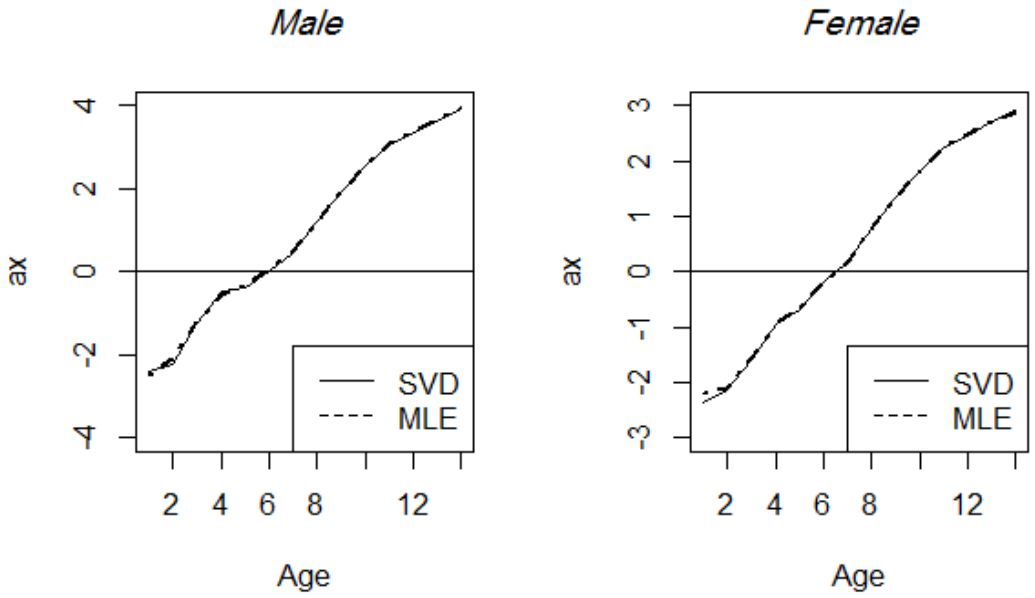
Table (4.21) Comparison between SVD and MLE for estimation of a_x for lung cancer.

Age	**Male**	**Female**

	SVD	MLE	SVD	MLE
5-9	-2.3936008	-2.48703008	-2.3535010	-2.1966573
10-14	-2.2321204	-2.09432884	-2.1380211	-2.0855666
15-19	-1.2540668	-1.22450790	-1.6145002	-1.5685070
20-24	-0.4781030	-0.57469015	-0.9532129	-0.9451742
25-29	-0.3663091	-0.35102896	-0.6822896 -	0.6588752
30-34	0.0162689	0.04328598	-0.1978954	-0.1695632
35-39	0.4718716	0.49722615	0.1714027	0.1921928
40-44	1.1973920	1.20430203	0.7748322	0.7754019
45-49	1.9046337	1.90416469	1.3440209	1.3456710
50-54	2.5889595	2.58533534	1.8404448	1.8371858
55-59	3.0610082	3.07773243	2.2675470	2.2595131
60-64	3.3577057	3.37980414	2.4544046	2.4729280
65-69	3.6424854	3.65489335	2.7047827	2.7176495
70-74	3.9286976	3.92969038	2.9080115	2.8867585

Source: Author calculation ilc and demography and R.

Figure (4.18) Comparison between SVD and MLE for estimation of a_x for lung cancer.



Source: Author plotted by ilc and demography and R.

If we take a look to table (4.21) and figure (4.18), we notice that there is slight difference between the estimation of parameter a_x from SVD and MLE, and this is very clear in the figure (4.18) for both sex (male,

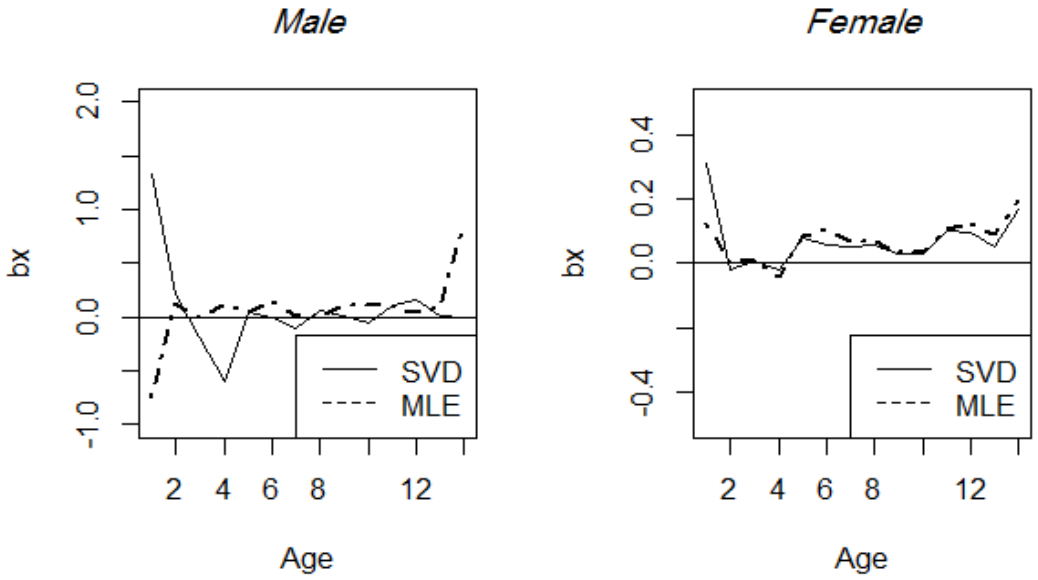
female). The maximum difference value of estimation of a_x for male is 0.1377916 in age-group (10-14), while for female is .1568437 in age-group (5-9).

Table (4.22) Comparison between SVD and MLE for estimation of b_x for lung cancer.

Age	Male		Female	
	SVD	MLE	SVD	MLE
5-9	1.337894858	-0.739474294	0.311528114	0.1249235639
10-14	0.203900633	0.133800117	-0.020155534	0.0004487961
15-19	-0.179014610	-0.008422668	0.009752804	0.0111812489
20-24	-0.595744360	0.129146698	-0.020111052	-0.0436408872
25-29	0.048773895	0.052706689	0.079049840	0.0825098864
30-34	-0.008071041	0.143112927	0.057697675	0.1080296991
35-39	-0.100369745	0.005607565	0.051034575	0.0660035497
40-44	0.056773560	-0.008251806	0.057117687	0.0655079933
45-49	0.008372227	0.111165902	0.027837393	0.0344939491
50-54	-0.048284805	0.104612823	0.030817762	0.0378819634
55-59	0.114600679	0.087534014	0.099173565	0.1054686606
60-64	0.156950197	0.040522396	0.096185995	0.1247160782
65-69	0.006384318	0.087139744	0.050966660	0.0896117518
70-74	-0.002165808	0.860799893	0.169104514	0.1928637467

Source: Author calculation ilc and demography and R.

Figure(4.19) Comparison between SVD and MLE for estimation of b_x for lung cancer.



Source: Author plotted by ilc and demography and R.

If we take a look to table (4.22) and figure (4.19), we will notice that the e stimation of parameter b_x. from SVD and MLE is a difference for male. a

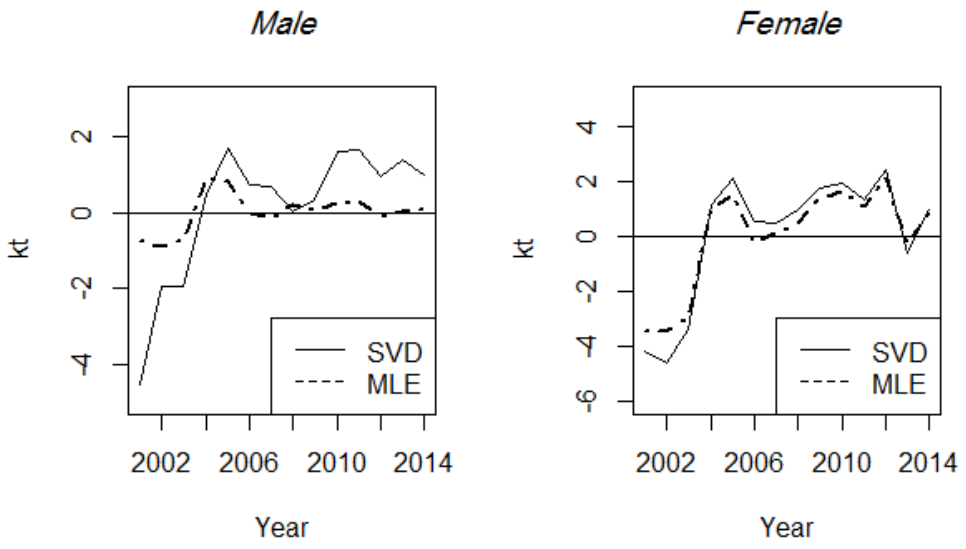
nd is slight difference for female , and this is very clear in the figure (4.19) for both sex (male, female).The maximum difference value of estimation of b_x for male is 2.077369 in age-group (5-9), while for female is 0.18866046 in age-group (5-9).

Table (4.23) Comparison between SVD and MLE for estimation of k_t for lung cancer.

Year	Male		Female	
	SVD	MLE	SVD	MLE
2001	-4.52625305	-0.78370685	-4.1957395	-3.4550937
2002	-1.93422071	-0.88547995	-4.5891530	-3.4175455
2003	-1.93393405	-0.72001820	-3.3751181	-2.9335145
2004	0.46087834	0.89754008	1.1693789	1.0571542
2005	1.71249586	0.83409272	2.1449693	1.5076962
2006	0.74305146	-0.01138519	0.5684068	-0.2544051
2007	0.70544229	-0.13434391	0.4722695	0.1393346
2008	0.01372249	0.20286007	0.9711164	0.4870224
2009	0.34186040	0.05924104	1.7808479	1.3801507
2010	1.61211845	0.24886670	1.9337381	1.6361492
2011	1.66845237	0.27205262	1.3501663	1.0673409
2012	0.95335131	-0.08322350	2.4316543	2.1486644
2013	1.41541856	0.02935940	-0.6172925	-0.2163707
2014	1.01183937	0.07414499	0.9801094	0.8534167

Source: Author calculation ilc and demography and R.

