A THREE-STATE MARKOV CHAIN MODEL OF PLASMODIUM FALCIPARUM PARASITEMIA TRANSMISSION IN GHANA

By

MICHAEL ASAMANI POBBI

(BSc. MATHEMATICS AND STATISTICS)

A Thesis submitted to the Department of Mathematics, Kwame Nkrumah University of Science and Technology in partial fulfilment of the requirements for the degree of

Master of Philosophy

College of Science,

OCTOBER 2012

DECLARATION

I hereby declare that this submission is my own work towards the MPhil. (Applied Mathematics) degree and that, to the best of my knowledge, it contains neither material previously published by another person nor material which has been accepted for the award of any other degree of the university, except where due acknowledgment has been made in the text.

…… …… ……

MICHAEL ASAMANI POBBI Signature Date

PG5072010

Certified by:

…… …… ……

Dr. F. T. ODURO Signature Date

(SUPERVISOR)

Certified by:

…… …… ……

Mr. K.F. DARKWAH Signature Date

(HEAD OF DEPARTMENT)

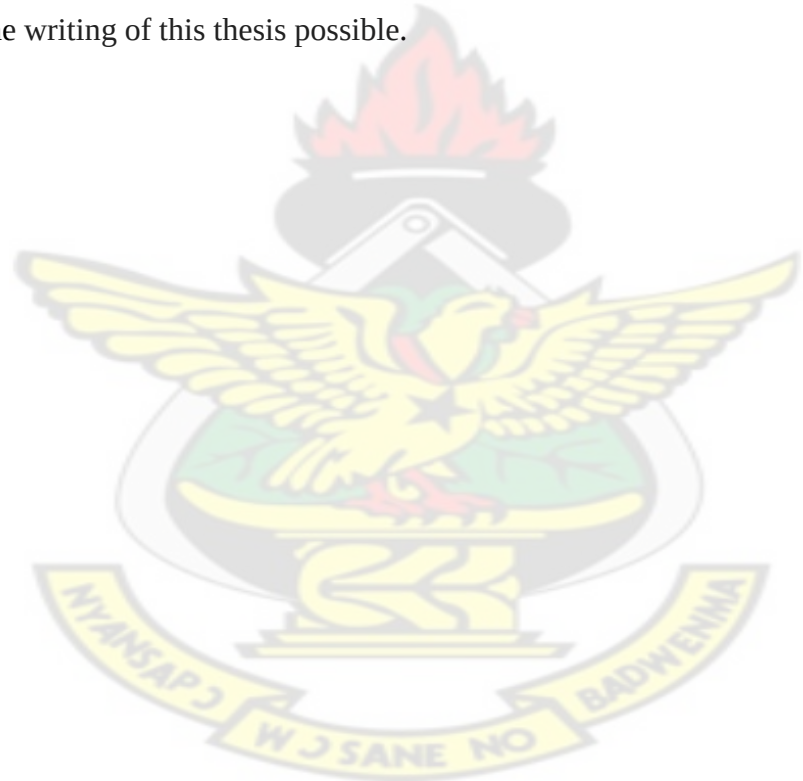
ACKNOWLEDGEMENT

My sincere appreciations first and foremost go to Almighty God who has seen us through this thesis.

I will like to thank my supervisor Dr. F. T. Oduro for his patience, guidance, constructive criticisms and encouragement which enabled me, produce this thesis.

Not forgetting Mr. K. F. Darkwa and the entire department of Mathematics, for their encouragements and advice.

Finally to my family and all others who have contributed in one way or the other to make the writing of this thesis possible.



ABSTRACT

In this thesis, a three state Markov chain model is used to describe the malaria transmission dynamics, using Ghana data from the Ghana Health Service and World Health Organization. The states of the model are defined as susceptible, Infected and dead and the time step for a transition to occur is defined as 8 days. The model is based on the assumptions that individuals are transferred at constant rate between states, and that only one transition is possible between two consecutive surveys. The model is used to determine the steady state probability distributions, the life expectancy of an individual and finally the results are interpreted in terms of malaria control issues. The expected time to a first infection is found to be 11 days and the total duration of the disease (non severe) is found to be 17 days. The life expectancy from the onset of the survey is found to be 55 years for both individuals who are initially infected and those who are initially susceptible.

DEDICATION

I dedicate this work to the most high God who in his own diverse ways has brought me this far and also to my family, friends and loved ones.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACT	Artemisinin based combination theraphy
AS/AQ	Artesunate- amodiaquine
AL	Artemether- lumefantrine
DHIMS	District Health Information Management System
DHS	Demographic and Health Survey
FY	Ghana Health Service
GHS	Fiscal Year
IMaD	Improving Malaria Diagnostics
IPT	Intermittent preventive treatment of infants
IRS	Indoor residual spraying
ITN	Insecticide-treated net
MOH	Ministry of Health
PMT	President's malaria initiative
WHO	World Health Organization

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

The chapter describes the disease under study and the burden of the disease worldwide and nationally. It further discusses the dynamics of the disease, the problem hoped to be resolved, the objectives of the theses, the methodology, reason for the research and scope of the theses.

1.1 BACKGROUND OF STUDY

Malaria in humans is the state of infection with the protozoan parasites of the genus Plasmodium. There are over 120 species of plasmodium but only four of such are responsible for malaria. These are *Plasmodium falciparum, P. vivax, P. malariae* and *P. ovale*. *Plasmodium falciparum* causes the most severe form of the disease, and is responsible for half of the clinical cases and 90% of the deaths from malaria (Nicolas, 2008). Recently a fifth species, *P. knowlesi* has been reported to be responsible for malaria in countries like Thailand --and Philippines. The parasite is indirectly transmitted via a vector, which is a mosquito, hence Malaria is a vector borne disease. The main vector is an infected female *Anopheles* mosquito.

1.1.1 Burden of Malaria

Malaria is one of the leading causes of death in the developing world today. Every year, malaria causes an estimated 1.3–3 million deaths and around half a billion clinical episodes (Breman et al., 2001). The majority of deaths occur in children under the age of 5 years. There are no accurate statistics available, as most cases occur in rural areas, where a large proportion of the population does not have access to hospitals or health care in general.

Malaria today occurs mostly in tropical and subtropical countries, particularly in sub-Saharan Africa and Southeast Asia (Figure A3.1, in appendix 3). In developing countries malaria may account for as much as 40% of public health expenditure, 30-50% of hospital admissions, and up to 50% of outpatient visits to health facilities (WHO, 2006). Critically, malaria is not just caused by poverty; the burden of malaria disease is also an important factor contributing to that poverty. Economic growth in countries with high malaria transmission has historically been lower than in countries without malaria. Some economists believe that malaria is responsible for a growth penalty of up to 1.3% per year in some African countries (Gallup and Sachs, 2001). Not only does malaria result in lost life, and lost productivity due to illness and premature death, malaria also hampers children's schooling and social development through both absenteeism and permanent neurologic damage (Holding and Snow, 2001). Although interventions have been made to mitigate the plague in Ghana, the situation is worrying and is a major cause of deaths. About 3.5 million people contract malaria every year. Approximately 20,000 children die from Malaria every year (25 per cent of the deaths of children under the age of five). Even if a child survives, the consequences from severe malaria such as convulsions or brain dysfunction can hamper long-term development and schooling (United Nations Children's Fund, 2007). Approximately 10% of children in Ghana will die before their 5th birthday, a quarter of those mortalities attributable to malaria. Every year, an average of 15,000 children below the age of five are recorded as dying from the disease, but these fatalities are not limited to children alone. 10% of pregnant women in Ghana also die as a result of this disease, contributing to the total of 13.2% of all deaths nationwide caused by malaria. These statistics compounded with the challenge of chronic illness such as HIV creates public health challenges that are difficult to handle (Appiagyei et al., 2011).

The annual economic burden of malaria is estimated at 1-2 per cent of the Gross Domestic Product in Ghana. Outpatient attendance over the last 19 years illustrates the increasing burden of malaria in Ghana. Communicable diseases accounted for about two third of outpatient visits, but their relative share has changed over time. While there is an overall consistent decrease of other infectious and parasitic diseases (from 31.8% in 1985 to 19.5% in 2003), there has been an increase in malaria cases (from 37.1% in 1985 to 44.7% in 2004) (Adams et al., 2004).

1.1.2 Disease Dynamics

Life Cycle of the parasite in Human Malaria infection begins when female anopheles mosquito bite a human and inject infectious cells known as sporozoites into the person's bloodstream. The sporozoites enter and multiply in the liver to form merozoites. The merozoites leave the liver to invade red blood cells usually after the fifth day. Inside the red blood cells, the merozoites multiply rapidly until the blood cells burst. When these cells burst the released merozoites infect other red blood cells. Some merozoites divide to form gametocytes, immature male and female gametes (cells involved in sexual reproduction). The gametes are involved in transmission of the disease. Symptoms generally appear about the time the red blood cells burst. The bursting cells release waste and toxins along with the merozoites. Fever develops as the immune system responds to the toxins and waste in the blood. Figure 1.1 shows the three main phases of the life cycle of malaria.

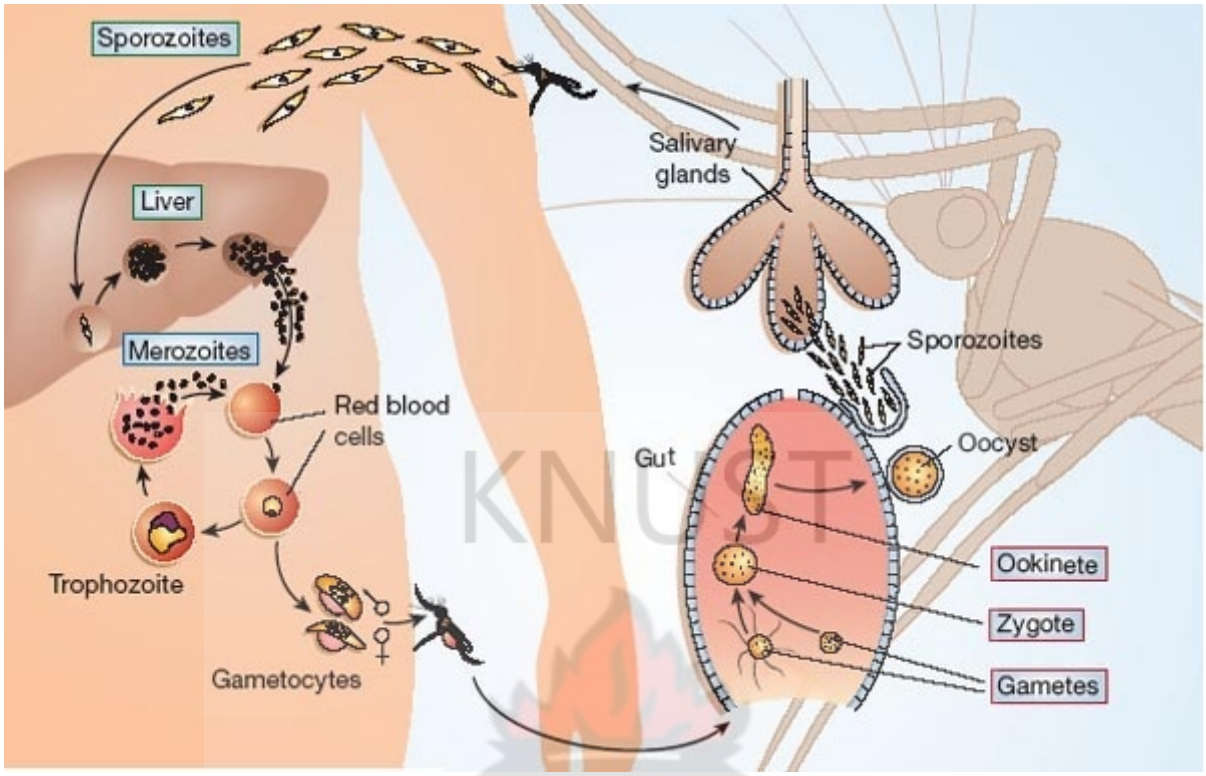


Figure 1.1: The three main phases of malaria cycle.

source: http://www.nature.com/nature/journal/v433/n7022/images/433113a-f1.2.jpg

Malaria is an acute febrile illness. In a non-immune individual, symptoms appear seven days or more (usually 10–15 days) after the infective mosquito bite (WHO, 2012). The first symptoms – fever, headache, chills and vomiting – may be mild and difficult to recognize as malaria. The four malaria parasites all produce fevers and anemia, and, if untreated, can open up a Pandora's box of complications. Some fevers erupt and then disappear; others do not. *P. falciparum* infection is incapable of such relapses (Webb, 2008). For both *P. vivax* and *P. ovale*, clinical relapses may occur weeks to months after the first infection, even if the patient has left the malarious area. These new episodes arise from dormant liver forms known as hypnozoites (absent in *P. falciparum* and *P. malariae*); special treatment – targeted at these liver stages – is required for a complete cure. If a sufferer is cleared of a falciparum infection, she or he is free of the disease unless and until reinfected by another parasite-laden mosquito. If not treated within 24 hours, *P. falciparum* malaria can progress to severe illness often

leading to death. Children with severe malaria frequently develop one or more of the following symptoms: severe anaemia, respiratory distress in relation to metabolic acidosis, or cerebral malaria. In adults, multi-organ involvement is also frequent. In malaria endemic areas, persons may develop partial immunity, allowing asymptomatic infections to occur (WHO, 2012). *Falciparum* can also produce cerebral malaria, a condition that may lead to dangerous sequel such as epilepsy, blindness, cognitive impairments, and behavioral disturbances, or it may lead to coma and death (Webb, 2000). Severe malaria is almost exclusively caused by *P. falciparum* infection, and usually arises 6–14 days after infection. Consequences of severe malaria include coma and death if untreated—young children and pregnant women are especially vulnerable. Splenomegaly (enlarged spleen), severe headache, cerebral ischemia, hepatomegaly (enlarged liver), hypoglycemia, and hemoglobinuria with renal failure may occur. Renal failure is a feature of blackwater fever, where hemoglobin from lysed red blood cells leaks into the urine. Severe malaria can progress extremely rapidly and cause death within hours or days (health2spread, 2012). In most severe cases of the disease, fatality rates can exceed 20%, even with intensive care and treatment **(**Kain et al., 1998).