Figure(4.20) Comparison between SVD and MLE for estimation of k_t 2001-2014 for lung cancer.



Source: Author plotted by ilc and demography and R.

If we take a look to table (4.23) and figure (4.20), we will notice that the e stimation of parameter k_t from SVD and MLE is a difference for male, a nd female, and this is very clear in the figure (4.20) for both sex (male, fe male). The maximum difference value of estimation k_t is 3.742546 in ye ar=2001 for male, while for female is 1.171608 in year 2002.

Table(4.24) Comparison between SVD and MLE for Errors for lung cancer.

Sex	Method	ME	MSE
Male	**SVD**	0.01145	0.46193
	MLE	0.00714	0.12385
Female	**SVD**	0.00523	0.08022
	MLE	0.02407	0.06613

Source: Author calculation ilc and demography and R.

Table (4.24) shows the errors from the two methods (SVD, MLE) to estimate the parameters, and they are satisfactory well, but MLE is better for male with errors (ME=0.00714,MSE=0.12385). While for female the SVD is the better with errors (ME=0.00523, MSE=0.08022).

4.3.2.4 Forecast $\mathbf{k_t}$ and Age-specific Cancer Mortality Rate:

After obtained mortality index k_t from MLE for male and SVD for femal e. We obtained forecast the mortality index from equations(2.31) , drift ($\hat{d}$) from equation(2.32), standard error ($\widehat{s.e}$) from equation(2.54) and erro r ($\hat{\sigma}^2$) from equation(2.37) .

Table (4.25) Estimation of drift, standard error and errors of RWD (0,1,0) for lung cancer.

Sex	Male	Female
Method	**MLE**	**SVD**
$\widehat{\boldsymbol{d}}$	0.0660	0.3981
$\widehat{\boldsymbol{s.e}}$	0.1468	0.4739
$\widehat{\boldsymbol{\sigma}}^{\boldsymbol{2}}$	0.3036	3.163

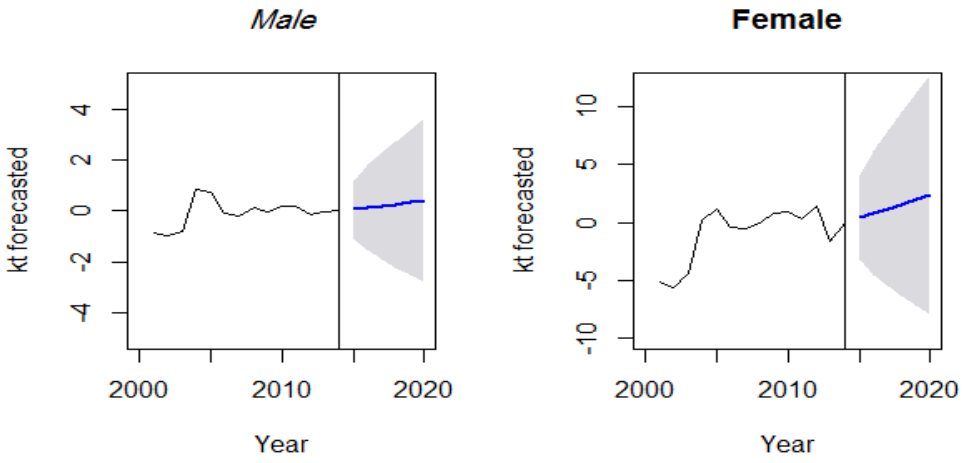
Source: Author calculation by ilc and demography, forecast and R.

Table (4.26) Forecast Mortality index for period 2015–2020 for lung cancer.

Sex	Male			Female		
	MLE			SVD		
Year	$k_t$forecast	lower	Upper	$k_t$forecast	lower	Upper
2015	0.0659886	-1.054736	1.186713	0.7362127	-8.318554	0.7432376
2016	0.1319772	-1.508596	1.772550	1.4724254	-0.363421	2.4518879
2017	0.1979658	-1.877213	2.273145	2.2086381	-1.971617	3.7238672
2018	0.2639544	-2.206003	2.733912	2.9448508	-3.353775	4.7698089
2019	0.3299430	-2.511613	3.171499	3.6810635	-4.591483	5.6713001
2020	0.3959316	-2.802134	3.593998	4.4172762	-5.726583	6.4701837

Source: Author calculation by ilc and demography, forecast and R.

Figure(4.21) Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for lung cancer.



Source: Author plotted by ilc and demography and R.

Tables (4.25) and (4.26) show the values of drift, standard error, errors and k_t for male and female, and figure (4.21) shows the trend of mortality index. It is increasing overtime for both sex (male-female). The age-specific cancer mortality rate, $m_{x,t}$ is now forecasting for years 2015-2020 from the equation(2.30).

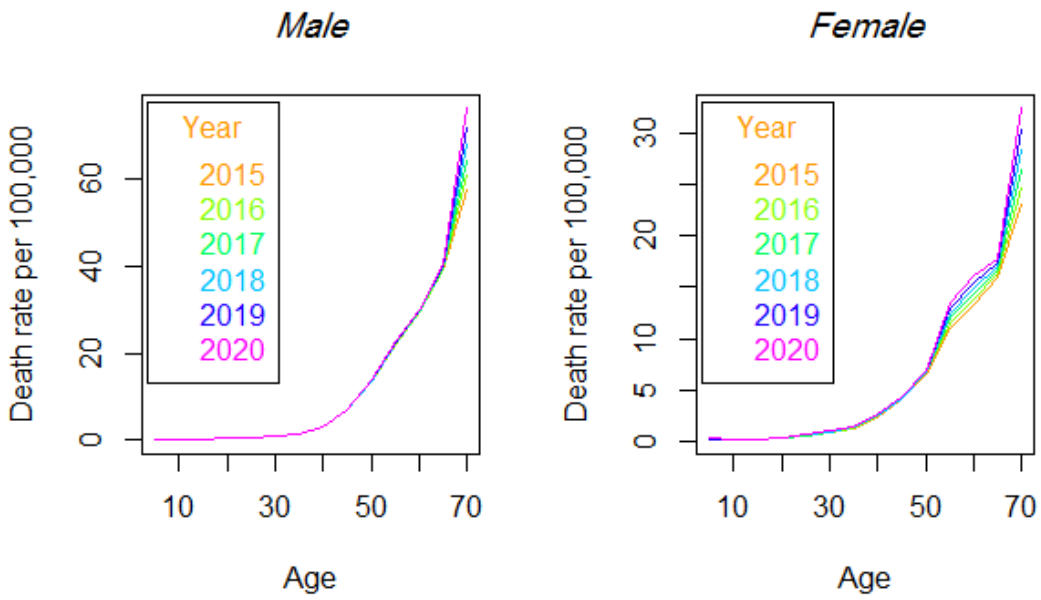
Table (4.27) Forecast age-specific mortality rate for lung cancer per 100.000 for period 2015 – 2020.

Sex	Age	2015	2016	2017	2018	2019	2020
Male	5-9	0.07	0.07	0.07	0.06	0.06	0.06
	10-14	0.13	0.13	0.13	0.13	0.13	0.13

	15-19	0.29	0.29	0.29	0.29	0.29	0.29
	20-24	0.57	0.58	0.58	0.59	0.59	0.60
	25-29	0.71	0.71	0.71	0.72	0.72	0.72
	30-34	1.07	1.08	1.09	1.10	1.11	1.12
	35-39	1.65	1.65	1.65	1.65	1.65	1.65
	40-44	3.33	3.33	3.33	3.33	3.32	3.32
	45-49	6.82	6.87	6.92	6.97	7.02	7.07
	50-54	13.46	13.56	13.65	13.75	13.84	13.94
	55-59	21.98	22.10	22.23	22.36	22.49	22.62
	60-64	29.53	29.61	29.69	29.77	29.85	29.93
	65-69	39.14	39.36	39.59	39.82	40.05	40.28
	70-74	57.42	60.77	64.32	68.08	72.06	76.27
Female	5-9	0.04	0.04	0.04	0.04	0.04	0.04
	10-14	0.12	0.13	0.13	0.14	0.14	0.15
	15-19	0.10	0.11	0.11	0.11	0.11	0.12
	20-24	0.01	0.01	0.01	0.01	0.01	0.01
	25-29	0.04	0.04	0.04	0.03	0.03	0.03
	30-34	0.04	0.04	0.04	0.04	0.03	0.03
	35-39	0.23	0.23	0.23	0.22	0.22	0.22
	40-44	0.44	0.44	0.44	0.44	0.44	0.44
	45-49	0.64	0.64	0.64	0.63	0.63	0.63
	50-54	0.71	0.70	0.69	0.68	0.67	0.65
	55-59	1.61	1.61	1.61	1.61	1.61	1.61
	60-64	1.09	1.07	1.05	1.04	1.02	1.00
	65-69	3.35	3.41	3.47	3.53	3.60	3.66
	70-74	2.87	2.88	2.88	2.89	2.89	2.90

Source: Author calculation by ilc and demography, forecast and R.

Figure(4.22) Forecast age-specific mortality rate 2015-2020 for lung cancer.



Source: Author plotted by ilc and demography and R.

Table(4.27) and figure(4.22) show the age-specific cancer mortality rates are increasing for age group (40-44) to (70-74) year-old for male, while for female age-specific cancer mortality rates are increasing for aged group (30-39) to (70-74) year-old. When comparing both sex, the male have higher cancer mortality rate than female overtime .

Table (4.28) Model's forecast errors based on mortality rate across ages f or lung cancer .

Sex	Method	MPE
Male	**MLE**	0.07195
Female	**SVD**	0.04939

Source: Author calculation ilc and demography, forecast and R .

The table (4.28) shows the errors from the forecasting age-specific cancer mortality rate for both sex (male, female), and they are satisfactory well for both sex (male, female).

4.2.3 Colon Cancer:

4.2.3.1 The Singular Value Decomposition (SVD):

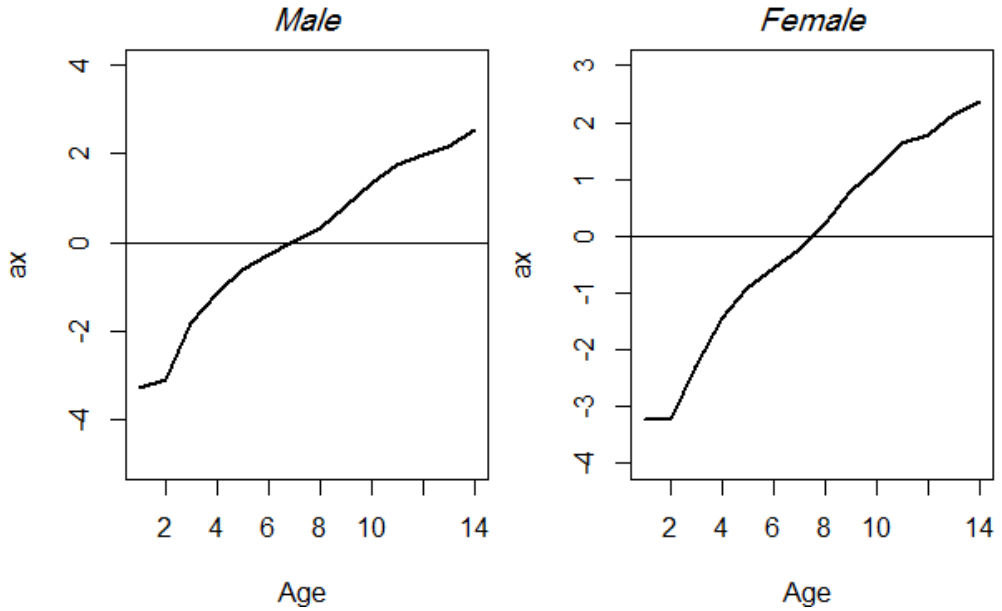
According to LC model we obtained the parameter a_x first from equation (2.13). We had the following table and figure as a result.

Table (4.29) Estimation of a_x by SVD for colon cancer.

Age	Male	Female
5-9	-3.53290945	-3.2355569
10-14	-3.10000329	-3.2325193
15-19	-1.81588120	-2.2813181
20-24	-0.98485968	-1.4495881
25-29	-0.61660929	-0.9240679
30-34	-0.28417757	-0.5687064
35-39	0.01832391	-0.2438334
40-44	0.33843088	0.2126980
45-49	0.81913661	0.7942193
50-54	1.34541791	1.2034246
55-59	1.75761235	1.6409298
60-64	1.98511767	1.7947444
65-69	2.16650715	2.1408554
70-74	2.53548953	2.3805798

Source: Author calculation ilc and demography and R.

Figure (4.23) General pattern of mortality *ax* by SVD for colon cancer .



Source: Author plotted by ilc and demography and R.

Table (4.29) shows the values of a_x, which represents the general pattern (age shape) of mortality by age x for both sex (male-female), and it is increasing overtime for both sex (male, female). Figure (4.23) shows the pattern of a_x and it has up trend for both sex (male, female), and this indicates that the younger ages have lower mortality rate than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate.

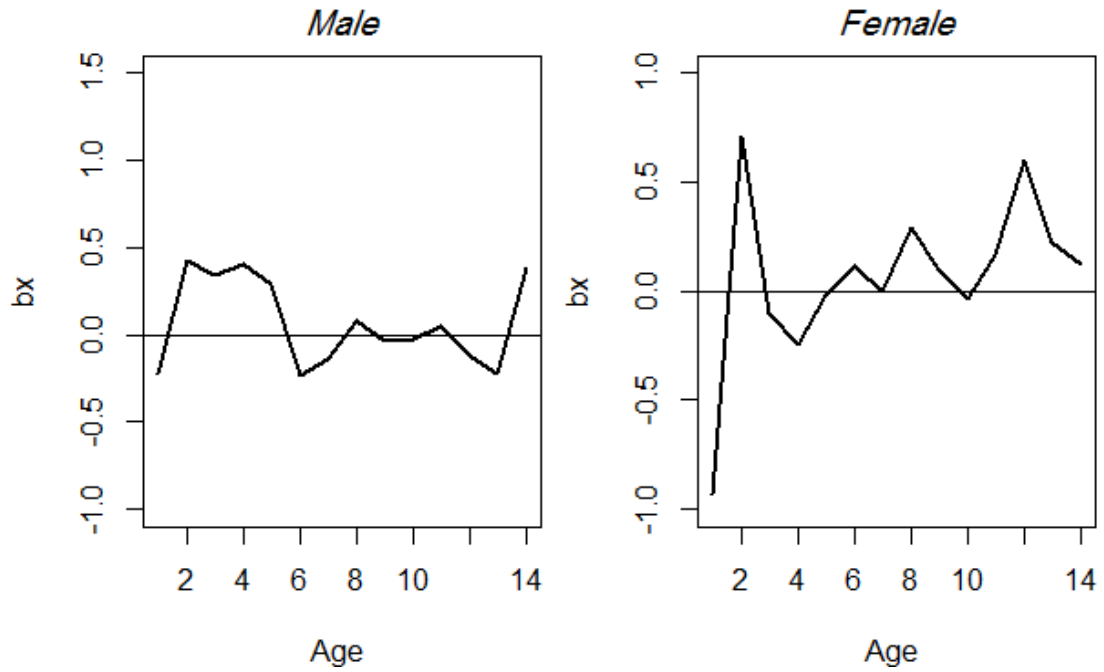
The second step is estimated the parameter b_x from the equation(2.17).

Table(4.30) Estimation of b_x by SVD for colon cancer .

Age	Male	Female
5-9	-0.659039790	-0.93021179
10-14	0.227754073	0.71187756
15-19	0.295034154	-0.09940891
20-24	1.122244900	-0.24818266
25-29	0.363905258	-0.01556198
30-34	0.003015176	0.11288387
35-39	-0.088629174	-0.00399170
40-44	0.068362719	0.29574963
45-49	-0.009866323	0.09731455
50-54	0.029212021	-0.03653695
55-59	-0.064217863	0.16559789
60-64	-0.238798260	0.60059430
65-69	-0.156492757 0.1075158	0.22492371
70-74	66	0.12495250

Source: Author calculation ilc and demography and R.

Figure(4.24) General pattern of b_x by SVD for colon cancer .



Source: Author plotted by ilc and demography and R.

Table (4.30) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality changes. The figure (4.24) shows the cancer mortality change for younger ages for male, and the cancer mortality among younger ages have highest values. While for female the mortality for younger ages have highest values. The high values of b_x indicate improvement in mortality at these ages, while the negative values at some ages indicate that mortality rate is increasing.

The parameter k_t first estimated from the equation(2.18) and re-estimated of k_t from equation(2.19).

Table(4.31) First and second estimation of k_t by SVD for colon cancer .

Year	Male		Female	
	1st estimation	2nd estimation	1st estimation	2nd estimation
2001	1.146942967	0.74160675	-0.817822591	-2.14419286
2002	-0.406075192	-0.14210725	-0.260261369	-1.38491806
2003	-0.325588844	-0.14228783	0.116706072	-0.68547055
2004	0.798716265	0.95131117	-0.164051532	0.16879074
2005	0.870350407	1.74953844	-0.793184651	0.43941683
2006	0.339109120	0.41057080	-0.002809343	-0.11913314
2007	0.004599658	1.06728620	-0.075122539	0.36403319
2008	-0.134249018	-0.09566960	0.022888123	-0.09131581
2009	-0.182098401	-0.09645981	-0.346507254 0.5	-0.55406383
2010	-0.032816015	0.04873011	53070949 0.7588	0.04140939
2011	-0.411749723	1.46094517	12619 0.3036730	0.96866951
2012	-0.580389202	1.40399144	51 0.186170334	1.12355344
2013	-0.462729122	1.26948739	0.518438131	0.30366522
2014	-0.624022900	1.91101848		0.78413057

Source: Author calculation ilc and demography, forecast and R.

Figure(4.25) General pattern of *kt* 2001–2014 by SVD for colon cancer .

Source: Author plotted by ilc and demography and R.

Table (4.31) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female), which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure (4.25) shows the

mortality index k_t has non-linear trend overtime for male and female. The low values of k_t indicate the mortality trend is decline**.**

4.2.3.2 Maximum Likelihood Estimation (MLE):

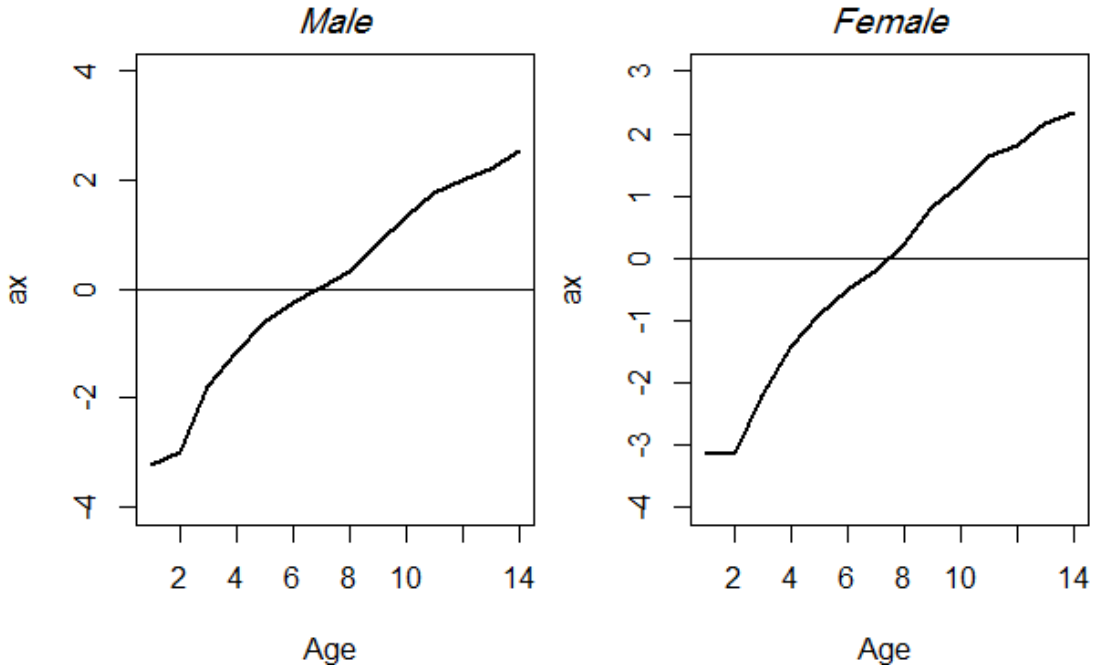
After fitting the technique of (MLE), we obtained at the following results.