Falciparum malaria can be treated to prevent complications. According to WHO guidelines 2010 on the treatment of *falciparium* malaria, artemisinin-based combination therapies (ACTs) are the recommended first line antimalarial treatments for uncomplicated malaria caused by *P. falciparum*. The following ACTs are recommended by the WHO: artemether plus lumefantrine; artesunate plus amodiaquine; artesunate plus mefloquine; artesunate plus sulfadoxine-pyrimethamine and dihydroartemisinin plus piperaquine. The choice of ACT in a country or region will be based on the level of resistance to the constituents in the combination. Artemisinin and its derivatives should not be used as monotherapy in

uncomplicated falciparum malaria. As second-line antimalarial treatment, when initial treatment doesnot work or stops working, it is recommended to use an alternative ACT known to be effective in the region, such as: Artesunate plus tetracycline or doxycycline or clindamycin; Quinine plus tetracycline or doxycycline or clindamycin. Any of these combinations are to be given for 7 days (WHO, 2010).

Factors which contribute to the transmission of *P. falciparum* are numerous. Some of these factors include are temperature, immunity of population, seasonal changes, sanitation, altitude and so on. *Falciparum* is not transmitted at stable temperatures below 19^{o} C. This is one condition which largely affects the distribution of the disease. With climate changes the distribution of the disease is expected to change across the globe.

An individual may also be immune to the parasite, but a question to ask is how much of the parasite can a person receive to remain immune? Natural defence mechanisms (or innate factors) against malaria are most apparent in populations continually exposed to malaria parasites. For example, inherited conditions such as sickle cell anaemia and beta-thalassaemia, which cause deformities in red blood cells and are common in people from malarious regions, make it more difficult for malaria parasites to infect red blood cells (Cross, 2004).

Seasonal changes as well affect spread of the *falciparum* parasite since mosquitoes breed more during the rainy seasons than in the dry seasons. Much work have been done to aid explain the spread and prevention of malaria which is increasing in recent times, yet much have not been achieved predicting disease parameters due to the complexity of the disease transmission.

Understanding of the dynamics of diseases is very essential. Mathematical models have been useful in description of diseases but not sufficient. Today we are faced with the need to predict the dynamics and transmission of indirectly transmitted diseases with a greater accuracy and over longer periods of time, and more often with limited empirical data (Ngwa, 1999). Quantifying and understanding disease dynamics will help in the prevention and control of emerging infectious diseases.This could be however challenging, since existing data on patients from hospitals are not accurate enough for such work. Also due to endemicity of the disease many people tend to self medicate and may not report to hospitals. In exploring the dynamics of diseases, it is desirable to quantify parameters such as force of infection, disease prevalence, and infection and recovery probabilities. These can be achieved by mathematical and statistical methods; however can be very involving. We can describe the process of an individual being infected with a disease as an experiment: with the outcomes: being infected, not infected, recovery or even death. By sampling of individuals in various disease states (e.g. uninfected and infected) at discrete time intervals it is possible to apply stochastic processes to such a data. Markov chain models, a discrete time stochastic process, is one method that is used to describe disease dynamics at the individual level, thence creating room for inferences regarding populations (Zipkin et. al., 2010). Parameters can be estimated at the individual level but can be extended to the population level.

1.2 PROBLEM STATEMENT

Every year, malaria causes an estimated 1.3–3 million deaths and around half a billion clinical episodes. The majority of deaths occur in children under the age of 5 years. Malaria today occurs mostly in tropical and subtropical countries, particularly in sub-Saharan Africa and Southeast Asia. The annual economic burden of malaria is estimated at 1-2 per cent of the Gross Domestic Product in Ghana. Outpatient attendance over the last 19 years illustrates the increasing burden of malaria in Ghana (Adams et. al., 2004). More understanding based on mathematical models will help in the prevention and curbing of such devastating impacts. This study attempts to construct one such a model.

1.3 OBJECTIVES

The main objective for this work is to develop an S-I-D Markov chain model of malaria transmission dynamics using Ghana data. The specific objectives of this study are to:

1) Determine the steady state probability distributions of the model.
2) Compute the life expectancy of an individual using the model and available data.
3) Interpret the results in terms of malaria control issues.

1.4 METHODOLOGY

Secondary data was obtained from the Ghana Health Service (GHS), World Health Organization (WHO), and the Central Intelligence Agency (CIA). The first two sources provided the incidence rate from 1997 to 2008 and admission and mortalities from 1997 to

2009 respectively. The third source, which was used to verify results, presents the national demographic data for 2008.

The time step for a transition to occur was defined and also the states for the Markov model were partitioned so that it was only possible to enter one state after one time step. The model was based on the assumptions that individuals are transferred at constant rate between states, and only one transition is possible between two consecutive surveys.

The probability of transiting into a new state is given as the frequency of moving from a previous state to the new state divided by the total frequency of moving from the previous state to any other state. Subsequently transition probabilities for the time step are computed and the transition matrix populated using the time step and available data.

The long run behavior of the process was determined by computing the steady states of the first order Markov process. The Markov model was then extended to estimate the first transition probabilities distribution for an individual to become infected and to recover respectively and then the expected time of an individual to first infection and recovery. Finally the fundamental matrix was determined which also allows for the calculation of the life expectancy using person below 5 years as sampling unit.

1.5 JUSTIFICATION

Malaria is on the rise though intervention from government and non government agencies had been made. Outpatient attendance over the last 19 years illustrates the increasing burden of malaria in Ghana (Adams et. al., 2004). About 3.5 million people contract malaria every year and in every 30 seconds malaria kills a child – about 3,000 children every day. Approximately 20,000 children die from Malaria every year (25 per cent of the deaths of

children under the age of five) in Ghana. The annual economic burden of malaria is estimated 1-2 per cent of the Gross Domestic Product in Ghana (United Nations Children's Fund, 2007). Although much work have made, further studies to quantify and understand disease dynamics will help in the prevention and control of emerging infectious diseases (such as HIV/AIDS and influenza) in Ghana, especially with the changing climatic conditions. It would also help in understanding the development and spread of drug resistant strains (especially with regard to malaria and tuberculosis) and hence aid policy making in Ghana (Bhadra et. al., 2009). Forecasting emerging epidemics can also be done by parameter estimation for infectious disease models.

1.6 SCOPE AND LIMITATION

The study is focused at explaining the dynamics of malaria using first step discrete Markov models. A three state Markov model in which the third model happens to be an absorbing state will be formulated for Ghana. The complexities of interaction between the vector and the host, the human body in response to *P. falciparum* and many other interactions makes explaining malaria dynamics a bit difficult. However, we choose a much simplified but efficient model to explain basic dynamics.The most important progeny of the methodology is to calculate life expectancies using discrete data. It is assumed that the defined states clearly and correctly characterize the process and that an individual sampled would be in one of the states at the defined time step.

1.7 THESIS ORGANIZATION

The remaining sections of this thesis are organized as follows. Chapter 2 reviews works in different fields where Markov processes were applied and also other theoretical works relevant to this study have been touched on. Chapter 3 also describes the Markov Chain modelling, some general applications of the model, how the model can be applied to malaria transmission dynamics and finally how we compute for our parameters. Chapter 4 discusses the calculated results for our expected disease metrics, the average life expectancies, simulation and how these results can be incorporated in malaria control. Finally, Chapter 5 is a conclusion based on the results obtained and the recommendations needed to help in the control of malaria.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

In this chapter relevant literature and other theoretical works relevant to this study have been touched on. A few works in different (non epidemiology) fields where Markov processes has been applied are first reviewed to explain the dynamics of processes, then the application of Markov process in explaining disease dynamics is discussed narrowing down to the use of discrete first order Markov processes in explaining disease dynamics.

2.2 APPLICATIONS OF MARKOV MODEL EXPERIMENTS TO DYNAMICS OF EVENTS IN VARIOUS FIELDS

In a paper in the field of economics written by Ciecka et al. (2003), the researchers explored the rich implications of the Markov nature of the increment-decrement model. They wanted to predict the future of the concept of worklife expectancy. In 1982, the Bureau of Labour Statistics (BLS) introduced the increment-decrement model of labour force activity in Bulletin 2135. A subsequent BLS publication, Bulletin 2254, in 1986 also used the increment-decrement methodology. Worklife expectancies were the most important progeny of their methodology. They showed that the increment-decrement model is a valuable construct that enables one to describe many aspects of labour market activity that should be of interest to economists, sociologists, and demographers.

A three-state Markov model was developed for speech on the telephone lines. The model considered the alternate occurrence of the telephone calls and the intercall gaps on the telephone lines. During a phone call there occur several talkspurts (speech without break) and pauses (silence without break) in the user's speech. The three types of events, intercall gaps (large gaps) talkspurts and pauses occurring on the telephone lines are assumed to have negative exponential density functions with different transition rate parameters. The steady state probability distribution, average and variance of the number of busy channels and the system utilization are evaluated as a function of call loss fraction (P_{0f}). The Synchronous Time Division Multiplexing (STDM) system with number of channels in the group equal to 6 and 24 is considered. The model is also applied to Time Assignment Speech Interpolation (TASI) system and the relationships among the number of user terminals, speech Freeze out Fraction (FOF) and system utilization are obtained for various values of P_{0f}. The STDM and TASI systems applied to the group of 6 channels are simulated on the EC-1030 computer to check the validity of the analytical results. The results of this study are portrayed on graphs and may be used as guide lines in the design of TASI systems (Kekre et al., 1977).

Hidden Markov models have recently been used to model single ion channel currents as recorded with the patch clamp technique from cell membranes. The estimation of hidden Markov models parameters using the forward-backward and Baum-Welch algorithms can be performed at signal to noise ratios that are too low for conventional single channel kinetic analysis; however, the application of these algorithms relies on the assumptions that the background noise be white and that the underlying state transitions occur at discrete times. To address these issues, they presented an "*H*-noise" algorithm that accounts for correlated background noise and the randomness of sampling relative to transitions and discuss three

issues that arise in the practical application of the algorithm in analyzing single channel data. First, they described a digital inverse filter that removes the effects of the analog antialiazing filter and yields a sharp frequency roll-off. This enhances the performance while reducing the computational intensity of the algorithm. Second, the data may be contaminated with baseline drifts or deterministic interferences such as 60-Hz pickup. We propose an extension of previous results to consider baseline drift. Finally, we describe the extension of the algorithm to multiple data sets (Venkataramanan et al., 2002).

In a research by Kekre (1978), a queueing model with finite waiting room, Poisson arrivals, multiple synchronous outputs and the number of active servers varying through the birth and death process is studied. Variation of active servers is obtained by providing a switch in each server. The switches are controlled through the birth and death process having finite population. The relationships between overflow probabilities, buffer size and expected queueing delay due to buffering are obtained. An efficient algorithm for computation of steady state probabilities is developed. Two digital voice-data systems; Data multiplexing in speech with Synchronous Time Division Multiplexing (STDM) and Time Assignment Speech Interpolation (TASI) systems are considered for the application of the queueing model studied. The results of this study are portrayed on graphs and may be used as guidelines for designing a buffer in digital systems. The two voice-data systems using the developed model are simulated on the EC-1030 computer to check the validity of the analytical results. Although this problem arose in the study of data interpolation in STDM and TASI systems, the queueing model developed is quite general and may be useful for other industrial applications.

A stochastic approach, namely, a Markov chain, was been adopted to simulate the dynamics of complex reactions in a flow chemical reactor without either solving directly differential equations governing the performance of the reactor or obtaining a closed form solution to such equations. All calculations were carried out iteratively with the aid of a computer. The results were in good agreement with the known results obtained from the deterministic approach. The present technique can be applied to a variety of flow reactors with complex chemical reactions, which cannot be handled easily by deterministic approaches (Too, 1983).

Fan et al. (1985) employed a Markov chain approach to generate the residence time distribution (RTD) curves of various flow models. This technique is shown to be efficient and useful in modeling complex flow systems which consist of ideal mixing cells, plug flow zones, split streams, recycle streams and any combination of these for which the mathematical description is extremely difficult if not impossible. The expressions for the mean and variance of the residence time distributions for a flow system are elucidated from the stochastic viewpoint. Two examples were simulated numerically; the agreement between the RTD's derived from the exact statistical treatment and those based on the present Markov chain approach appears satisfactory. The technique proposed in this study enables us to generate the RTD curve for any flow system and to match it to the experimental data. Furthermore, the applicability of the Markov chain approach for generating the cycle time distribution curve for a circulating system is discussed.

A continuous time Markov chain (Markov process), was employed to analyze and model, in a unified fashion, both the kinetics of unimolecular reactions and mixing accompanied by flow in a continuous flow reactor under an unsteady state operation; the results reduce to those under the corresponding steady state operation in the limit as

In another research by Hiligsmann et al. (2009), the researchers developed and validated an original Markov micro simulation model to accurately assess the cost-effectiveness of prevention and treatment of osteoporosis. Although Markov models are increasingly used in economic evaluations of treatments for osteoporosis, most of the existing evaluations are cohort-based Markov models missing comprehensive memory management and versatility. They developed a Markov micro simulation model with a lifetime horizon and a direct health-care cost perspective. The patient history was recorded and was used in calculations of transition probabilities, utilities, and costs. They carried out an example calculation for alendronate therapy, to test the internal consistency of the model. Then, external consistency was investigated by comparing absolute lifetime risk of fracture estimates with epidemiologic data.

The results revealed that, for women at age 70 years, with a twofold increase in the fracture risk of the average population, the costs per quality-adjusted life-year gained for alendronate therapy versus no treatment were estimated at €9105 and €15,325, respectively, under full and realistic adherence assumptions. Also all the sensitivity analyses in terms of model parameters and modelling assumptions were coherent with expected conclusions and absolute lifetime risk of fracture estimates were within the range of previous estimates, which confirmed both internal and external consistency of the model. They concluded that Micro simulation models present some major advantages over cohort-based models, increasing the reliability of the results and being largely compatible with the existing state of the art, evidence-based literature. The developed model appeared to be a valid model for use in economic evaluations in osteoporosis.