Table (4.32) Estimation of a_x by MLE for colon cancer .

Age	Male	Female
5-9	-3.66637596	-3.1131758
10-14	-2.99544866	-3.1525122
15-19	-1.76819721	-2.1945513
20-24	-1.09702779	-1.4138014
25-29	-0.61624948	-0.9045201
30-34	-0.25596000	-0.5164398
35-39	0.02772012	-0.2234811
40-44	0.34389374	0.2224095
45-49	0.83107735	0.8053181
50-54	1.34510871	1.2028160
55-59	1.77697451	1.6388940
60-64	2.02234346	1.8011937
65-69	2.21900972	2.1573556
70-74	2.53552174	2.3364841

Source: Author calculation ilc and demography and R.

Figure(4.26) General pattern of ax by MLE for colon cancer .



Source: Author plotted by ilc and demography and R.

Table(4.32) shows the values of a_x ,which represents the general pattern (age shape) of mortality by age for both sex (male-female), and Figure(4.26) shows the pattern of a_x and shows the values of $a_x$is increasing over time for both sex (male, female), and this indicates that they have up trend in mortality and the younger ages have lower mortality rate than older ages. The negative trend in a_x is in accord with improvement in cancer mortality rate .

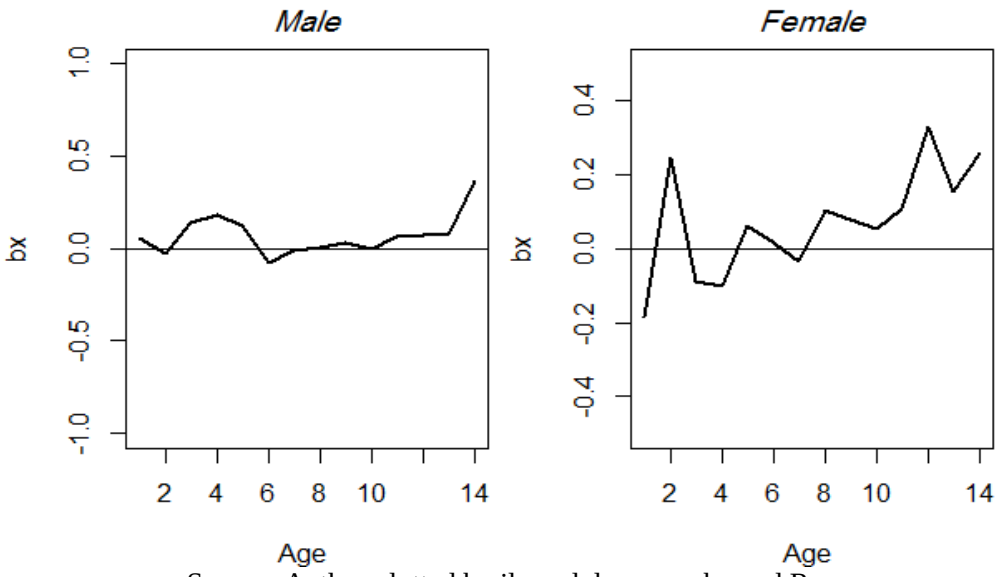
The second step is estimated the parameter b_x from the equation(2.30).

Table (4.33) Estimation of b_x by MLE for colon cancer .

Age	Male	Female
5-9	-0.821667657	-0.18479741
10-14	-0.051230827	0.24778118
15-19	0.305852583	-0.08815489
20-24	0.247693920	-0.10261376
25-29	0.278103135	0.06052411
30-34	-0.177637626	0.01467529
35-39	-0.031097756	-0.03567370
40-44	0.007978901	0.10526351
45-49	0.068626016	0.07931886
50-54	-0.007935564	0.05453435
55-59	0.131536718	0.10732095
60-64	0.139910407	0.32818242
65-69	0.136732105	0.15384091
70-74	0.773135645	0.25979818

Source: Author calculation ilc and demography and R.

Figure(4.27) General pattern of b_x by MLE for colon cancer .



Source: Author plotted by ilc and demography and R.

Table(4.33) shows the values of b_x which represents the tendency of mortality at age x to change as the general level of mortality changes. the figure (4.27) shows the b_x has a negative values for younger ages for male, and positive value for older ages while the middle ages have invariant values, for female the values of b_x is no invariant, and the younger and the older ages have highest values . The high values of b_x indicate improvement in mortality at all ages , while the negative values at some ages indicate that mortality is increasing.

The parameter k_t estimated from the equation(2.30).

Table(4.34) Estimation of kt by MLE for colon cancer .

Year	Male	Female
2001	-0.928203921	-2.04627162
2002	-1.027160713	-1.45314184
2003	-0.594061609	-0.99320113
2004	0.654405599	-0.16336410
2005	1.161618504	0.09847927
2006	-0.041611141	-0.19240538
2007	0.039622075	0.42254482
2008	-0.218416341	0.36222290
2009	0.007840427	-0.55027730
2010	0.172756241	0.62763718
2011	0.134331804	1.20181095
2012	0.217326561	1.39765791
2013	0.012394591	0.27970006
2014	0.409157921	1.00860828

Source: Author calculation ilc and demography, forecast and R.

Figure(4.28) General pattern of k_t for 2001-2014 by MLE for colon cancer .

Source: Author plotted by ilc and demography and R.

Table(4.34) shows the values of mortality index k_t for the period 2001–2014 for both sex (male-female) ,which it captures the main time trend on the logarithmic scale in death rates at all ages. Figure(4.28) shows the mortality index k_t has nonlinear trend for male and female. The high values of k_t indicate there is no improvement of cancer mortality rate. for male have highest mortality rate in year2005, while for female have highest mortality in 2010 to 2014.

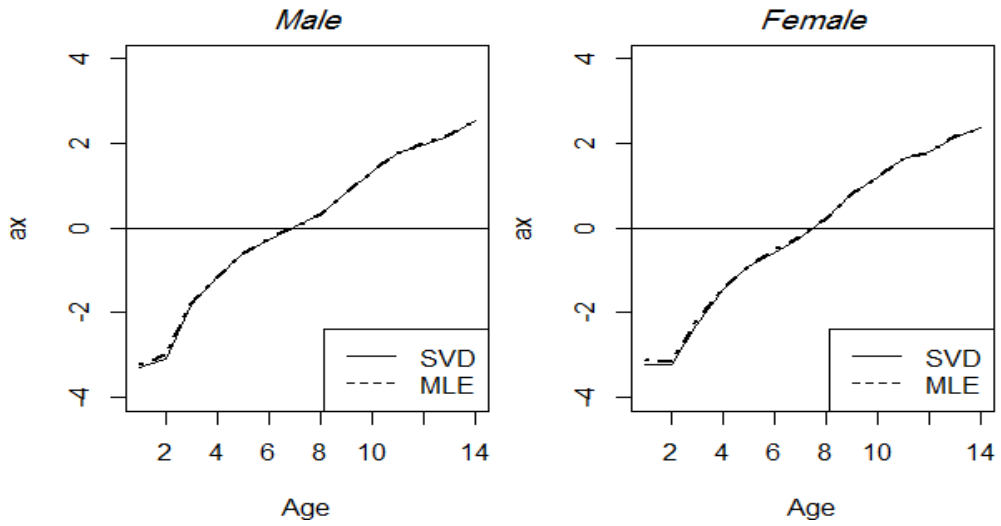
4.2.3.3 Comparison between SVD and MLE:

Table (4.35) Comparison between SVD and MLE for estimation a_x for colon cancer .

Age	Male		Female	
	SVD	MLE	SVD	MLE
5-9	-3.53290945	-3.66637596	-3.2355569	-3.1131758
10-14	-3.10000329	-2.99544866	-3.2325193	-3.1525122
15-19	-1.81588120	-1.76819721	-2.2813181	-2.1945513
20-24	-0.98485968	-1.09702779	-1.4495881	-1.4138014
25-29	-0.61660929	-0.61624948	-0.9240679	-0.9045201
30-34	-0.28417757	-0.25596000	-0.5687064	-0.5164398
35-39	0.01832391	0.02772012	-0.2438334	-0.2234811
40-44	0.33843088	0.34389374	0.2126980	0.2224095
45-49	0.81913661	0.83107735	0.7942193	0.8053181
50-54	1.34541791	1.34510871	1.2034246	1.2028160
55-59	1.75761235	1.77697451	1.6409298	1.6388940
60-64	1.98511767	2.02234346	1.7947444	1.8011937
65-69	2.16650715	2.21900972	2.1408554	2.1573556
70-74	2.53548953	2.53552174	2.3805798	2.3364841

Source: Author calculation ilc and demography, forecast and R.

Figure(4.29) Comparison between SVD and MLE for estimation a_x for colon cancer .



Source: Author plotted by ilc and demography and R.

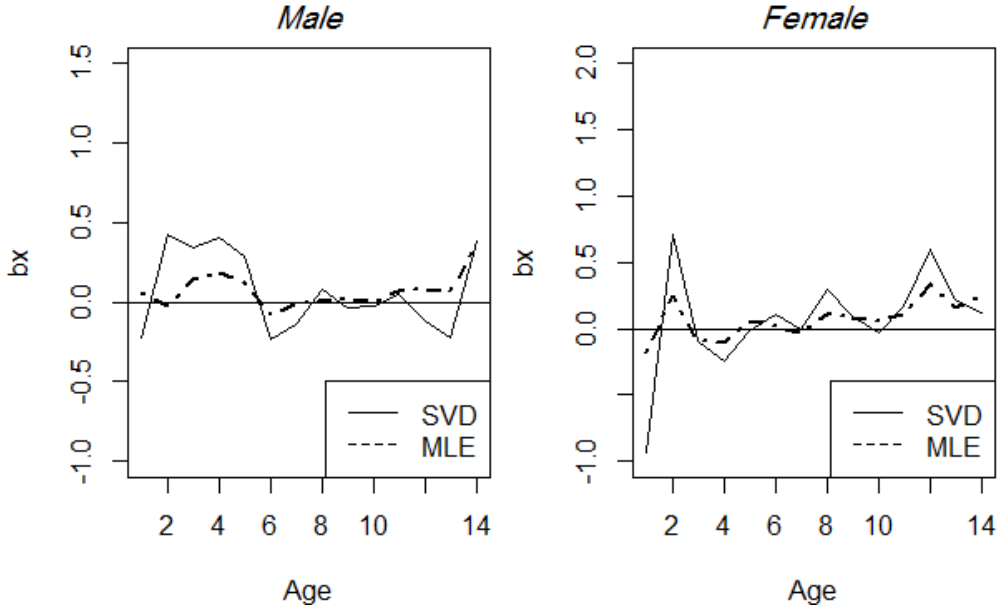
If we take a look to table(4.35) and figure(4.29), we will notice that the es timation of parameter a_x from SVD and MLE is slight difference , and thi s is very clear in the figure (4.29) for both sex (male, female). The maxim um difference value of estimation a_x for male is 0.1334665in age-group (5-9), while for female is 0.1223812 in age-group (5-9).

Table (4.36) Comparison between SVD and MLE for estimation of b_x for colon cancer .

Age	Male		Female	
	SVD	**MLE**	**SVD**	**MLE**
5-9	-0.659039790	-0.821667657	-0.93021179	-0.18479741
10-14	0.227754073	-0.051230827	0.71187756	0.24778118
15-19	0.295034154	0.305852583	-0.09940891	-0.08815489
20-24	1.122244900	0.247693920	-0.24818266	-0.10261376
25-29	0.363905258	0.278103135	-0.01556198	0.06052411
30-34	0.003015176	-0.177637626	0.11288387	0.01467529
35-39	-0.088629174	-0.031097756	-0.00399170	-0.03567370
40-44	0.068362719	0.007978901	0.29574963	0.10526351
45-49	-0.009866323	0.068626016	0.09731455	0.07931886
50-54	0.029212021	-0.007935564	-0.03653695	0.05453435
55-59	-0.064217863	0.131536718	0.16559789	0.10732095
60-64	-0.238798260	0.139910407	0.60059430	0.32818242
65-69	-0.156492757	0.136732105	0.22492371	0.15384091
70-74	0.107515866	0.773135645	0.12495250	0.25979818

Source: Author calculation ilc and demography, forecast and R.

Figure(4.30) Comparison between SVD and MLE for estimation b_x for colon cancer .



Source: Author plotted by ilc and demography and R.

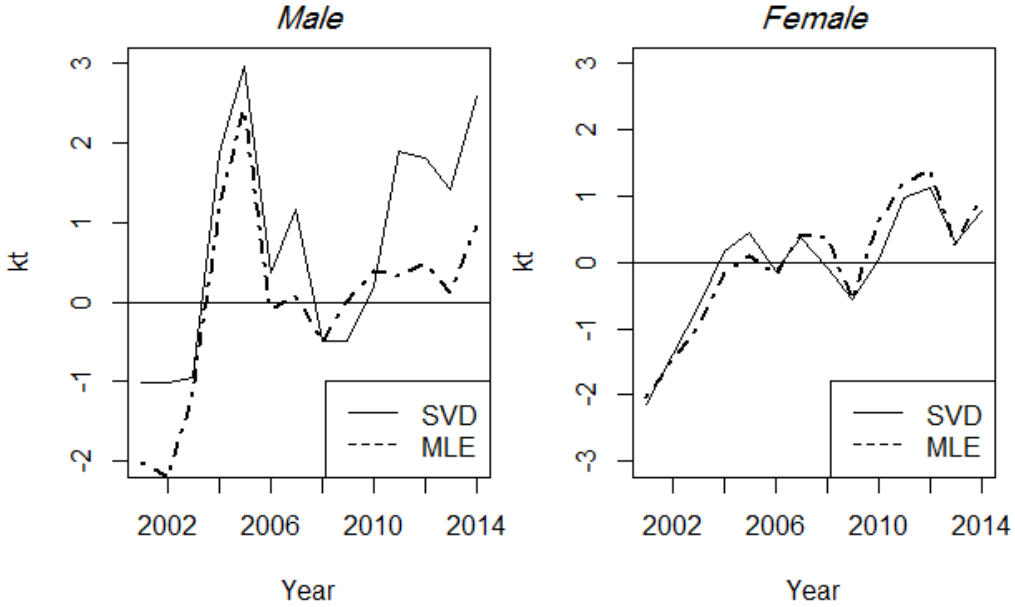
If we take a look to table(4.36) and figure(4.30), we will notice that the es timation of parameter b_x from SVD and MLE is a difference for male and female , and this is very clear in the figure (4.30) for both sex (male, fema le). The maximum difference value of estimation b_x for male is 0.874551 in age-group (20-24), while for female is 0.74541445 in age-group (5-9).

Table(4.37) Comparison between SVD and MLE for estimation of k_t for colon cancer .

Year	Male		Female	
	SVD	**MLE**	**SVD**	**MLE**
2001	0.74160675	-0.928203921	-2.14419286	-2.04627162
2002	-0.14210725	-1.027160713	-1.38491806	-1.45314184
2003	-0.14228783	-0.594061609	-0.68547055	-0.99320113
2004	0.95131117	0.654405599	0.16879074	-0.16336410
2005	1.74953844	1.161618504	0.43941683	0.09847927
2006	0.41057080	-0.041611141	-0.11913314	-0.19240538
2007	1.06728620	0.039622075	0.36403319	0.42254482
2008	-0.09566960	-0.218416341	-0.09131581	0.36222290
2009	-0.09645981	0.007840427	-0.55406383	-0.55027730
2010	0.04873011	0.172756241	0.04140939	0.62763718
2011	1.46094517	0.134331804	0.96866951	1.20181095
2012	1.40399144	0.217326561	1.12355344	1.39765791
2013	1.26948739	0.012394591	0.30366522	0.27970006
2014	1.91101848	0.409157921	0.78413057	1.00860828

Source: Author calculation ilc and demography and R.

Figure(4.31) Comparison between SVD and MLE for estimation k_t 2001-2014 for colon cancer .



Source: Author plotted by ilc and demography and R.

If we take a look to table(4.37) and figure(4.31), we will notice that the es timation of parameter k_t. from SVD and MLE is a difference for male, w hile for female is slight difference, and this is very clear in the figure (4.3 1) for both sex (male, female). The maximum difference value of estimati on k_t for male is 1.669811 in year 2001, while for female is 0.5862278 in year 2010.

Table(4.38) Comparison between SVD and MLE for Errors for colon cancer .

Sex	Method	ME	MSE
Male	**SVD**	0.05377	0.28382
	MLE	0.00506	0.11065
Female	**SVD**	-0.00401	0.13561
	MLE	0.02958	0.07824

Source: Author calculation ilc and demography and R.

Table(4.38) shows the errors from the two methods (SVD,MLE), and they are satisfactory well for estimating the parameters, but MLE is better than SVD for male with errors (ME=0.00506, MSE=0.11065). While for female the SVD is better than MLE with errors (ME=-0.00401, MSE=0.13561) .

4.2.3.4 Forecast k_t and Age-specific Cancer Mortality Rate:

After obtained mortality index k_t from MLE for male and SVD for female. We obtained forecast the mortality index from equations(2.31) , drift ($\hat{d}$) from equation(2.32), standard error ($\widehat{s.e}$) from equation(2.54) and error ($\hat{\sigma}^2$)from equation(2.37).

Table(3.39) Estimation of drift, standard error and errors of RWD (0,1,0) for colon cancer .

Sex	**Male**	**Female**
Method	**MLE**	**SVD**
$\widehat{d}$	0.1029	0.2253
$\widehat{s.e}$	0.1476	0.1598
$\widehat{\sigma}^2$	0.3069	0.3598

Source: Author calculation by ilc and demography, forecast and R.

Table(4.40) Forecast mortality index for period 2015 – 2020 for colon cancer**.**

Sex	**Male**			**Female**		
Method	**MLE**			**SVD**		
Year	k_t**forecast**	**lower**	**upper**	k_t**forecast**	**lower**	**Upper**
2015	0.1028740	-1.023911	1.229659	0.2252556	-0.9947564	1.445268
2016	0.2057480	-1.443697	1.855193	0.4505113	-1.3354035	2.236426
2017	0.3086220	-1.777779	2.395023	0.6757669	-1.5832564	2.934790
2018	0.4114960	-2.071818	2.894810	0.9010226	-1.7877530	3.589798
2019	0.5143699	-2.342552	3.371292	1.1262782	-1.9670169	4.219573
2020	0.6172439	-2.598116	3.832604	1.3515339	-2.1298549	4.832923

Source: Author calculation by ilc and demography, forecast and R.

Figure (4.32) Fitted and forecasted mortality index with 95% prediction line from 2001-2020 for colon cancer.

Source: Author plotted by ilc and demography and R.

Table (4.39) and (4.40) show the values of drift, standard error, errors and k_t and figure (4.32) shows that the k_t increases overtime for both sex (male, female).

The age-specific cancer mortality rate, $m_{x,t}$ is now forecasting for years 2015-2020 from the equation(2.30).

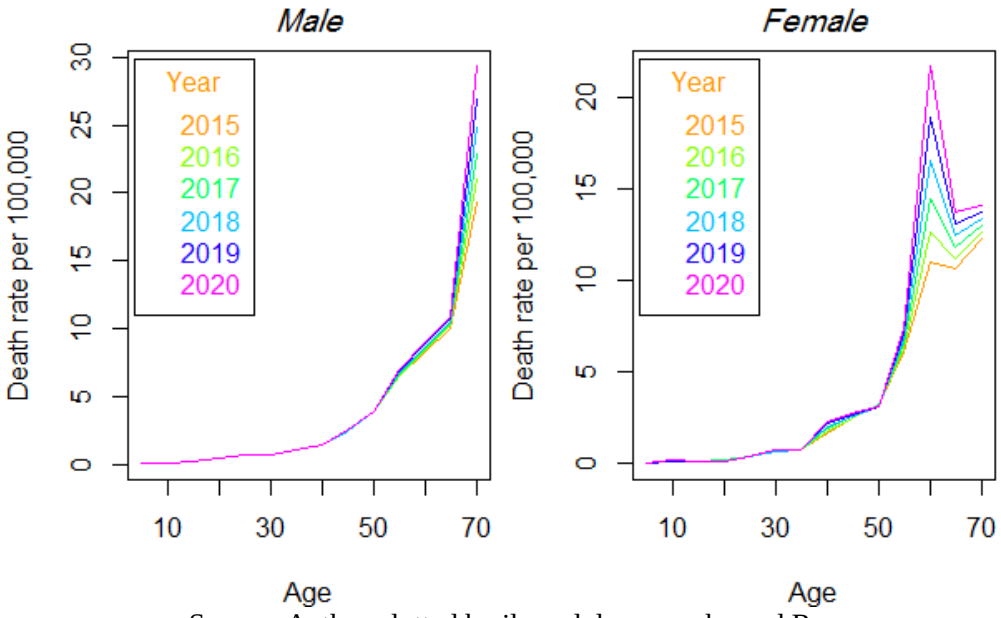
Table (4.41) Forecast age-specific cancer mortality rate for colon cancer per 100,000 for the period 2015–2020.