Markov chain Monte Carlo methods are frequently used in the analyses of genetic data on pedigrees for the estimation of probabilities and likelihoods which cannot be calculated by existing exact methods. In the case of discrete data, the underlying Markov chain may be reducible and care must be taken to ensure that reliable estimates are obtained. Potential reducibility thus has implications for the analysis of the mixed inheritance model, for example, where genetic variation is assumed to be due to one single locus of large effect and many loci each with a small effect. Similarly, reducibility arises in the detection of quantitative trait loci from incomplete discrete marker data. The paper aimed to describe the estimation problem in terms of simple discrete genetic models and the single-site Gibbs sampler. Reducibility of the Gibbs sampler was discussed and some current methods for circumventing the problem outlined (Sheehan, 2000).

In the paper by Patten et al. (2005), Markov models were extended to describe the longitudinal course of the disorder. They described that most epidemiological study of major depression report period prevalence estimates. These are of limited utility in characterizing the longitudinal epidemiology of this condition. They stated also that Markov models provide a methodological framework for increasing the utility of epidemiological data. Markov models relating incidence and recovery to major depression prevalence have been described in a series of prior papers.

In a paper by Sesso et al. (2002), the researcher stated that the use of Markov chain models is a technique commonly used to simulate long-term progressive diseases. These models represent recurring events associated with an ongoing risk, assuming that patients reside in one of a finite number of health states. Subjects may transition from one health state to

another during a defined interval of time called a cycle. The life expectancy benefits of antihypertensive treatment, based on both systolic and diastolic blood pressure reduction, was estimated with a cardiovascular disease event Markov model with prospective data from 57 573 men and women. Seven patient states were defined, including (1) no cardiovascular disease, (2) stroke, (3) myocardial infarction, (4) revascularization, (5) history of cardiovascular disease, (6) noncardiovascular disease death, and (7) cardiovascular death. Risk functions were developed from gender-specific multivariate Cox proportional hazards models for primary events and age-, smoking-, and diabetes-adjusted models for secondary events. At baseline we assumed (1) hypothetical pretreatment blood pressures of 160/95 or 150/90 mm Hg; (2) strategies A and B lower blood pressure by 20/13 and 13/8 mm Hg, respectively; and (3) baseline age of 35 years. For subjects initially at 160/95 mm Hg, those with antihypertensive treatment, antihypertensive treatment and diabetes, or antihypertensive treatment, diabetes, and currently smoking had corresponding gains in life expectancy of 2.43, 2.80, and 2.43 years for Strategy A. An initial blood pressure of 150/90 mm Hg resulted in similar gains. Compared with Strategy B, with blood pressure reductions of 13/8 mm Hg, Strategy A provided additional gains in life expectancy of 0.84, 0.99, and 0.87 years for those with antihypertensive treatment, antihypertensive treatment and diabetes, or antihypertensive treatment, diabetes, and currently smoking. The initial blood pressure level did not affect the magnitude of life expectancy gains for equivalent blood pressure reductions. Greater gains in life expectancy among hypertensive and diabetic women suggest that blood pressure lowering may yield greater benefits in selected subgroups.

Unique to survival analysis of veterinary clinical data is classification of observations from euthanized animals. The first study highlighted limitations of Kaplan-Meier product limit

analysis (KM) of veterinary clinical data. Three data sets with different outcome proportions (alive, lost-to-follow-up, dead due to disease, dead due to other, euthanized due to disease, euthanized due to other) were used. Different classifications of observations from euthanized animals caused inconsistent conclusions of significant differences between strata within data sets. At times, ranking of median survival time estimates for strata was reversed. The KM was found inappropriate to evaluate observations from euthanized animals. This finding, coupled with restriction of KM to two-state description of disease (alive to outcome), prompted exploration of an alternate analysis method. Markov models allow modeling of multiple health states and outcomes. A 5-state, time-homogeneous, Markov chain was used for a cohort of 64 dogs with generalized lymphoma. The model contained two transient (WELL, TOXIC) and three absorbing (DEAD, EUTHANASIA, LOST-TO-FOLLOWUP) states. The transition probability matrix (P) was used to iterate future transitions and survival probabilities. Matrix solution and Monte Carlo simulation were used to estimate survival time. Estimates appeared reliable. Markov modeling was extended for comparison of vaccine-associated sarcoma progression after treatment in a cohort of 294 cats. For a 5-state model, transition probabilities derived from exponential transformation of incidence rates were used to construct P for each treatment – NONE (no surgery), SX (surgery) and SX+RAD (surgery and radiation). Monte Carlo estimates of durations in transient states and expected survival showed SX+RAD prolonged expected survival significantly longer than SX than NONE. Commitment to repeated treatment with surgery and radiation did prolong expected survival of cats with vaccine-associated sarcoma. Assumptions of Markov modeling did not appear prohibitive for analysis of veterinary clinical data and further exploration (Hosgood, 2002).

The application of Markov models in recovery and restoration is a natural extension to their success in modeling ecological succession and disturbance. Due to the continual shifting of ecosystem function during the restoration, however, and the immigration and extinction of local species, the model must be reconsidered in terms of its ecological analogues. Modeling groups of species, classified by functional groups, is suggested. An extended Markov model, the hidden Markov model, is introduced as a method of linking the structure and function of the ecosystem in the modeling construct (Tucker et al., 2004).

Commenges (1999) in his work presented the influence of covariates and different durations and time-dependent variables synthesized using explanatory processes, and also general additive model for transition intensities. He first discuss the main assumptions which can be made for multi-state models: the time-homogeneity and semi-Markov assumptions, the problem of choice of the time scale, the assumption of homogeneity of the population and also assumptions about the way the observations are incomplete, leading to truncation and censoring. Different inference approaches, including penalized likelihood, are considered. Finally three examples of application in epidemiology are presented and some references to other works are given.

Data from three national surveys conducted by the Canadian national statistical agency (Statistics Canada) were used in this analysis. These data were integrated using a Markov model. Incidence, recurrence and recovery were represented as weekly transition probabilities. Model parameters were calibrated to the survey estimates. The population was divided into three categories: low, moderate and high recurrence groups. The size of each category was approximated using lifetime data from a study using the WHO Mental Health

Composite International Diagnostic Interview (WMH-CIDI). Consistent with previous work, transition probabilities reflecting recovery were high in the initial weeks of the episodes, and declined by a fixed proportion with each passing week. They concluded that Markov models provide a framework for integrating psychiatric epidemiological data. The study extended the Markov approach by distinguishing several recurrence categories.

Many chronic diseases have a natural interpretation in terms of staged progression. Multistate models based on Markov processes are a well-established method of estimating rates of transition between stages of disease. However, diagnoses of disease stages are sometimes subject to error. The paper presented a general hidden Markov model for simultaneously estimating transition rates and probabilities of stage misclassification. Covariates could be fitted to both the transition rates and the misclassification probabilities. For example, in the study of abdominal aortic aneurysms by ultrasonography, the disease was staged by severity, according to successive ranges of aortic diameter. The model was illustrated on data from a trial of aortic aneurysm screening, in which the screening measurements were subject to error. General purpose software for model implementation was developed in the form of an R package (Jackson et al., 2003).

Debanne et al. (2000) developed a computer-implemented, multivariate Markov chain model toproject tuberculosis (TB) incidence in the United States from 1980 to 2010 in disaggregated demographic groups. Uncertainty in model parameters and in the projections is represented by fuzzy numbers. Projections were made under the assumption that current TB control measures will remain unchanged for the projection period. The model predicted that the rate of decline in the number of cases among Hispanics will be slower than among white non-Hispanics and black non-Hispanics—a prediction supported by the most recent data.

A mathematical model of three-dimensional transport of gas-phase contaminants in indoor environments based on a Markov chain, the Markov model, was extended to solid-phase contaminant transport in indoor environment. The performance of the model was affirmed by comparison with simplified transport models, and a previously published study of poly-dispersed particulate transport in a quiescent environment. The experiment was conducted to measure the transport and fate of mono-dispersed fluoresce in-tagged particles with nominal aerodynamic diameters of 3 μm and 14 μm under turbulent and quiescent conditions in a room-scale chamber. Advection and turbulence were characterized using 3-axis anemometery and tracer gas studies. Parameterization of turbulence in the Markov model was explored: Turbulent diffusion coefficients and fluctuating velocities were associated with mixing time through simulation, but the simulated fluctuating velocities were not strongly correlated with measured fluctuating velocities. The Markov model did not replicate the experimental results with high fidelity, but this may be due in part to limitations in the anemometry, and the complexity of the aerosol release. The public health significance of the mathematical modeling was demonstrated in the context of Mycobacterium tuberculosis exposure and infection risk in commercial passenger aircraft. Infection risk to passengers, and the impact of exposure reduction strategies were quantitatively assessed using the Markov model in conjunction with Monte Carlo simulation. The expected infections ranged from 10 -6 to 10-1 per 169 susceptible passengers, for an exponential (k = 1) dose-response function. Use of respiratory protection, surgical masks or filtering face piece respirators, by the infectious source and or the susceptible passengers reduced infection risk by a factor of 2-10 (Jones, 2008)

Surveillance data for communicable nosocomial pathogens usually consist of short time series of lownumbered counts of infected patients. These often show overdispersion and

autocorrelation. Inferences that depend on such analyses cannot be considered reliable when patient-to-patient transmission is important. Hence a new method was proposed for analyzing these data based on a mechanistic model of the epidemic process by developing a 'structured' hidden Markov model where the underlying Markov chain is generated by a simple transmission model. Both structured and standard (unstructured) hidden Markov models were applied to time series for three important pathogens. Results showed that both methods can offer marked improvements over currently used approaches when nosocomial spread is important. Also in comparison to the standard hidden Markov model, the new approach was more parsimonious, biologically plausible, and allowed key epidemiological parameters to be estimated (Cooper and Lipsitch, 2004).

Markov simulation model to assess the impact of changing trends in coronary heart disease incidence on requirements for coronary artery revascularization procedures in Western Australia was done by Mannan et al. (2010). Different CHD incidence trend scenarios were developed to explore the effect of changing CHD incidence on requirements for coronary artery bypass graft (CABG) and percutaneous coronary interventions (PCI), together known as coronary artery revascularization procedures (CARPs) by applying a validated Markov simulation model to the population of Western Australia. Results showed that, the most dominant component of CHD incidence is the risk of CHD hospital admission for those with no history of CHD and if this risk levelled off and the trends in all other risks continued unchanged, then the projected numbers of CABGs and PCIs are only minimally changed. Further, the changes in the projected numbers remained small even when this risk was increased by 20 percent (although it is an unlikely scenario). However, when the other CHD incidence components that had also been declining, namely, the risk of CABG and that of CHD death for those with no history of CHD, were also projected to level off as these were

declining in 1998-2000 and the risk of PCI for those with no history of CHD (which was already increasing) was projected to further increase by 5 percent, it had a substantial effect on the projected numbers of CARPs. They concluded that there was need for dramatic changes to several CHD incidence components before it has a substantial impact on the projected requirements for CARPs. Continued monitoring of CHD incidence and also the mix of initial presentation of CHD incidence is required in order to understand changes to future CARP requirements.

In a work by Lekone and Finkenst¨adt (2006), a stochastic discrete-time susceptible-exposed-infectious-recovered (SEIR) model for infectious diseases was developed with the aim of estimating parameters from daily incidence and mortality time series for an outbreak of Ebola in the Democratic Republic of Congo in 1995. The incidence time series exhibit many low integers as well as zero counts requiring an intrinsically stochastic modeling approach. In order to capture the stochastic nature of the transitions between the compartmental populations in such a model they specified appropriate conditional binomial distributions. In addition, a relatively simple temporally varying transmission rate function was introduced which allowed for the effect of control interventions. They developed Markov chain Monte Carlo methods for inference that were used to explore the posterior distribution of the parameters. The algorithm was further extended to integrate numerically over state variables of the model, which are unobserved. This provided a realistic stochastic model that can be used by epidemiologists to study the dynamics of the disease and the effect of control interventions.

Markov chain models are powerful tools; applicable to the study of disease dynamics that allow straightforward calculations of easily interpretable metrics of interest including probabilities of infection or recovery, expected times to initial infection, duration of illness and life expectancies for susceptible and infected individuals (Zipkin et al., 2010).

They presented the basic principles and assumptions behind Markov chain modelling with an intuitive interpretation of parameter estimates and a step-by-step guide (including software code) for implementing this approach in the study of wildlife diseases. They also included an explanation of the estimation process necessary to implement Markov chain modelling (i.e. estimating the probability of state transitions between consecutive time steps) from typical survey data. They demonstrated the usefulness and ease of calculation of Markov chains through an example using a house finch Carpodacusmexicanus–Mycoplasma gallisepticum (MG) system. Their results showed how semi-weekly transition estimates of susceptible and infected individuals can be used to estimate a wide array of seasonal disease-associated metrics.

A model, using stochastic processes, was developed by Verma et al., (1983) to estimate some epidemiological parameters of malaria in a homogeneous population from longitudinal data. Assessments of transition probabilities from one state of health to the other were made taking "lost to follow-up" as a competing risk. The model was based on the assumptions that individuals are transferred at constant rate between states, and only one transition is possible between two consecutive surveys. It showed a good fit to the observed data; the model was also simple to understand and could easily be used if computer facilities are not available.

In a paper by Gani et al. (1971), Markov chain methods were applied to chain binomial models in epidemics. In both Greenwood and Reed-Frost chain binomial models, it was shown that the susceptibles and susceptibles together with infectives respectively form Markov chains. These chains are used to obtain probabilities for the duration time and the total number of cases in an epidemic. A study of chain binomial models as Markov chains imbedded in continuous time processes is made. A practical application of the effects of inoculation on an epidemic was carried out, and some numerical results for the mean duration times and mean numbers of cases given.

Plasmodium falciparum has a complex transmission cycle. Public health planning and research would benefit from the ability of a calibrated model to predict the epidemiologic characteristics of populations living in areas of malaria endemicity. The paper described the application of Bayesian calibration to a malaria transmission model using longitudinal data gathered from 176 subjects in Ndiop, Senegal, from July 1, 1993, to July 31, 1994. The model was able to adequately predict *P. falciparum* parasitemia prevalence in the study population. Further insight into the dynamics of malaria in Ndiop was provided. During the dry season, the estimated fraction of nonimmune subjects goes down to 20% and then increases up to 80%. The model-predicted time-weighted average incidences contributed by nonimmune and immune individuals are 0.52 cases per day and 0.47 cases per day, respectively. The median times needed to acquire infection (conversion delay) for nonimmune and immune individuals are estimated at 39 days and 285 days, respectively (Cancré et al., 2000).