Sex	Age	2015	2016	2017	2018	2019	2020
Male	5-9	0.02	0.02	0.01	0.01	0.01	0.01
	10-14	0.05	0.05	0.05	0.05	0.05	0.05
	15-19	0.20	0.21	0.21	0.22	0.23	0.23
	20-24	0.38	0.39	0.40	0.41	0.42	0.43
	25-29	0.62	0.64	0.66	0.68	0.70	0.72
	30-34	0.71	0.69	0.68	0.67	0.66	0.65
	35-39	1.01	1.01	1.01	1.00	1.00	1.00
	40-44	1.42	1.42	1.42	1.42	1.42	1.422
	45-49	2.38	2.39	2.41	2.43	2.45	2.46
	50-54	3.82	3.82	3.82	3.81	3.81	3.81

	55-59	6.32	6.41	6.50	6.59	6.68	6.77
	60-64	8.12	8.23	8.35	8.47	8.60	8.72 10
	65-69	9.87	10.01	10.15	10.29	10.44	.58
	70-74	18.75	20.31	21.99	23.81	25.78	27.91
Female	5-9	0.02	0.01	0.01	0.01	0.01	0.01
	10-14	0.08	0.10	0.11	0.13	0.15	0.18
	15-19	0.09	0.09	0.09	0.09	0.08	0.08
	20-24	0.18	0.17	0.16	0.15	0.15	0.14
	25-29	0.39	0.39	0.39	0.39	0.39	0.38
	30-34	0.63	0.65	0.67	0.68	0.70	0.72
	35-39	0.78	0.78	0.78	0.78	0.78	0.78
	40-44	1.67	1.78	1.90	2.04	2.18	2.33
	45-49	2.44	2.50	2.55	2.61	2.66	2.72
	50-54	3.21	3.18	3.16	3.13	3.11	3.08
	55-59	6.10	6.33	6.57	6.82	7.08	7.35
	60-64	11.03	12.63	14.46	16.56	18.96	21.70
	65-69	10.67	11.23	11.81	12.43	13.07	13.75
	70-74	12.26	12.61	12.97	13.35	13.73	14.12

Source: Author calculation by ilc and demography, forecasting and R.

Figure (4.33) Forecast age-specific mortality rate 2015-2020 for colon cancer.



Source: Author plotted by ilc and demography and R.

Tables(4.41) and figure(4.33) show the age-specific cancer mortality rates are increasing for all age-group over time for male and female .

Table(4.42) Model's forecast Model's forecast errors based on mortality rate across ages for colon cancer .

Sex	Method	MPE
Male	**MLE**	0.06156
Female	**SVD**	0.07052

Source: Author calculation by ilc and demography, forecast and R.

Table (4.42) shows the errors from forecasting age-specific cancer mortality rate for both sex (male, female) .

CHAPTER FIVE

(Conclusions and Recommendations)

5.1 Conclusions
5.2 Recommendations

Conclusions:

The results showed that :

1. The cancer age-specific mortality rate was increasing overtime and age-group for all cancer for both sex(male, female).
2. The cancer age-specific mortality rate of male was higher than female for all cancer overtime and age-groups.
3. The cancer age-specific mortality rate of lung had highest rate overtime and it was (57.42,57.42,64.23,68.08,72.06 and 76.27) than colon (12.91,14.30, 5.84, 17.55, 19.45 and 21.55) and after that oral (3.11,3.10,3.09,3.08,3.06 and 3.05).

- **Oral Cancer:**

1. The two methods (SVD, MLE) were satisfactory with errors (ME=0.00016, 0.03162) and (MSE=0.25208, 0.15594) respectively for males. while for females (ME=0.02856, 0.07680) , (MSE=0.32310,0.22285) respectively.
2. SVD was better than MLE for both sex (male, female), with error= (ME=0.00016, (MSE=0.35183) for male, while for female (ME=0.02856, MSE=0.32310).
3. The errors of forecasting age-specific cancer mortality rate across ages for male and female (MPE=0.14960,0.25731) respectively.
4. The highest age-specific cancer mortality rate for male found in age-group (70-74) in years (2015,2016,2017,2018,2019 and 2020) and it was (3.11,3.10,3.09,3.08,3.06 and 3.05) respectively.
5. The highest age-specific cancer mortality rate for female found in age-group (65-69) in years (2015, 2016, 2017, 2018, 2019 and 2020) and it was (3.35,3.41,3.47,,.53,3.60 and 3.66) respectively.

- **Lung cancer:**
  1. The two methods (SVD, MLE) were satisfactory with errors (ME=0.01145,0.00714) and (MSE=0.46193,0.12385) respectively for male , while for females (ME=0.00523, 0.02407) and (MSE=0.08022,0.06613) respectively.
  2. MLE is better than SVD for male with error=(ME=0.00714, MSE=0.12385), while for female SVD is the better than MLE with error (ME=0.00523,MSE=0.08022).
  3. The errors of forecasting age-specific cancer mortality rate for male. and female (MPE=0.07195,04939) respectively.
  4. The highest age-specific cancer mortality rate for male found in age-group (70-74) in years (2015, 2016, 2017, 2018, 2019 and 2020) and it was (57.42,57.42,64.23,68.08,72.06 and 76.27) respectively.
  5. The highest age-specific cancer mortality rate for female found in age-group (65-69) for all years (2015,2016,2017,2018,2019 and 2020) and it was (3.35,3.41,3.47,3.53,3.60 and 3.66) respectively.
- **colon cancer:**
  1. The two methods (SVD, MLE) were satisfactory with errors (ME=0.05377,0.00506) and (MSE=0.28382,0.11065) respectively for male, while for females (ME = -0.00401,0.02958) and (MSE=0.13561,0.07824) respectively.
  2. MLE was better than SVD for male with error=(ME=0.00506, MSE=0.11065), while for female SVD was better than MLE with error (ME=-0.00401,MSE=0.07824) respectively.
  3. The errors of forecasting age-specific cancer mortality rate across ages for male, and female (MPE=0.06156 ,0.07052) respectively**.**
  4. The highest age-specific cancer mortality rate for male found in age-group (70-74) in years (2015,2016,2017,2018,2019 and 2020)

and it was (18.75,20.31,21.99,23.81,25.78 and 27.91) respectively.

5. The highest age-specific cancer mortality rate for female found in age-group (70-64) in years (2015, 2016, 2017, 2018, 2019 and 2020) and it was (12.26,12.61, 12.79, 13.35, 13.73 and 14.12) respectively.

Recommendations:

Recommended to

1. Use statistical modeling to forecast mortality rate to improve the understanding the cancer mortality rate impact of life
2. Use Lee-Carter model to forecast mortality rate ,because it is simple and also provides a description of mortality change that is easy to understand. The model has a few variables and combines demographic and statistical models other than the mortality models.
3. Apply SVD to estimate the model's parameters, because it has small error.
4. Mortality data are some of the best sources of information about the health of living communities. So we suggest to register information in a correct way because incomplete data affect the performance of the model to forecast.
5. ***In the last twenty years*** discovering modern methods or technical developments ***in statistics.*** These discovers and develops in statistical have a direct impact on the content that should be taught in our universities as change course content and structure, in both introductory and advanced courses for statisticians student and those from other disciplines.

6. The cancer mortality rate is increasing due time, so health sector must make plans and programs to reduce the mortality rate especially for lung cancer for male.

References:

1. MedicineNet.com.

2. Benjamin Gompertz. (2010). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*. **115**: 513-583 **.**

3. Lee, Ronald D. and Carter, Lawrence R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*. **87**: 659-671.

4. James Korley Attuquaye. (2015). The impact of varying time scales on the quality of cancer projections based on the Bayesian age-period-cohort model . Norwegian University of Science and Technology. Norway.

5. Jaione Etxeberria , Mar´ıa Dolores Ugarte , Tom´as Goicoa and Ana F. Militino. (2015) .*Statistical Journal* .**13**: 21–40.

6. Zoltan Butt, Steven Haberman and Han Lin Shang. (2015). Lee-Carter Mortality Models using Iterative Fitting Algorithms.

7. World Health Organization Mortality database. http://www.who.int/healthinfo/statistics/mortality_rawdata/en.(downloaded 5/11/2017) .

8. John R. Wilmoth, (1993), Computational Method of Fitting and Extrapolating the Lee-Carter model of mortality Change. Department of Demography, University of California, Berkeley. Technical Report

9. Ronald Lee, (2000), The Lee-Carter Method For Forecasting Mortality, With Various Extensions And Applications. North American Acturial Journal. (4,1): 80-91.

10. Lawrence R. Carter and Alexia Prskawetz**,** (2005), Examining Structural Shifts in Mortality Using the Lee-Carter Method. Max Planck Institute for Demographic.

11. Steven Haberman and Maria Russolillo, (2005), Lee-Carter mortality forecasting: application to the Italian population. Actuarial Research Paper No. 167.
12. Booth, Rob J. Hyndman, Leonie Tickle, Piet de Jong, (2006), Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions. Demographic Research. 15: 289-310.
13. . Claia Pedroza. (2006). A Bayesian forecasting model: predicting U.S. male mortality. Biostatistics (7,4): 530–550.
14. Jenny Zheng Wang, (2007), Fitting and Forecasting Mortality for Sweden: Applying the Lee-Carter Model. Dept. of Mathematical Statistics, Stockholm University.
15. Sándor Baran, József Gáll, Márton Ispány**,** Gyula Pap. (2007), Forecasting Hungarian mortality rates using the Lee-Carter method. Acta Oeconomica, (57,1):25–38.
16. Marie-Claire Koissi and Arnold F. Shapiro, (2008), The Lee-Carter Model Under The Condition Of Variables Age-Specific Parameters. 43rd Actuarial Research Conference, Regina, Canada.
17. Marie Claire Koissi, Arnold Shapiro , GÄoran HÄognÄ and Ronald Lee. (2008). Fitting and Forecasting Mortality Rates for Nordic Countries Using the Lee-Carter method.
18. Mariachiara Di Cesare and Mike Murphy, (2009), Forecasting Mortality, Different Approaches For Different Cause Of Deaths, The Cases Of Lung Cancer; Influenza, Pneumonia, Bronchitis; And Motor Vehicle Accidents, British Actuarial Journal British Actuarial Journal. 15:185-211.
19. Jackie Li, (2010) Projections of New Zealand Mortality Using the Lee-Carter Model and its Augmented Common Factor Extension, Population Association of New Zealand 36:27-53.