Multistate Markov models are frequently used to characterize disease processes, but their estimation from longitudinal data is often hampered by complex patterns of incompleteness. Two algorithms for estimating Markov chain models in the case of intermittent missing data in longitudinal studies, a stochastic EM algorithm and the Gibbs sampler, were described. The first was viewed as a random perturbation of the EM algorithm and was appropriate when the M step is straightforward but the E step was computationally burdensome. It lead to a good approximation of the maximum likelihood estimates. The Gibbs sampler was used for a full Bayesian inference. The performances of the two algorithms were illustrated on two simulated data sets. A motivating example concerned with the modelling of the evolution of parasitemia by *Plasmodium falciparum* (malaria) in a cohort of 105 young children in Cameroon was described and briefly analyzed (Deltour, 1999).

A three-state Markov model taking into account clinical signs of malaria infections by P. falciparum was described in a research by Richard et al., (1993). They defined three states considered are the noninfected (state 0), the infected exhibiting no clinical signs (state 1), and the infected with clinical signs (state 2). Methods for estimating the transition rates from longitudinal data were indicated. Their model was used to assess the effect on children of an intervention trial on the use of mosquito nets impregnated with insecticide. The trial was conducted in West Africa (Burkina Faso) between 1985 and 1987. The analysis showed that the intervention was most effective on transition rates between state 1 and state 2.

CHAPTER 3

METHODOLOGY

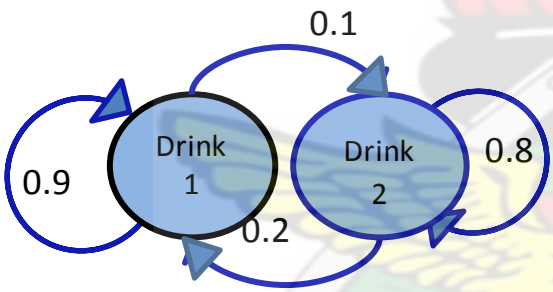
3.0 INTRODUCTION

In this chapter, the theory of Markov processes is discussed in terms of its definition, simulation of Markov chains, the steady states, classification states, absorbing states, as well as how to estimate relevant transition probabilities for the study.

3.1 MARKOV CHAINS

Suppose that with each element s of a sample space S of some random experiment E, we associate a function X (t,s), where t belongs to

A change of state of the system is referred to as a transition, and the probabilities associated with various state-changes are called transition probabilities. The probability of moving from state *I* at time t to state *j* at time t+1 given we are currently in state *I*, p_{ij} defined as;



These equations imply that a given process will be in some state *k* after exactly *m* (less than *n*) states when moving from state *I* to state *j* after *n* steps. Where

3.3 CLASSIFICATION OF STATES OF A MARKOV CHAIN

A state j is said to be **accessible** from a state I (written $I \rightarrow j$) if a system started in state I has a non-zero probability of transitioning into state j at some point. Formally, state j is accessible from state I if there exists an integer $n \geq 0$ such that

transient. Since a recurrent state definitely will be revisited after each visit, it will be visited infinitely often if the process continues forever. If the process enters a certain state and then stays in this state at the next step, this is considered a return to this state. Hence, the following kind of state is a special type of recurrent state.

A state *I* is called **absorbing** if it is impossible to leave this state. Therefore, the state *I* is absorbing if and only if

$$P = \begin{array}{c} \text{state} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

process starts in either of these states, it can never leave these two states. Furthermore, whenever the process moves from one of these states to the other one, it always will return to the original state eventually.

Another useful property of Markov chains is periodicities. The period of state *I* is defined to be the integer

distribution (or invariant measure) if its entries

the process enters state j from state I, since q_{ij} is the transition rate from state I to state j given that the process is in state i. By summing over all

We extension of our example above we define a discrete three states compartmental model with the following states: susceptible (state 0), infected (state 1) and dead (or removed; state 2). The model is graphed as below:

Figure 3.2: A graph of the three state Markov Chain.

Definition of states in the Model:

S=Uninfected state. It consists of humans who have not being exposed before and individuals who have recovered from an infection.

I= Infectious state; consist of infected humans and carriers of the disease

D= Death or removed state.

Parameters of the Chain (Probabilities of transition):

p_{00}= Probability of remaining in a susceptible state.

P_{01}= transition Probability from susceptible state to infectious state.Referred to in literature as discrete time force of infection.

P_{02}= transition probability from susceptible to a removed state.

P_{10}= transition probability from infectious state to susceptive state.

P_{11}= Probability of remaining in an infectious state.

P_{12}= transition probability from an infectious state to a removed state.

P_{22}= Probability of remaining in an removed state.

These p_{ij}'s in the study of diseases actually give us a measure the frequency of both disease occurrence and deaths from the disease. The rates also tell us how fast the disease is occurring in a population.

The figure 3.2 shows that a susceptible individual that receives the Plasmodium parasites from a mosquito bite can remain susceptible, become infected or die in a given time step. Similarly, infected individuals can recover, remain infected or die. However once an individual dies, it remains dead and hence a transition to another state is not possible. The susceptible and infected states are therefore transient states while death is an absorbing state because once an individual dies, it remains dead.

We could also have a four state compartmental model where we can separate the Recovered individuals from S as below.

Figure 3.3: A graph of a four state S-I-R-D Markov chain.

Although the latter model is an expanded form of the three-state discrete model, the dynamics are very different when critically considered. This time we may find the immunity of recovered individuals or the immunity of naïve individuals separately if the additional state is a recovery state or immune state respectively. Until date to actual values of immunity have not being found and to pursue for these parameters would require large resources, years, corporate expertise Also note that in this model the S are those who have never being exposed to the parasite.

Additional state therefore can be added to account for disparities between I and recovered individuals or to incorporate other details of interest. For example, an immune state can be included to encompass individuals that have recovered from infection and are incapable of contracting the disease in the future. Similarly, infected individuals can recover, remain infected or die. The incubation period for the *P. falciparum* parasite (the delay from infected bite to fever) within the human is from 7 to 14 days. However due to the endemicity in Ghana most people would seek for cure when clinical signs to manifest we chose an average of 8 days begin. Fortunately, this disease is curable when proper, early and prompt treatment is sought. Smelters(1997) indicates that When the correct chemotherapy, diet and adequate rest

are given to a patient, the maximum length for recovery should be seven days, three days for detention in the hospital and four for rest at home depending on the protocol of the health facility. A more recent work (Chuks et al., 2009) in a case study of Volta Region in Ghana however indicated that it takes from 1 to 14^{+} days length of stay of patient at health facility on admission for malaria. Hence averagely individual spends a close to eight days or more depending on the conditions afore mentioned.

The most important stage in Markov modelling is to determine the one step transition probabilities. To estimate this, it is very important to choose a reasonable time step for which a transition will occur. An illogical time step could result in wrong results and interpretation. With previously explained time frame for infection, recovery and death due to disease, we can choose the time step for sampling as 8 days during which much severe instances would not be observed. Hence the model will be based on mild cases of the disease.

The death state may not be necessarily death in all cases, *falciparum* can also produce cerebral malaria, a condition that may lead to dangerous sequel such as epilepsy, blindness, cognitive impairments, and behavioral disturbances, or it may lead to coma which is unlikely to be reversed. Practically blood sample would be taken from Individuals would be screened to find out if the blood is infected after eight days. Then if they are, appropriate drugs would be administered to them and still monitored after the eighth day to determine if they are free of the disease. Death may before or after the first sampling time or even after administration of drugs.

By observing individuals over time (screening blood sample after each time step) and recording the disease status during a given sampling occasion, we can study the dynamic of any disease. The resulting time-series data is be represented as a matrix, X of 0, 1 and 2 s

indicating whether a given individual was uninfected (0), infected (1) or dead (2) . A detection history of X= 0, 0, 1, 1, 0 indicates that an individual was infected only during the third sampling occasion (x3) and recovered during the fifth sampling occasion (x5). However an individual could recover during a sampling occasion and be infected during the next sampling occasion due to multiple bites from the mosquito. Since death is absorbing, no additional data for that individual are necessary (that is even if we considered a sixth sampling period). An individual who received the plasmodium parasite could remain infected (carrier) due to acquired immunity in the individual. Hence we can consider an infection as a "success" though quite unpleasing to use and not infected as a "failure". Also an individual who is infected during our sampling period (t+1) does not depend on the fact that another individual was infected or not. In other words transitions are independent of each other.

We can obtain a time series data by sampling persons during time steps whereby the only possible states of an individual at a current time would be S={susceptible, infected}since sampling cannot be made on a dead individual.

With the observed data sequence $\{X^{(n)}\}$, we can easily find the transition frequency F_{ji} in the sequence by counting the number of transitions from state *j* to state *I* in one step and represent them as a one-step transition matrix for the sequence $\{X^{(n)}\}$as follows:

the population are not very appropriate for estimating the needed probabilities. An example is , the proportion of malaria on all deaths in children under five was reported as 33% (health facility data) with an estimated 14,000 deaths in children under five were attributable to malaria in 2008 (GHS, 2008). However the verbal autopsy component of the 2008 DHS household survey reports that malaria accounted for 43% of all deaths in children aged 29 days to 5 years, and that roughly half deaths in children under five occurred at home (PMI, 2011). This and many reasons may result in inconsistent report of the disease process in the population.

Let us consider an example. If we have 100 people in a community and observed 20 infected case due to any disease during a year, the probability of infection will be 20/100. However if each individual can be infected more than once in a year we could have as much as the total population being infected. So both the numerator and denominator in the equation (9) will change.

To avoid such inconsistency we can examine the distribution of the process. Since individuals make transitions at different continuous periods and a transition is also independent of the time (time homologous) the arrival of the process in a state follows a Poisson distribution with exponential interarrival times. We can therefore find the probability density function from the cumulative distribution function. The cumulative distribution function describes the probability of transition before time *t* and thus can be used to derive the probability of transition from the rate of transition. With a constant rate of transition, the probability p is given as;

The one-step state transition probability matrix P in this study is given as;

state 0 1 2

Where the rows (indexed as I = 0–2) represent the state of a process for a given individual at time n and the columns (indexed as j = 0–2) indicate the state of the process at the following time step n + 1. As death is an absorbing state, the probability of becoming susceptible or infected is zero. With the time step already defined, at any given time, an individual is assumed to be exactly one of these three states (although a perfect clinical state may not be detected), which is reflected by the fact that each of the rows always sum to one (5). Probabilities p_{01} p_{02} are also known as the discrete force of infection and the discrete mortality rate respectively and p_{10} and p_{12} are the force of recovery and the mortality rate due to malaria or case fatality rate respectively.

3.6.2 Determining the N-step transition matrix

We need to determine the n-step transition probabilities of the Markov process since the elements of the matrix P^n only provide information on the Markov chain at the nth time step; nothing can be inferred about the state of the process during any of the n – 1 time steps. However we must populate our transition matrix, P before finding P^n. To do this we calculate or transition probabilities p_{ij}'s from the given data.

These p_{ij}'s as in the study of diseases actually give us a measure the frequency of both disease occurrence and deaths from the disease and how fast the disease is occurring in a population. It also provides quantification for the prognostics of *P. Falciparum* parasitemia which is necessary to describe the severity of a disease to establish priorities for clinical services and public health programs to establish a baseline for natural history of the disease

so that as new treatment become available, the effects of these treatment can be compared with the expected outcome without them. Furthermore, if different types of therapy are available for a given disease, such as surgical procedures, to compare the effectiveness of the various type of therapy there is need for such quantitative means for expressing the prognosis in group receiving treatment (Gordis, 2009).

From (5) the total for each row equal 1 hence there only need to find two of the three parameters in each row of the transition matrix. For row one I computed for p_{01} and p_{02} in row two I calculated for p_{10} and p_{12}.

To obtain p_{01} we would use (11). The incidence rate (provided by GHS) usually gives a measure of the infection rate. Since everyone in the population of Ghana is at risk the incidence rate could be a good measure for the force of infection in the nation. However we could encounter some problems when using the incidence. We know that the infection rate lies in the range 1/14 to 1/6, that is the inverse of the times for infection, hence the incidence should be within this range. With this rate we can find the probability of infection during 8 days using (11).

To calculate the next parameter, p_{02} I used data from both GHS complimented by the data from WHO survey. We would calculate for the discrete rate of mortality for the year as the number of people who died as a result of all causes over the number of all people who were admitted in the WHO data. This was verified using he mortality rate of Ghana presented by CIA. Discrete mortality rate (probability of death) is defined using (9) as follows;

3.7 ESTIMATING DISEASE METRICS

Diagnostic and screening test permit the categorization of sick and healthy individual. Once a person is identified as having a disease, the question arises. “How can we characterize the natural history of the disease in quantitative term”? In Epidemiology some measures are to find the 5 year survival rate, and also the life tables. Markov modeling allows for computation of very important quantitative measures which can be used in characterizing malaria.

3.7.1 Estimation the First transition probability distributions

It is often desirable to also make probability statements about the number of transitions made by the process in going from state *I* to state *j* for the first time. The probability that a process initially in state I will be in state j after exactly n time steps is simply the elements of matrix P^n, denoted as p^n_{ij}. We can use these elements (p^n_{ij}) of the transition matrix P to calculate our disease metrics. We know from the definition of Markov process that the probability that a susceptible individual becomes infected after one time step is simply p_{01} and the probability that it remains susceptible is p_{00}. We define our initial time step as n and the interval n, n+1, n+2,…,n+m. Thus, the probability that a susceptible individual first becomes infected after two time steps is simply the probability that it remained susceptible for exactly one time step and then became infected during the next time step:

between the m – 1 and m time steps is given by:

3.7.2 Estimating The Expected Time To First Transition.

We can now use the probability distributions to determine the expected time to first infection (i.e. the average time to initial infection for a susceptible individual given that the individual does become infected) and the expected time to recovery (i.e. the average duration of infection). The expected time for an individual in a given state at time n to first enter another state, denoted

two transient states, 0 and 1 (susceptible and infected) and one absorbing state, 2 (dead), therefore:

of remaining susceptible and mortality respectively will be affected. Also an increase in recovery rates will affect the malaria mortality rate and the rate of remaining infected. Similarly, in terms of probabilities, p_{01} is relates p_{02} and p_{00} and also p_{10} relates p_{11} and p_{12}. From (5) we have;

$k1p_{01}=1-k2p_{02}-k3p_{00}$ and $k4p_{10}=1-k5p_{12}-k6p_{11}$

For some positive constants ki, where i=1, 2,…, 6. Hence a decrease in p_{01} as a result of a decrease of k1 will yield increase in p_{02} and p_{00} respectively. Similarly an increase in p_{10} as a result of an increase of k4 will yield decreases in p_{12} and p_{11} respectively (using 5).

With this idea we would simulate using the model and data whiles purposeful changes are made in the infection probability and recovery probability to find how sensitive the life expectancy and other computed disease metrics are towards theses changes.

CHAPTER 4

RESULTS

4.1 INTRODUCTION

In this chapter we discuss the source and nature of data. How the transition probability matrix was obtained is also shown. The characteristics of the model are then obtained. Next, the first step transition probabilities (to infection and recovery) are determined and their corresponding overall transition probabilities respected. These results are plotted against the sampling period (n+m) to determine the nature of the distribution. Finally, the life expectancy was found using the model and simulations were conducted to see the impact of the probabilities on the life expectancy.

4.2 SOURCE AND NATURE OF DATA

The data sources for this study are essentially secondary and retrospective. In general, three sets of data are used for this study. The primary sources of empirical information relates to the statistical returns of the Ghana Health Service, the World Health Organisation and Central Intelligence Agency. The choice of these data is predicated on the fact that representative information over time on the subject matter is gathered nationwide.

The GHS is responsible for generating and analyzing data emanating from the functioning of the health services through the Center for Health Information Management (CHIM). The mission of the Health Ministry is to contribute to socio-economic development and wealth creation by promoting health and vitality, ensuring access to quality health, population and nutrition services for all people living in Ghana and promoting the development of the local

health industry. The sector goal is to ensure a healthy and productive population that reproduces itself safely.

The medium Team Policies and priorities of the Health sector are to: Ensure mother and child are healthy through scaling up implementation of high impact and rapid delivery intervention; Promote good nutrition, food security and food safety; Prevent, control and manage communicable diseases such as HIV/AIDS, Malaria, Tuberculosis, Buruli Ulcer, Guinea Worm, Leishmaniasis, Lympahatic Filariasis, schistosomiasis; Collaborate effectively with relevant MDAs and stakeholders to improve housing, personal hygiene, environmental sanitation and access to potable water; Reduce risk factors associated with non communicable diseases such as tobacco and alcohol use, lake of exercise, poor eating habits, unsafe driving and stress; Strengthen referrals systems and clinical management of diseases as well as prevention and management of blindness and promotion of mental health; Strengthen surveillance and response to epidemics and emergencies and so on.

The World Health Organization (WHO) is a specialized agency of the United Nations (UN) that is concerned with international public health. It was established on 7 April 1948, with headquarters in Geneva, Switzerland and is a member of the United Nations Development Group. Its predecessor, the Health Organization, was an agency of the League of Nations.

The constitution of the World Health Organization had been signed by all 61 countries of the United Nations by 22 July 1946, with the first meeting of the World Health Assembly finishing on 24 July 1948. It incorporated the Office International d'Hygiène Publique and the League of Nations Health Organization. Since its creation, WHO has been responsible for playing a leading role in the eradication of smallpox. Its current priorities include

communicable diseases, in particular, HIV/AIDS, malaria and tuberculosis; the mitigation of the effects of non-communicable diseases; sexual and reproductive health, development, and ageing; nutrition, food security and healthy eating; substance abuse; and drive the development of reporting, publications, and networking. WHO is responsible for the World Health Report, a leading international publication on health, the worldwide World Health Survey, and World Health Day.

WHO identifies its role as one of six main objectives: providing leadership on matters critical to health and engaging in partnerships where joint action is needed; shaping the research agenda and stimulating the generation, translation and dissemination of valuable knowledge setting norms and standards and promoting and monitoring their implementation; articulating ethical and evidence-based policy options; providing technical support, catalysing change, and building sustainable institutional capacity; and monitoring the health situation and assessing health trends.

The World Factbook (also known as the CIA World Factbook) is a reference resource produced by the Central Intelligence Agency of the United States with almanac-style information about the countries of the world. The official paper copy version is available from the National Technical Information Service and the Government Printing Office. Other companies—such as Skyhorse Publishing—also print a paper edition. The Factbook is available in the form of a website, which is partially updated every week. It is also available for download for use off-line. It provides a two- to three-page summary of the demographics, geography, communications, government, economy, and military of 267 entities including U.S.-recognized countries, dependencies, and other areas in the world. In researching the *Factbook*, the CIA uses the sources some of which are: Antarctic Information Program

(National Science Foundation), Armed Forces Medical Intelligence Center (Department of Defense), Bureau of the Census (Department of Commerce), Bureau of Labor Statistics (Department of Labor) and so on. Other public and private sources are also consulted.

4.3 PRELIMINARY ANALYSIS OF DATA

4.3.1Malaria Admissions

The table for malaria admission cases in Ghana from 2000 to 2009 are shown in table A2.2 in the appendix. Generally inpatient cases of malaria for all age group appear to have increased from 2000 to 2008. Malaria accounted for over 28 % of total inpatient cases reported at various hospitals and in the case of children below 5, malaria accounted for over 36% of the total child inpatient cases. The latter shows that children were at higher risk within the time frame considered. A plot of the cases of admission against the years is shown in figure 4.1.

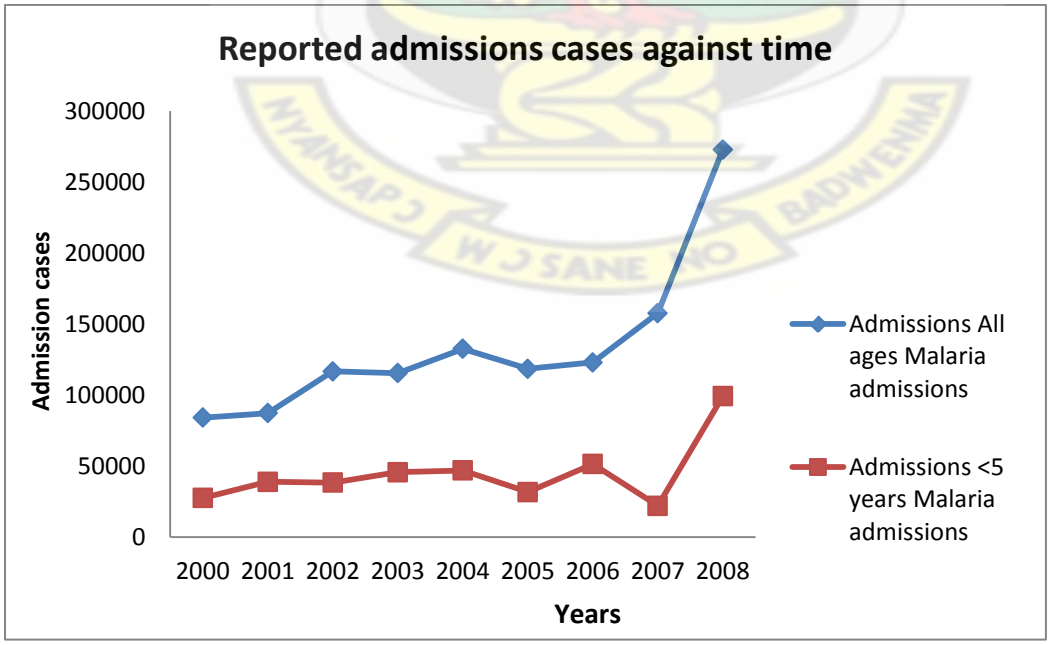


Figure 4.1: The plot of malaria admission cases from 2000 to 2009 in Ghana.

4.3.2 Malaria Incidence

Data from 1997 to 2008 was obtained for the incidence reported across the country (Table A2.1 in appendix 2). A quick look at the table shows an increasing trend of the number of cases for malaria. With the lowest incidence recorded in 1997 and gradually rises to the highest in 2008. However we could observe a sharp decline in 2007 which seems a bit unusual. Incidence is related to the population size so with increasing populations we expect the incidence to rise and whereby there is an intervention, we expect a mild rise in the number of incidence, hence the observed value in 2007 may be an outlier. The time series plot of the data is shown in figure 4.2 below.

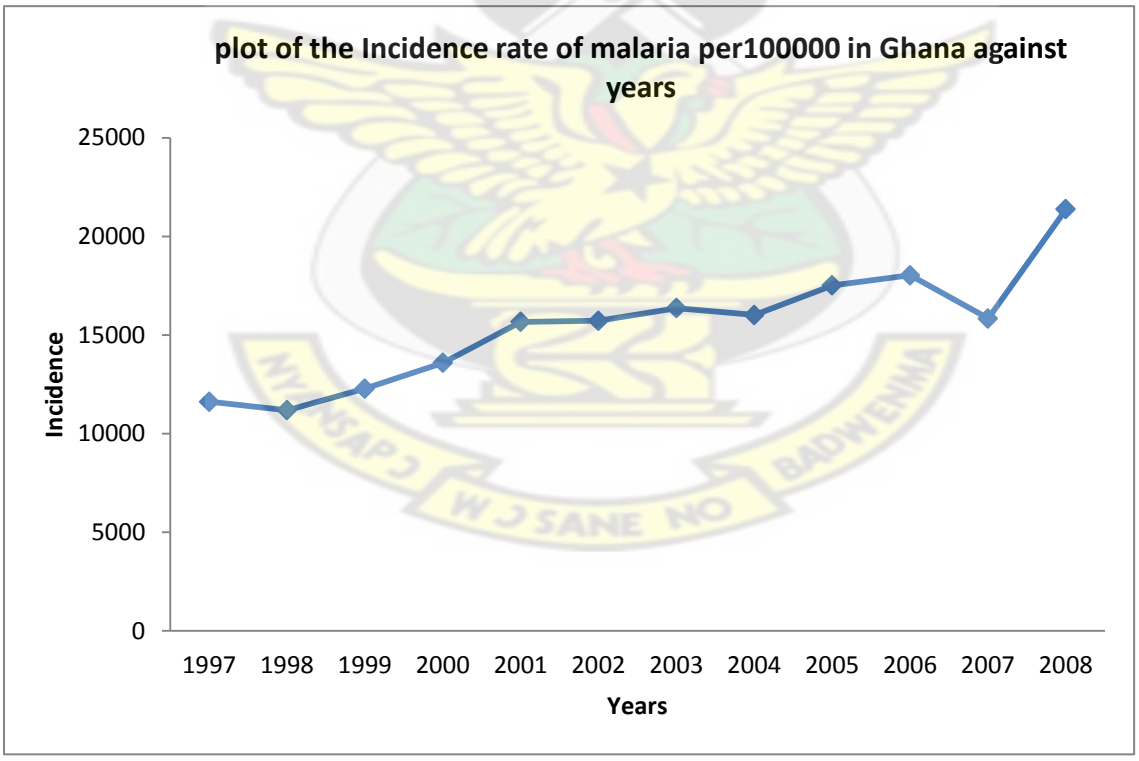


Figure 4.2: plot of the Incidence rate of malaria per100000 in Ghana against years

4.3.3 Malaria Mortalities

The reported mortalities cases due to all causes and due to malaria are presented in table A2.3 (in appendix 2). Mortality due to malaria in children accounted for over 21% of the total deaths recorded from 2000 to 2008. A plot of the mortalities due other causes and the mortality as a result of malaria infection for all ages from 2000-2008 is shown in figure 4.3 below.

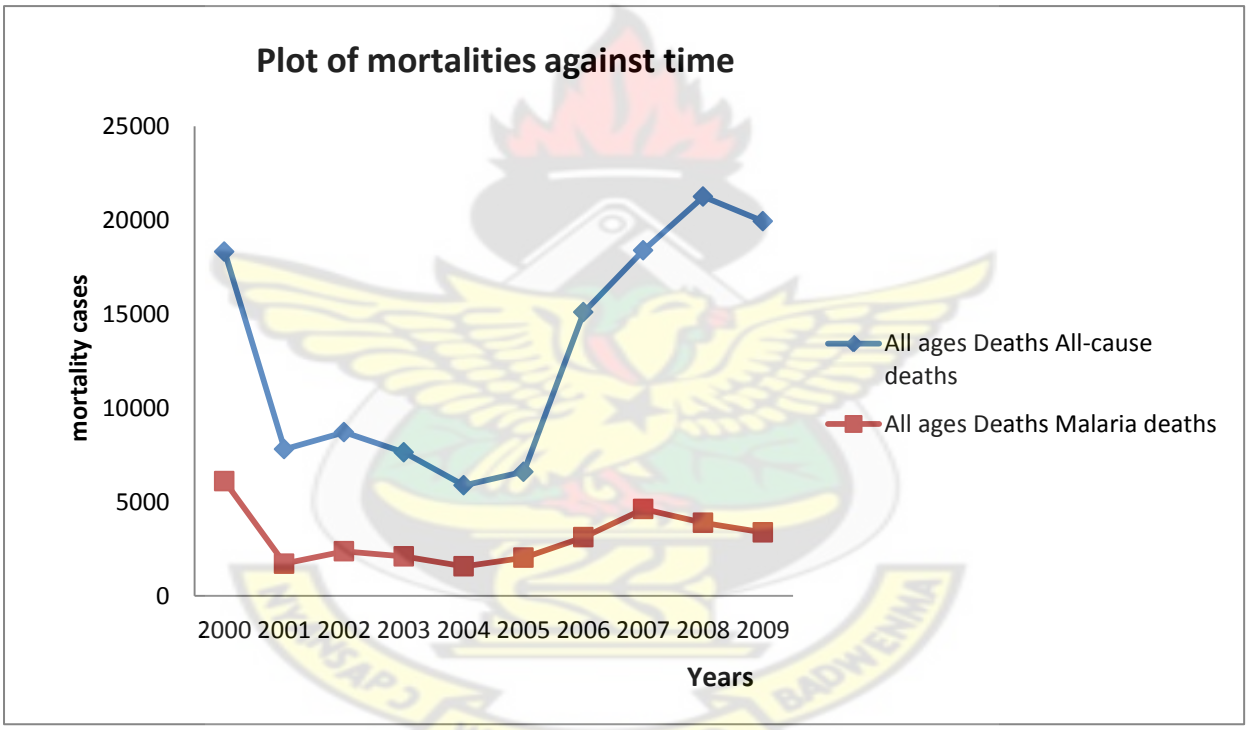


Figure 4.3: A plot of death cases due to malaria and other cases from 2000-2008 in Ghana.