20. Angela U. Chukwu and E. O. Oladipupo, (November 2012) Modeling Adult Mortality in Nigeria. Studies in Mathematical Sciences. 5:1-12.
21. Wasana Aberathna ، Lakshman Alles ، W. N.Wickremasinghe and Isuru Hewapathirana, (2014), Modeling and Forecasting Mortality in Sri Lanka. Sri Lankan Journal of Applied Statistics. (15-3)141-170.
22. Farid Flici, (April 2015), Mortality forecasting for the Algerian population with considering cohort effect.
23. Wan Zakiyatussariroh Wan Husin, Mohammad Said Zainol and Norazan Mohamed Ramli, (2015) Performance of the Lee-Carter State Space Model in Forecasting Mortality.
24. Wouter van Wel (2015),Mortality Modeling and Forecasting using Cross-Validation Techniques. marble. (1.92).
25. Lucia Andreozzi, Maria Teresa Blacona ،Nora Arnesi, The Lee Carter Method For Estimating And Forecasting Mortality.
26. Rosella Giacomettia, Marida Bertocchib, Svetlozar T. Rachevc, Frank J. Fabozzid, A comparison of the Lee-Carter model and AR-ARCH model for forecasting mortality rates.
27. National Cancer Institute.
28. Johns Hopkins. (2006).Mortality and Morbidity Data Sources for measuring Mortality. Module 6a. School of public Health.
29. Heather Booth and Leonie Tickle. (2008). Mortality modelling and forecasting: a review of methods. ADSRI Working Paper No. 3The Australian Demographic & Social Research Institute. Australia.
30. Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Khalaf-Allah, M., 2011. Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics* **48**: 355–367.
31. Tabeau, E., Van Den Berg Jeths, A. and Heathcote, C. (2001). *Forecasting Mortality in Developed Countries: Insights from a Statistical*

Demographic and Epidemiological Perspective, Kluwer Academic Publishers.

32. Laurent Callot, Niels Haldrup and Malene Kallestrup Lamb.(2014). Deterministic and stochastic trends in the Lee-Carter mortality model. Research Paper **.**

33. Federico Girosi and Gary King. (2007). Understanding the Lee-Carter Mortality Forecasting Method.

34. Lee R. and Miller T. (2001). Evaluating the Performance of the Lee-Carter method for Forecasting Mortality. *Demography* **38**: 537–549.

35. Jack C. Yue, Sharon S. Yang and Hong-Chih Huang. (2008). A Study of the Lee-Carter Model with Age-Shifts.

36. A.E. Renshaw and S. Haberman. (2003). Lee-Carter mortality forecasting with age-specific enhancement. . *Insurance: Mathematics and Economics.* **33**: 255-272.

37. Arlt Josef ,Arltová Markéta and Klůfa Jindřich. (2011). The Diagnostic Checking of the Lee-Carter Mortality Forecasting Method. *Statistical Inst*. **58**:6139-6144.

38. Heather Booth, Leonie Tickle and Len Smith. (2002). Evaluation of the Variants of the Lee-Carter Method of Forecasting Mortality: A Multi-Country Comparison. .*researchgate.* **31(1)**:13-34.

39. Federico Girosi and Gary King. (2008) .*Demographic Forecasting*. Princeton University Press.

40 .Jolliffe I.T. (2002). *Principal component analysis***.** *2nd ed*. Springer-Verlag. New York.

41. Peter J. Brockwell and Richard A. Davis. (2002). *Introduction to time series and forecasting . 2nd ed* .Springer texts in statistics.

42. Guy Faguet. (2015). *The Conquest of Cancer, A distant goal* .Springer.

43. Salim EI, Jazieh AR, Moore MA.(2011). Lung cancer incidence in the arab league countries: risk factors and control. *Asian Pac J Cancer Prev.* **12(1):** 17-34.

44. World Health organization WHO, Fact sheet February 2017.

45. Abeer Al-Jaber, BDS, MPH, Lubna Al-Nasser, BDS, MPH and Ashraf El-Metwally, MD. (2016) . Epidemiology of oral cancer in Arab countries, *Saudi Medical Journal* . **37**: 249-255 .

46. American Cancer Society. (2010) .The global economic cost of cancer. **No.005444**.

47. American Cancer Society, (2015), Oral cancer, Inc. **N.300208-Rev.1/15**

48. Oral cavity and Orophayngeal Cancer, American Cancer Society

49.Cancer Quest, Emory Winship Cancer Institute .

50. Harlad S. Goodman and Alice M.Hrwitz, oral cancer p.235.

51. Understanding Lung Cancer. (2014). Cancer Council Australia.

52. Tina M. St. John, M.D. ,(2005). *with Every Breath: A Lung Cancer Guide Book*. The Lung Cancer Caring Ambassadors Program. www.lungcancercap.org.

53. Travis WD, Colby TV, Corrin B, et al. (1999). *Histologic typing of lung and pleural tumors*. *3rd Ed*. Springer Verlag. Berlin, Germany.

54. National Comprehensive Cancer Network Staff (NCCN). (2014). *Colon Cancer*.**1.2014**. National Comprehensive Cancer Network (NCCN).

55. National Cancer Institute. What You Need To Know About Cancer of the Colon and Rectum.

56. American Joint Committee on Cancer (AJSS). (2010). IIIinois. The original Source for this material is the AJCC Cancer staging Manual, *7th ed*. Springer Science + Business Media LLC, WWW.spinger.com.

57. Trsten Hothron and Brain S. Everrit. (2014). *A Handbook of statistical Analyses Using R*. 3ed.

Appendix

Appendix:

The numbers of cancer's death for both sex:

cancer		oral		lung		colon	
year	age	male	female	male	female	male	female
2001	5	5	1	2	1	0	3
2001	10	4	1	9	7	2	1
2001	15	4	5	13	6	6	6
2001	20	7	3	15	11	11	9
2001	25	3	3	21	10	15	7
2001	30	3	7	31	20	28	14
2001	35	7	7	32	17	21	17
2001	40	10	12	48	29	27	16
2001	45	11	14	94	47	33	32
2001	50	20	6	142	59	48	34
2001	55	17	10	147	46	37	25
2001	60	22	6	182	56	40	29
2001	65	12	12	195	65	51	31
2001	70	15	10	155	52	36	31
2002	5	2	1	4	3	2	0
2002	10	8	1	3	3	2	2
2002	15	3	2	9	13	6	2
2002	20	6	6	14	10	6	9
2002	25	4	3	13	10	15	9
2002	30	7	2	13	9	14	11
2002	35	8	5	38	22	23	14
2002	40	19	15	57	31	25	20
2002	45	21	8	98	40	46	33
2002	50	18	14	118	58	40	32
2002	55	16	15	145	43	38	42
2002	60	14	13	211	75	36	28
2002	65	17	10	179	48	32	38
2002	70	22	6	156	57	37	35
2003	5	5	1	2	5	2	1
2003	10	1	2	4	7	2	0
2003	15	5	4	16	5	6	8
2003	20	5	7	11	14	10	5
2003	25	5	4	10	9	9	14
2003	30	5	4	21	13	24	17
2003	35	12	9	45	20	27	20
2003	40	12	9	70	35	27	24

2003	45	15	12	99	53	33	29
2003	50	24	16	151	65	53	36
2003	55	17	13	188	76	52	36
2003	60	26	17	198	56	66	43
2003	65	24	13	211	51	51	41
2003	70	16	14	172	60	50	33
2004	5	2	3	4	6	3	2
2004	10	3	0	3	3	2	0
2004	15	2	3	15	7	9	2
2004	20	7	1	24	13	17	12
2004	25	8	5	15	10	17	8
2004	30	5	1	22	19	15	14
2004	35	15	9	45	22	25	25
2004	40	17	7	57	45	24	31
2004	45	14	15	120	63	38	29
2004	50	26	19	175	100	45	38
2004	55	28	17	177	87	54	42
2004	60	24	15	193	83	49	39
2004	65	21	18	249	58	54	41
2004	70	20	9	193	68	37	28
2005	5	2	4	4	3	0	5
2005	10	1	1	9	3	2	1
2005	15	10	10	11	14	13	7
2005	20	4	4	21	12	22	6
2005	25	4	4	20	14	28	17
2005	30	4	6	28	22	13	14
2005	35	7	9	34	26	19	15
2005	40	10	8	63	56	30	17
2005	45	17	12	122	65	51	30
2005	50	21	16	190	88	46	56
2005	55	30	12	211	107	56	50
2005	60	21	15	227	94	56	38
2005	65	19	14	218	81	41	44
2005	70	11	10	197	53	58	36
2006	5	1	0	5	4	1	1
2006	10	3	2	10	4	0	1
2006	15	7	2	12	8	13	4
2006	20	6	2	14	17	13	4
2006	25	5	3	21	13	17	15
2006	30	5	2	21	18	18	10
2006	35	9	4	23	27	23	24
2006	40	8	12	77	38	33	20
2006	45	15	12	118	70	33	30