The plot for all-causes mortality shows a rapid decline from 2000 to 2001, then a mild rise from 2001 to 2002. The mortality afterward gradually reduced from 2002 until the lowest reported case was observed in 2004. The trend then rapidly rose from 2004 until the highest case was observed in 2008 then the trend declined. Generally there tends to be an upward trend in the all-cause mortality cases. Similarly the malaria mortality shows a rapid decline

from 2000 to the lowest reported mortalities in 2001, then a mild rise from 2001 to 2002. The mortality afterward gradually reduced from 2002 until 2004 and then rapidly rose from 2004 until the highest case was observed in 2007 then the trend declined. Generally this tends to be an upward trend however very slowly comparing with the increasing population over the years.

The plots or the mortalities rates however give a better comparison of the two trends in the nation taking population into consideration (Figure 4.4).This shows downward trend in the proportion of mortalities for both all-cause and malaria cause. The trend of malaria deaths cases could be due to intervention policies and strategies and increasing government and non government expenditure in financing malaria control over the years. Table A2.6 shows the list of Governmental and external financing in malaria control.

A very important situation is to observe is the rate of children who die within each year compared to that of other ages. This is presented in table A2.4. A plot of these results shows that generally the malaria mortality rate in both age groups decline from 2000 to 2008. There rate at which children die over the years are higher as compared to those in all-ages. The trend for child mortality, however, seems to be narrowing to the mortality rate for all ages.

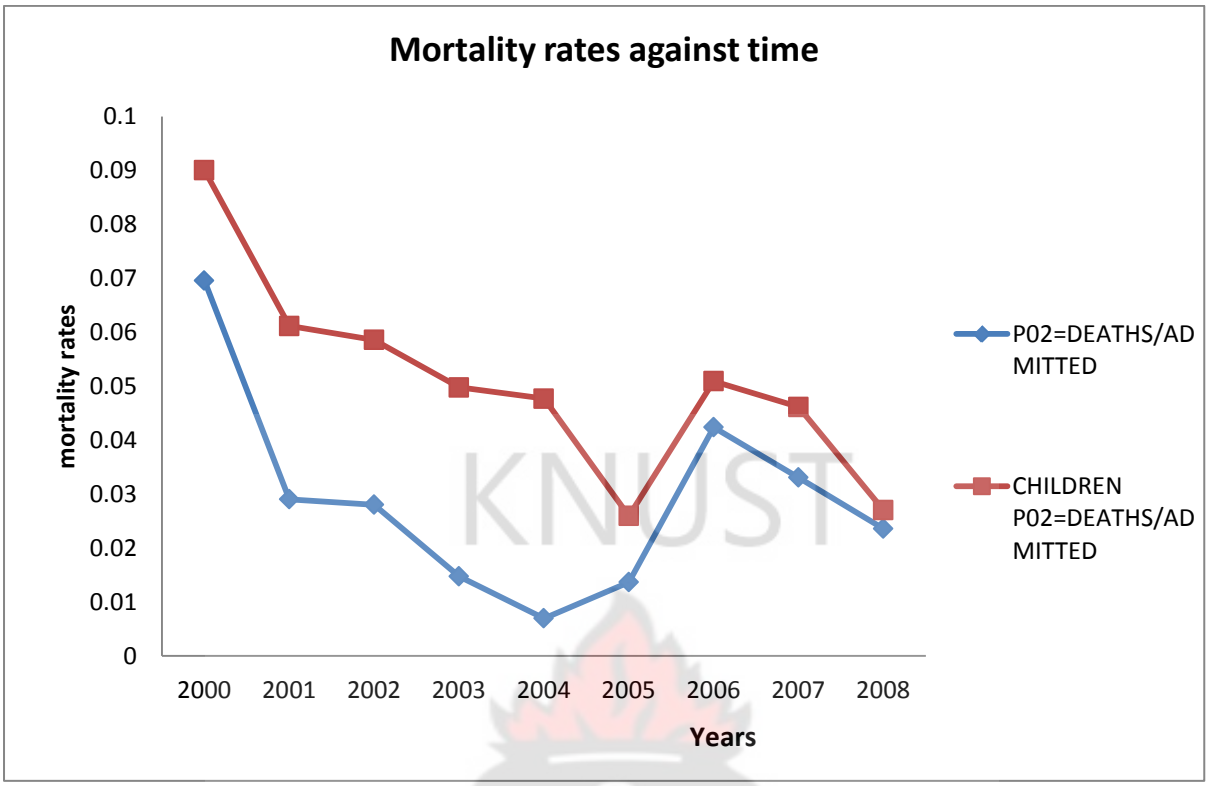


Figure 4.4: A plot of mortalities rate per person year for malaria and other cases from 2000-2008 in Ghana.

4.3.4 Recovery

The number of people who recovered from malaria infection after treatment was administered from within the time frame in Ghana for the various years are shown in table A2.4 in (appendix 2). Malaria cases accounted for 30% of total outpatient cases for all –groups and about 32.6% of total outpatient cases for children below 5 years. The plot of the cases of recovery in both groups is shown in the figure 4.5 below. Generally there tends to be an increase trend in cases for both all-age and below 5 age groups.

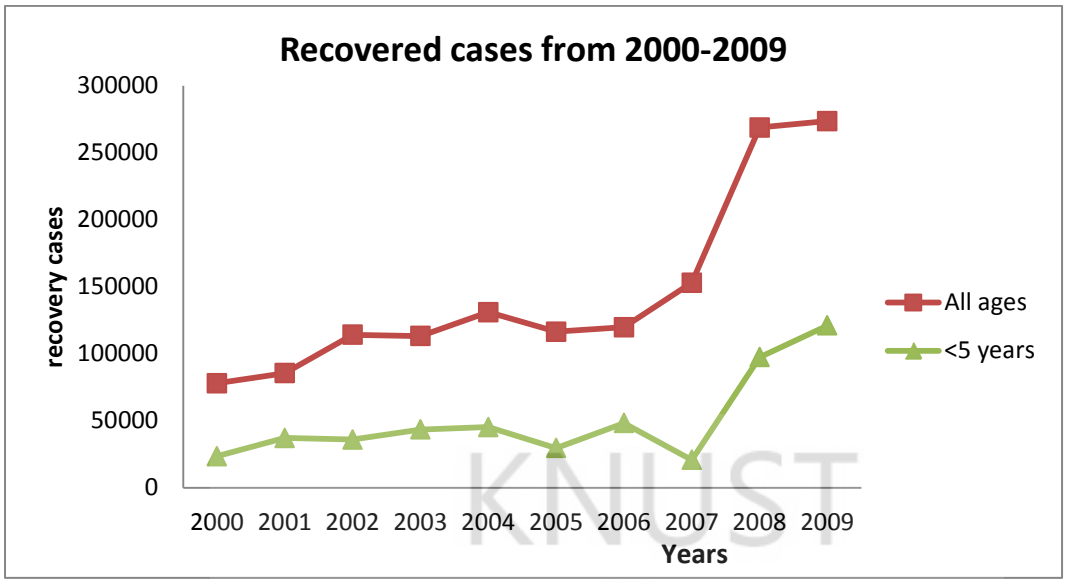


Figure 4.5: Recovered cases for all –age and below 5 groups from 2000-2009

However this may not be a true reflection considering the number of people who may have reported in each of the two groups, hence a need for the plot of the rates of recovery for all-ages and for below 5 years.

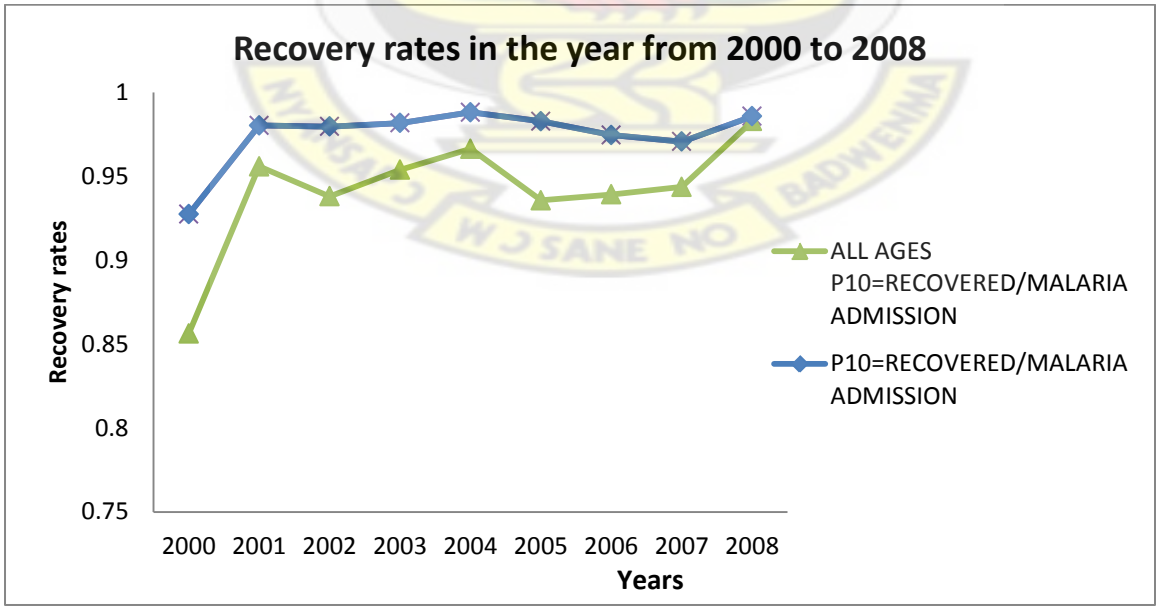


Figure 4.6: The recovery rates of all-ages and below 5 groups from 2000 to 2008

Figure 4.6 shows the proportion of people who recovered within each year from 2000 to 2008 in the county for all-ages and for children below 5 years. The proportion of people who recovered from malaria infection in a year in Ghana from 2000 to 2008 tends to be high in both groups. Noticeably was the increase in the proportion of people who recovered from 2000-2001.

4.4 MODEL BUILDING

4.4.1 Transition Probabilities

4.4.1.1 Infection probability.

Invasion of the blood by merozoites starts after the 5^{th} day after an infected mosquito hence the maximum rate of infection is given as 1/6. The incidence rate is a measure of the infection rate but may yield inaccurate result since not all individuals will be observed during the yearly period of study. The infection rate was calculated from the incidence rates (Table 4.1) using (9). Then we found the average from 2000 to 2008 as follows

4.1.1.2 Recovery probability.

Recovery probability was best estimated using (10) and monitored results of recovery in a research conducted by Nuguchi research institute (Table A2.4).

Continuous monitoring of the efficacy in the age group with the largest proportion of malaria problem- children under 5- pointed to the possibility of differences in drug responses between adults and children. Nonetheless, the efficacy of Artesunate – Amodiaquine has remained high and currently is over 90% after 28 days of treatment (MOHG, 2009).

The rate of recovery was estimated as follows;

population size as children who were admitted which is quite inappropriate since admission of children will be as a result of a severe condition hence such individuals have higher probabilities of dying. Estimation was done to explain the difference in p_{01} and p_{02}. Also with data from PMI (2011) report we have the number of children who reported to health facilities and the number who died. These parameters were used to estimate the malaria mortalities which also served as a guide as follows;

The Markov transition matrix for the S-I-D *P. falciparum* parasitemia is given as:

within an 8 day period. Which accounts for over 40000 below 5 deaths in an estimated population of 3.7 million children (alive) in Ghana are like to die after a year (Table A2.3b).

Also the rate from I to D in the model shows us that about 40 people (below 5 years) are likely to die out of 100000 under 5 children as a result of malaria infection during 8 days. This reflects about 15000 child deaths as a result of malaria and about 35% proportion of all death in the population in a year.

The probability of infection also indicates that about 7 out 10 children will be infected if exposed to mosquito bite in a period of 8 days. This is also a reflection of the high infection rate of 0.166 which implies that one would have symptoms of disease starting from 6 days in Ghana. This is confirmed by World Malaria Report (2011), which indicates that transmission in West Africa and specifically Ghana is above 100 cases per 1000 population (Figure A3.2, in appendices).

4.5 CLASSIFICATION OF THE MODEL STATES

By our classification of states in section 3.3 we see that S is reachable from I and I from S hence susceptible(S) and infectious (I) states of the model are communicating class.

Also since there is a positive probability of returning to S if we start from S and likewise a positive probability of returning to I, state S and I also exhibit a recurrent property. In other words if we start the process from a susceptible state we are always likely to return to a susceptible state. Same can be said of state I. Also D is recurrent although absorbing since remaining in an absorbing state is considered as a visit during the time.

If we are to move from S to D there are two possible paths: S-D and S-I-D. Since these two paths have an odd interval the chain is aperiodic. Since all states of the Markov chain are recurrent and aperiodic the chain is ergordic.

4.6 THE STEADY STATES PROBABILITIES

The probability,

the probability distribution of the initial state. In other words the transition probability into a susceptible state after a long time equals zero. Practically this means that in the long run, if malaria persists, then even if the person recovers that person will quickly move out of the recovery state. Similarly a person will quickly move out of a susceptible state. However,

Figure 4.8: plot of the probability distributions for an infection of susceptible individuals and recovery of infected individuals against the sampling times.

The figure above (Fig. 4.8) shows the probability that a susceptible individual first becomes infected with *P. falciparum* (solid black line) and the probability that an infected individual first recovers (dash black line) in exactly m steps (calculated in 8 days time steps).

We observe that the probability distributions decrease monotonically and faster for the probability distribution of first infection. This implies that currently it takes longer time to recovery than to be infected in Ghana.

4.8 OVERALL PROBABILITIES OF INFECTION AND RECOVERY

The infinite sum of the first transition probability distribution yields the overall probability of first infection which is in shown the table A2.5. The overall probabilities or first infection and first recovery are used to find the duration of infection and the duration of recovery respectively. However it may take a long time for the distribution to converge to the overall probability. The actual (converging) probability of infection was calculated by making use of (18) which is included in the matlab script. The results for the overall probabilities of infection and recovery are shown in Table A2.5 in appendices. This permitted for the estimation of the mean time for first infection and mean recovery is calculated using (19). The plot of the cummulative transition probability distribution is shown in figure 4.9.

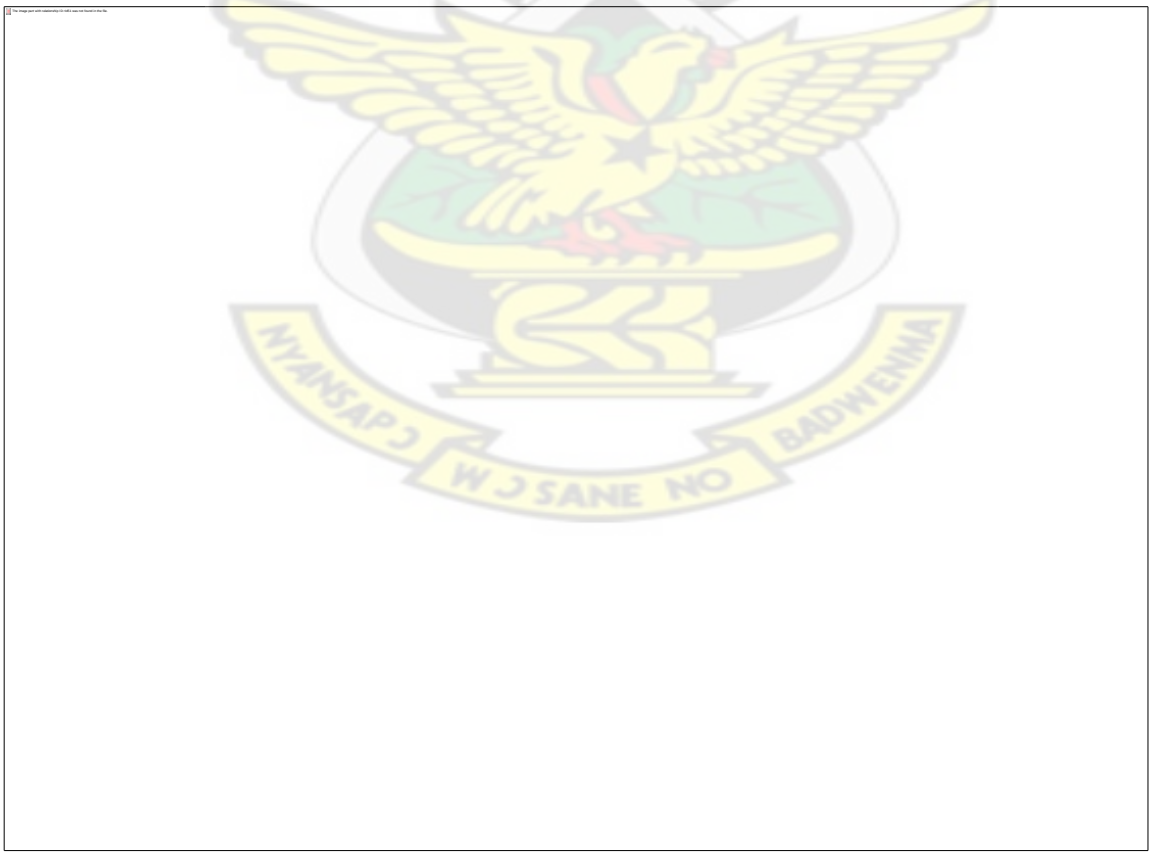


Figure 4.9: The plot of the cumulated probability distributions for an infection of susceptible individuals and recovery of infected individuals against the sampling times.

Figure 4.9 shows the cumulative probability that a susceptible individual has been infected (solid black line) and an infected individual has recovered (dash black line) over the duration of the study.

Interestingly, the overall probability of infection for a susceptible individual was very high (0.9994). The long-term transition probabilities for infection took almost 11 days to stabilize. Given that a susceptible individual becomes infected, the expected time to first infection is given as 11days (table 4.1).

Another interesting result is the overall probability of recovery for a susceptible individual which was also found to be high (0.9991). This result was used to calculate the expected time for recovery or the duration of the disease which is another useful quantity in characterizing a disease. Given that an infected individual recovers, the expected time to first recovery is given as approximately17 days (table 4.1). This was also consistent with a work by Wiser (1999) who stated that the duration of a typical *P. falciparum* malaria symptom if untreated ranges from 2-3 weeks.

Table 4.1: Estimated disease metrics for humans exposed to *P. falciparum* using Ghana data from 2000 to 2008.

Metric	Value
Probability of infection	0.9994
Probability of recovery	0.9991
Expected time to infection(steps)	1.4278
Expected time to recovery(steps)	2.1259
Expected time to infection (days)	11.4222
Expected time to recovery (days)	17.0068

4.9 LIFE EXPECTANCIES

The computed result of our fundamental matrix given by $(I-Q)^{-1}$ defined in section 3.6 are shown in appendix 3. This allows us to find the time it takes for an individual to enter an absorbing state. The first element of row one is the total expected time a person is expected to spend in the susceptible state before being absorbed and the second element is the time a person is expected to spend in the infectious state before being absorbed. The summation of these two gives us the total time for a person who was in a susceptible state at the onset of the survey period to be absorbed. Similarly the summation of the second row gives the life expectancy starting from an infectious state.

The life expectancies using (21); expected time to death from the onset of the experiment) for susceptible and infected individuals were found to be approximately 55 years (Table 4.2) in each situation. This is not the life expectancy at birth. Rather, since the sampling unit was on

individuals below 5 years adding the average age of this group will yield the life expectancy at birth.The result shows a similarly life expectancies for people who were previously susceptible and those who were previously infected. This explains the fact that people in susceptible states are not so different from those who are infected. Practically this means that recovery rates in Ghana are high with regards to current treatment and other recovery related factors. However the life expectancy for uninfected individuals are slightly higher than that for infected persons from the onset of the survey. This shows the fact that; the life expectancy of people in the nation could be increased if preventive measures are taken rather than depending on treatment. Although this points out to an interesting fact for policy making, we can however not make conclusive deduction on this slight difference until further test is conducted.

Table 4.2: The calculated life expectancy from 2000 to 2008 using first order Markov model.

Time	Life expectancy 2000-2008
Susceptible (sampling steps)	2525.4
Infected (sampling steps)	2525.4
Susceptible (Years)	55.1994
Infected (Years)	55.1990

`

4.10 SIMULATION

In order to determine the impact of infection probabilities and also of the recovery probabilities we conducted simulations using the Markov model and obtained estimates from the Ghana simultaneously.

We reduced both the probability of infection and the probability of recovery at 1% intervals until a 10% reduction in each situation. We cannot an explicit solution to k1 and k2, however changing k1 will produce a change in both k1 and k2. We therefore decreased the infection probability also by 5% at intervals of 1% of the existing (prevailing) probability and found the outcome for the life expectancy using the model. The mortality rate was kept constant so that logical interpretation could be inferred. The results for simulation are indicated in Table 4.3.

Table 4.3: Results of the simulated life expectancies by decreasing infection probability.

	Life expectancy (years)	
decrease in infection probability	**Susceptible**	**Infected**
0.0000	55.1994	55.1990
0.0035	55.2011	55.2007
0.0280	55.2132	55.2127
0.0315	55.2149	55.2145
0.0350	55.2167	55.2162

The plot for the simulated life expectancies as a result of decreasing the infection probability and keeping mortality probability constant is shown in figure 4.10. This showed increases in the life expectancy with decreases n infection probabilities.

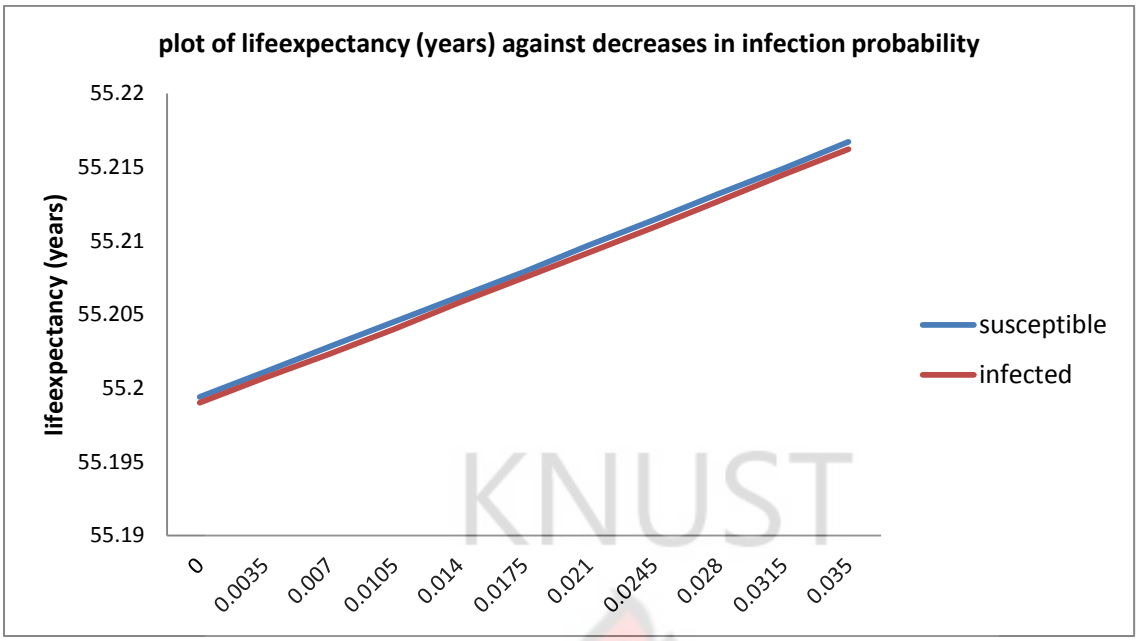


Figure 4.10: Simulated results for life expectancies of susceptible and infected individuals against decreases in infection probability.

Next, to determine the impact of recovery probabilities, we increased the recovery probability by 5% and then observed the outcome for the life expectancy using the model. It was logical to expect a decrease in the probability of death due to malaria if the recovery rates are increased so we decreased p_{12} by some constant as well.

Table 4.4: Results of the simulated life expectancies and transition times by increasing recovery probability.

	Life expectancy (years)		recovery duration
increase in recovery probability	**Susceptible**	**Infected**	
0	55.1994	55.1990	17.0068
0.0047	55.5959	55.5956	16.8387
0.0094	55.9948	55.9948	16.6739
0.0141	56.3963	56.3965	16.5123
0.0188	56.8004	56.8008	16.3539
0.0235	57.2070	57.2077	16.1984

The plot for the three way simulated life expectancies as a result of increasing the recovery probability and reducing malaria mortality probability and the probability of remaining susceptible is shown in figure 4.11. This also shows as increase in the life expectancy. Also interesting to note was the decreases in the duration of recovery (Table 4.4).

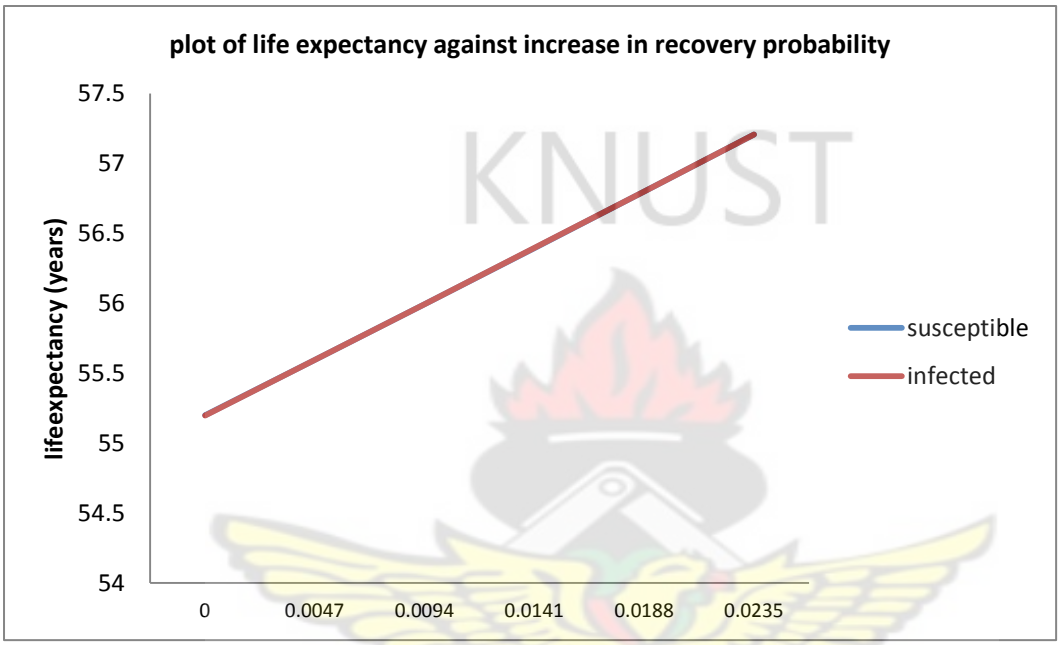


Figure 4.11: Simulated results for life expectancies of susceptible and infected individuals against increases in recovery probability.

4.11 INTERPRETATION OF RESULTS

The measures of occurrence from the disease showed how fast the disease is occurring in Ghana from 2000 to 2008. These quantifications of the *P. falciparum* parasitemia are necessary to describe the severity of a disease to establish priorities for clinical services and public health programs. Example the malaria mortality probability within the 8 days sampling time (0.0004) though looks small contributes to over of 15000 child deaths as a result of malaria and about 33% proportion of all death in the population (approximately 3.7 million below age 5 children in Ghana) in a year. This is worrying and calls for more anti-malaria

policies, funding of drugs, increase in hospital facilities and services and education of public on the prevention of the disease.

The estimation of 8 days (time step) probabilities of infection and recovery in terms of seasonal probabilities of infection and recovery for malaria was possible by the application of Markov models. The model also permitted the computation of expected time to infection and duration of infection and state-dependent life expectancy from the onset of the survey period. Hence the use the Markov model provided a convenient and accessible approach for computation of disease metrics that summarize relevant aspects of the host-pathogen interaction. Quantities such as expected time to infection and seasonal probabilities of infection and recovery have clearly interpretable meanings with respect to management efforts and can help set thresholds of acceptable disease persistence.