2006	50	26	11	213	97	46	49
2006	55	27	16	213	96	73	43
2006	60	27	8	216	82	51	41
2006	65	23	17	255	90	44	39
2006	70	21	12	186	54	44	46
2007	5	5	4	6	11	1	3
2007	10	3	4	3	4	4	3
2007	15	5	4	9	7	8	2
2007	20	3	4	20	17	9	10
2007	25	3	5	25	22	13	13
2007	30	8	2	22	17	19	7
2007	35	9	7	29	27	23	19
2007	40	13	6	70	51	31	26
2007	45	7	12	115	64	34	40
2007	50	22	18	195	81	56	51
2007	55	34	19	255	92	60	50
2007	60	31	14	264	97	74	58
2007	65	15	14	225	69	53	33
2007	70	23	13	170	67	51	50
2008	5	2	3	10	1	0	0
2008	10	4	1	10	6	0	0
2008	15	7	1	11	7	5	5
2008	20	5	6	20	17	11	10
2008	25	5	4	22	16	15	12
2008	30	8	11	28	26	16	11
2008	35	4	8	40	34	22	14
2008	40	11	7	59	48	23	26
2008	45	17	13	141	83	38	37
2008	50	18	22	173	103	51	41
2008	55	29	21	250	113	66	48
2008	60	23	13	269	108	66	60
2008	65	19	5	245	87	47	60
2008	70	14	11	266	73	49	46
2009	5	1	2	7	9	1	0
2009	10	0	4	3	6	2	1
2009	15	7	2	12	8	7	0
2009	20	5	6	26	16	12	10
2009	25	5	7	23	17	10	11
2009	30	4	5	36	33	15	16
2009	35	4	0	33	30	23	21
2009	40	11	7	73	50	26	21
2009	45	15	16	147	83	53	36
2009	50	25	10	225	92	50	49

2009	55	23	21	284	128	67	59
2009	60	25	15	274	132	73	42
2009	65	24	11	225	99	46	38
2009	70	15	6	230	82	60	45
2010	5	3	5	6	5	0	1
2010	10	4	2	8	8	5	3
2010	15	4	2	18	12	6	5
2010	20	3	3	31	9	14	7
2010	25	4	5	23	21	20	9
2010	30	5	10	39	30	18	13
2010	35	4	5	60	38	22	15
2010	40	8	14	77	56	37	28
2010	45	21	10	143	70	31	34
2010	50	20	18	202	96	62	48
2010	55	29	15	314	127	77	68
2010	60	30	16	315	126	74	67
2010	65	36	15	243	121	60	44
2010	70	19	20	276	95	67	58
2011	5	1	1	6	5	0	1
2011	10	3	3	5	7	1	4
2011	15	2	1	12	11	5	0
2011	20	13	5	28	13	12	11
2011	25	8	3	25	26	21	14
2011	30	7	7	35	15	23	21
2011	35	11	8	33	38	25	20
2011	40	16	3	78	44	27	33
2011	45	15	11	148	69	48	50
2011	50	27	22	217	112	62	60
2011	55	30	10	303	136	113	72
2011	60	30	29	318	120	99	87
2011	65	23	14	308	110	70	60
2011	70	30	17	284	89	55	50
2012	5	2	2	5	6	2	1
2012	10	1	3	2	6	1	1
2012	15	5	6	21	8	13	5
2012	20	4	1	17	16	9	9
2012	25	9	1	23	24	14	19
2012	30	9	3	20	30	25	15
2012	35	4	4	50	34	29	17
2012	40	9	10	77	52	37	32
2012	45	20	10	139	76	48	61
2012	50	21	13	206	106	68	62
2012	55	25	18	329	135	87	74

2012	60	33	16	359	156	88	87
2012	65	21	10	303	122	82	81
2012	70	20	11	215	116	69	54
2013	5	4	2	8	4	0	2
2013	10	3	3	6	5	2	2
2013	15	3	1	9	5	5	3
2013	20	6	4	16	11	12	6
2013	25	4	5	31	22	18	16
2013	30	8	2	29	32	20	31
2013	35	11	8	48	46	27	26
2013	40	11	12	82	38	33	30
2013	45	14	17	129	81	52	56
2013	50	22	19	254	104	72	56
2013	55	37	24	362	110	98	72
2013	60	30	19	394	147	101	81
2013	65	24	11	367	119	99	84
2013	70	26	16	272	107	62	54
2014	5	2	1	5	4	3	1
2014	10	0	5	3	4	1	2
2014	15	6	3	13	9	5	5
2014	20	3	1	25	13	12	10
2014	25	6	2	29	21	23	17
2014	30	7	4	37	26	34	20
2014	35	11	11	42	34	36	23
2014	40	12	9	81	52	30	35
2014	45	18	14	132	80	58	55
2014	50	24	13	246	123	81	68
2014	55	38	22	335	142	82	80
2014	60	28	10	369	181	107	100
2014	65	20	27	346	134	122	84
2014	70	21	10	279	106	89	66

Source: WHO mortality database(Egypt).

The numbers of population for both sex:

year	age	male	female
2001	5	4354981	4039272
2001	10	4466094	4154159
2001	15	3992886	3672113
2001	20	2916009	2677604
2001	25	2311454	2523638
2001	30	2211454	2236594
2001	35	2112814	2142503
2001	40	1778032	1726474
2001	45	1542849	1404430
2001	50	1108442	1119924
2001	55	873397	767443
2001	60	772797	768067
2001	65	558905	469271
2001	70	612507	584406
2002	5	4444095	4114829
2002	10	4555898	4230958
2002	15	4072125	3739282
2002	20	2973308	2726078
2002	25	2356418	2569641
2002	30	2254056	2276850
2002	35	2153398	2181201
2002	40	1811672	1757412
2002	45	1572325	1429722
2002	50	1129494	1140176
2002	55	890312	781389
2002	60	787606	782048
2002	65	569941	477981
2002	70	624694	595271
2003	5	4531564	4197498
2003	10	4644945	4315397
2003	15	4151173	3813417
2003	20	3030743	2779921
2003	25	2401810	2620558
2003	30	2297299	2321625
2003	35	2194693	2224166
2003	40	1846160	1791849
2003	45	1602353	1457785
2003	50	1150971	1162587
2003	55	907363	796763
2003	60	802595	797423

2003	65	580889	487452
2003	70	636701	607114
2004	5	4620708	4283584
2004	10	4735938	4403562
2004	15	4232165	3891030
2004	20	3089689	2836378
2004	25	2448477	2673898
2004	30	2341838	2368670
2004	35	2237244	2269283
2004	40	1881788	1828061
2004	45	1633334	1487268
2004	50	1173160	1186113
2004	55	924916	812902
2004	60	818069	813557
2004	65	592136	497353
2004	70	180803	172566
2005	5	4712778	4364621
2005	10	4830310	4486040
2005	15	4316486	3963305
2005	20	3151227	2888845
2005	25	2497256	2723587
2005	30	2388500	2412236
2005	35	2281833	2311163
2005	40	1919286	1861547
2005	45	1665900	1514601
2005	50	1196538	1207953
2005	55	943345	827884
2005	60	834376	828537
2005	65	603941	506608
2005	70	184409	175784
2006	5	4785621	4481655
2006	10	4951638	4604165
2006	15	4382301	4015253
2006	20	3203060	2932074
2006	25	2539679	2758205
2006	30	2426296	2443170
2006	35	2317974	2340698
2006	40	1948989	1907942
2006	45	1690410	1554858
2006	50	1212686	1237371
2006	55	956015	847794
2006	60	853255	840361
2006	65	616704	514038

2006	70	381078	360685
2007	5	4878002	4572625
2007	10	5047224	4697621
2007	15	4466896	4096756
2007	20	3264892	2991590
2007	25	2588705	2814192
2007	30	2473133	2492762
2007	35	2362720	2388210
2007	40	1986612	1946670
2007	45	1723041	1586419
2007	50	1236096	1262487
2007	55	974469	865003
2007	60	869726	857419
2007	65	628609	524472
2007	70	388434	368006
2008	5	4065843	3819821
2008	10	4127349	3834924
2008	15	4517923	4291633
2008	20	4157635	3964226
2008	25	3255450	3338048
2008	30	2494728	2388271
2008	35	2384397	2419586
2008	40	2134842	2086915
2008	45	1937053	1853382
2008	50	1592218	1565757
2008	55	1252526	1084454
2008	60	927685	831683
2008	65	661294	570004
2008	70	415199	399641
2009	5	4161580	3909646
2009	10	4224698	3925153
2009	15	4623077	4391294
2009	20	4252796	4054684
2009	25	3329534	3414079
2009	30	2551237	2442273
2009	35	2438453	2474622
2009	40	2182940	2133693
2009	45	1980536	1894837
2009	50	1627500	1600382
2009	55	1280435	1108500
2009	60	947803	849969
2009	65	676013	582973
2009	70	424375	408641

2010	5	4265412	4007172
2010	10	4330443	4023320
2010	15	4738218	4500674
2010	20	4358058	4155045
2010	25	3411636	3498432
2010	30	2613911	2502494
2010	35	2498485	2535871
2010	40	2236659	2186333
2010	45	2029328	1941694
2010	50	1667511	1639886
2010	55	1312016	1135920
2010	60	971087	871006
2010	65	692772	597495
2010	70	434889	418810
2011	5	4356562	4092730
2011	10	4422983	4109223
2011	15	4839472	4596769
2011	20	4451188	4243760
2011	25	3484541	3573128
2011	30	2669769	2555925
2011	35	2551877	2590015
2011	40	2284456	2233014
2011	45	2072694	1983152
2011	50	1703145	1674900
2011	55	1340053	1160173
2011	60	991839	889603
2011	65	707576	610252
2011	70	444183	427752
2012	5	4470143	4200705
2012	10	4539043	4218732
2012	15	4966481	4718944
2012	20	4569576	4356649
2012	25	3578304	3669086
2012	30	2741843	2624749
2012	35	2621190	2659446
2012	40	2345970	2291960
2012	45	2128762	2035826
2012	50	1748147	1718162
2012	55	1375184	1190248
2012	60	1017278	912511
2012	65	725665	626205
2012	70	456396	439115
2013	5	4671950	4328151

2013	10	4072956	3796132
2013	15	3979957	3745131
2013	20	4227955	4017140
2013	25	4206955	4037141
2013	30	3545962	3444120
2013	35	2738971	2671093
2013	40	2314975	2259079
2013	45	2128977	2089072
2013	50	1880980	1858065
2013	55	1550983	1556054
2013	60	1167988	1185041
2013	65	805991	823029
2013	70	526994	552019
2014	5	4728480	4397298
2014	10	4223820	3959705
2014	15	4311587	4077106
2014	20	4563914	4365276
2014	25	4278670	4119799
2014	30	3411964	3319319
2014	35	2676907	2604220
2014	40	2391651	2348059
2014	45	2194173	2155938
2014	50	1897954	1878438
2014	55	1546889	1536902
2014	60	1129991	1131325
2014	65	789891	800462
2014	70	515620	533635

Source: WHO mortality database(Egypt).