The expected time to a first infection was given as 11 days with a very high probability (0.9994). It means that a susceptible individual has very high likelihood to be infected if exposed to mosquitoes during an average of 11 days. This is a reflection of the endemicity of the disease in Ghana. Hence IPT's using antimalarial drugs like Sulfadoxine-pyrimethamine (SP) and Chemoprophylaxis can be administered before 11 days as a measure to prevent occurrence of the disease.

The expected time to a first recovery of the duration of non severe *P. falciparum* was found to be 17 days with a very high probability (0.9991). It means that individuals in Ghana recover from non severe *P. falciparum* malaria infection at an average time of 17 days with a high probability.

We can use these estimated values of the model as baselines for the natural history, so that as new preventive methods become available, the effects on life expectancy, recovery times and infection time using these using these treatment can be compared with the expected outcome without them. We can compare to see the impact that purposeful changes in disease management can have on groups receiving treatment, which can also be used to compare the effectiveness of the various types of treatments.

Preventive methods are methods which are used to avoid infection from occurring. These consist of preventing mosquito bites, controlling mosquito breeding, the use of anti- malarial drugs and so on. Prevention in this study involves the reduction of probability of infection. By the outcome of our simulation, using the Markov model for Ghana, we notice that the life expectancy increases when the probability of infection is decreased, even when mortality rates are kept constant. The outcome of the two way simulation by decreasing the probability of recovery by 5% of the prevailing probability at 1% intervals and keeping the discrete mortality rate constant, showed a final gain in life expectancy in years of 0.0173 (Table 4.3).

This proves the fact that preventive measures which lead to reducing infection probabilities can help increase the average life span of people living in Ghana. Therefore prevention using any of these methods, namely: IRS to kill adult mosquitoes; the use of ITN's; the use of repellants; control of mosquito breeding; use of IPT and so on, will in tend save many lives which will result in an increase in the average life expectancy of Ghanaians.

Disease management deals with diagnosis and early treatment of disease and other related factors like nutrition, rest and so on, which are needed for recovery. This implies proper care for infected persons and in this study involves the probability of recovery and the malaria

mortality probability. Also from the outcome of the simulation, we also notice that the life expectancies increases when doth the probability of recovery is increased and the probability of malaria mortality is decreased. An increase to 5% in the probability of recovery by 1% intervals and a reduction by some percentage k5 in the probability of malaria mortality, showed a final gain of 2 years in life expectancy (Table 4.4). Although the increments were rather arbitrarily chosen if treatment methods and malaria death are monitored carefully, we can find k2 which should permit in predicting the life expectancies using the model.

The result thus proves the fact that improvements in disease management procedures which increase the recovery probabilities can help increase the average life of people living in Ghana. Another interesting finding is on the shortening of the duration of recovery from prevailing 17 days to almost 16 days after simulation. This clearly shows that increases in the recovery probability speeds up recovery time and could prevent much complication as a result of the disease.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

The measures of occurrence from the disease showed how fast the disease is occurring in Ghana from 2000 to 2008. These quantifications of the *P. falciparum* parasitemia transmission are necessary to describe the severity of a disease to establish priorities for clinical services and public health programs.

The estimation of 8 days (time step) probabilities of infection and recovery in terms of seasonal probabilities of infection and recovery for malaria was possible by the application of Markov models. The model also permitted the computation of expected time to infection, the duration of recovery and the state-dependent life expectancy from the onset of the survey period which have clear interpretable meanings with respect to management efforts and can help set thresholds of acceptable disease persistence.

The expected time to a first infection was given as 11 days with a very high probability (0.9994). This means that susceptible individuals have very high likelihood of being infected if exposed to mosquitoes during an average of 11 days. The expected time to a first recovery or the duration of non severe *P. falciparum* was found to be 17 days with very high probability (0.9991).

The outcome of our simulation by decreasing the probability of infection by 5% of the prevailing probability at 1% intervals showed a final gain in life expectancy in years of 0.0173 (Table 4.3). This proves the fact that prevention methods which lead to decreasing the infection probabilities can help increase the average life span of people living in Ghana. Therefore prevention using any of these methods, namely: IRS to kill adult mosquitoes; the use of ITN's; the use of repellants; control mosquito breeding; use of IPT's and so on, will in tend save many lives which will result in an increase in the life expectancy.

Also an increase to 5% in the probability of recovery by 1% intervals and a reduction in probability of malaria mortality, showed a final gain of 2 years in life expectancy (Table 4.4).

The result thus proves the fact that improvements in disease management conditions which lead to increases in the recovery probability can help increase the average life span of people living in Ghana. The shortening of the duration of recovery from prevailing 17 days to almost 16 days due to increases in recovery probability show that complications as a result of the disease can also be avoided with this procedure.

The state-dependent life expectancy from the onset of the survey period using the model was found to be approximately 55 years for both individuals who were had the disease(infectious state) and those who did not (susceptible state). This result is consistent with results from the Ghana health service which gives an average life expectancy at birth from 2003 to 2009 as 59 years.

Although the increments and decreases in simulation were rather arbitrarily chosen if monitoring is carefully done, can help determine amount of increase and decrease which can be used to predict the life expectancies using this model. The Markov model did not account for differences among classes of individuals below 5 (that is below 2 and above 2 years),

neither did it account for differences in the season (wet and dry season) yet separate models can be developed for each life stage and season to compare disease dynamics across an individuals' lifetime and across seasons respectively if adequate data from Ghana is obtained.

Finally, this research was intended to demonstrate the usefulness of Markov chain modelling in the study of diseases and I envision this research can serve as an entry point into the extensive literature and potential applications of Markov chains in malaria and disease modelling.

5.2 RECOMMENDATIONS

We recommend that measures be put in place by the government of Ghana to implement malaria management procedures which increase the recovery probability or recovery rates since this leads to an increase in the life expectancy of the Ghanaian and also reduces the duration of the disease.

We also recommend that preventive measures which lead to the reduction in the infection probability be adapted since this also increases the life expectancy of Ghanaians. Finally we recommend that intermittent preventive treatment (IPT) using anti malaria drugs like Sulfadoxine-pyrimethamine (SP), Chemoprophylaxis can be administered before 11 days to interrupt occurrence of malaria.

REFERENCES

Adams I., Darko D. and Accorsi S. (2004). "Malaria: a burden explored", Bulletin of Health Information, pp. 28-33.

Appiagyei A, Osei S. (2011). "REACH Against Malaria", www.reachghana.org., 4th March 2012.

Breman J. G., Egan A., and Keusch G. T. (2001). "The intolerable burden of malaria: a new look at the numbers", The American Society of Tropical Medicine and Hygiene, Vol. 64 (Suppl): iv–vii.

Cancré N.,Tall A., Rogier C. , Faye J., Sarr O., Trape J., Spiegel A. and Bois F.,(2000). "Analysis of an Epidemiologic Model of *Plasmodium falciparum* Malaria Infection in Ndiop, Senegal", American Journal of Epidemiology Volume 152, No. 8, pp. 760-770.

Ciecka, J. and Skoog, G. R. (2003). "The Markov (Increment-Decrement) Model of Labor Force Activity: New Results Beyond Worklife Expectancies", Journal of Legal Economics, Vol. 11, No. 1, pp. 1-22.

Cross C. (2004). "Immunity to malaria", www.welcometrust.org., 9th March, 2012.

Commenges, D. (1999). "Multi-state models in epidemiology", Lifetime Data Analysis, Vol. 5, pp. 315–327.

Cooper B. and Lipsitch M. (2004). “The analysis of hospital infection data using hidden Markov models”. Biostatistics, Vol. 5, pp. 223–237.

Deltour I., Richardson S. and Hesran J. (1999). “Stochastic Algorithms for Markov Models Estimation with Intermittent Missing Data”, International Biometric Society, Vol. 55, No. 2,pp. 565-573.

Gani J. and Jerwood D. (1971). “Markov Chain Methods in Chain Binomial Epidemic Models”, International Biometric Society, Vol. 27, No. 3, pp. 591-603

Gordis L. (2009). “Epidemiology”, 4th Edtion, Saunders Elsevier Inc., Philadelphia. pp. 37-41, 59-70.

Health2spread (2012). “Symptoms of malaria”, www.health2spread.com/symptoms-of-malaria.php, 4th April, 2012.

Hiligsmann M., Ethgen O., Bruyère O., Richy F. Gathon H. J and Reginster J.Y(2008). “Development and Validation of a Markov Microsimulation Model for the Economic Evaluation of Treatments in Osteoporosis”, International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Vol. 12, Issue 5, pp. 687–696.

Hillier F. S. and. Lieberman G. J. (2001). "Introduction to Operations Research" , 7th edition, McGraw-Hill companies, inc., 1221 avenue of the Americas, New York, Ny., 10020. pp. 802-812.

Holding, P. A. and Snow R.(2001). "Impact of Plasmodium falciparum malaria on performance and learning: review of the evidence". The American Society of Tropical Medicine and Hygiene, Vol. 64, pp. 68–75.

Hosgood G. L. (2002). "Markov Models to Estimate and Describe Survival Time and Experience in Cohorts with High Euthanasia Frequency", Elsevier Science, pp. xii-xiii.

International Association for Medical Assistance to Travelers (2011). "World malaria risk chart: Geographical distribution of principal malaria vectors, Plasmodium falciparum drug resistant areas, and guidelines for suppressive medication by country". IAMAT, Canada. pp. 1-5.

Jackson C. H., Sharples L. D., Thompson S. G., Duffy S. W., Couto E. (2003). "Multistate Markov models for disease progression with classification error", Journal of the Royal Statistical Society: Series D (The Statistician),Vol. 52, No. 2, pp. 193–209.

Jones, R. M., (2008), "Experimental evaluation of a Markov model of contaminant transport in indoor environments with application to tuberculosis transmission in commercial passenger aircraft", University of California, Berkeley, pp.341-342.

Kekre H. B. (1978). “Computers & Electrical Engineering: A finite waiting room queueing model with multiple servers having Markovian interruptions and its application to computer communications”, Elsevier. pp.x-xii.

Kekre H.B. and Saxena C. L. (1977). “Three-state Markov model of speech on telephone lines and optimal utilization of communication systems by TASI technique”, http://dx.doi.org/10.1016/0045-7906(77)90033-7, 14th November, 2011.

Killeen P. R.(2010). “Markov model of smoking cessation”, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3176613, 14th November, 2011.

Lekone, P. E. and Finkenst¨adt B. F. (2006). “Statistical Inference in a Stochastic Epidemic SEIR Model with Control Intervention: Ebola as a Case Study” International Biometric Society Vol. 62, pp. 1170–1177.

Macdonald G. (1957). “The Epidemiology and control of Malaria”, Oxford University Press, London, pp.1-9.

Maire N. (2008). “Stochastic simulation models of Plasmodium falciparum malaria epidemiology and control”, Trop. Med. Int. Health, Vol. 7, pp. 421–428.

Malaria Consortium (2012). “Intermittent Preventive Treatment”, http://www.malariaconsortium.org/page.php?id=114, 11th May, 2012.

Mannan H. R., Knuiman M. and Hobbs M. (2009). “Using a Markov simulation model to assess the impact of changing trends in coronary heart disease incidence on requirements for coronary artery revascularization procedures in Western Australia”, http://www.biomedcentral.com/1471-2261/10/2, 14th November, 2011.

Mba C. J. and Aboh I. K. (2009). “Prevalence and Management of Malaria in Ghana: A Case Study of Volta Region”, African Population Studies, Vol. 22 n°1, pp. 156-158.

Ministry of Health- Ghana Health Services (2010). “The Health Sector in Ghana: Facts and Figures” http://www.ghanahealthservice.org/malaria_control.php, 20th January, 2012.

Ministry of Health- Ghana Health Services (2008). “The Health Sector in Ghana: Facts and Figures” http://www.ghanahealthservice.org/malaria_control.php, 14th February, 2010.

Ministry of Health- Ghana Health Services (2007). “The Health Sector in Ghana: Facts and Figures”, http://www.ghanahealthservice.org/malaria_control.php, 14th February, 2010.

Nassar R., Fan L., Too J. R. (1981). “A stochastic treatment of unimolecular reactions in an unsteady state continuous flow system”, Chemical Engineering Science, Vol. 36, Issue 8, pp. 1307-1317.

Ngwa G. A., Shu W. S. (1999). “A Mathematical Model for Endemic Malaria with Variable Human and Mosquito Populations”, The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy, pp. 2.

Patten S. B. and Lee R. C. (2005). “Describing the Longitudinal Course of Major Depression using Markov Models: Data Integration across Three National Surveys”. http://www.pophealthmetrics.com/content/3/1/11, 12th November, 2011.

President’s Malaria Initiative (2011). “Malaria Operational Plan- FY 2011(Year 4)”, U.S. Global Malaria Coordinator, pp. 11.

Richard A., Richardson S., Maccario J. (1992). “A three-state Markov model of *Plasmodium falciparum* parasitemia”, Mathematical Biosciences,Vol. 117, Issues 1–2, pp. 283–300.

Sesso H. D, Chen R. S., L’Italien G. J., Lapuerta P., Lee W. C., Glynn R. J. (2002.), “Blood Pressure Lowering and Life Expectancy Based on a Markov Model of Cardiovascular Events” , International Society for Pharmacoeconomics and Outcomes, Research 7th International Meeting, Arlington, Va.

Sheehan N. A. (2000),“Application of Markov Chain Monte Carlo Methods to Genetic Analyses on Complex Pedigrees”, International Statistical Institute (ISI), Vol. 68, No. 1, pp. 83.

Too J. R., Fan L. T., Nassar R. (1983). “Markov chain models of complex chemical reactions in continuous flow reactors”, Elsevier Ltd.Vol. 7, Issue 1, pp. 1–12.

United Nations Children’s Fund (2007). “UNICEF Ghana Fact Sheet: Malaria” http://www.unicef.org/wcaro/WCARO_Ghana_Factsheet_malaria.pdf, 12th January, 2012.

Venkataramanan L. and Sigworth F. J. (2002). “Applying Hidden Markov Models to the Analysis of Single Ion Channel Activity”, Biophysical Journal Vol. 82, pp.1930–1942.

Verma B. L., Ray S. K., Srivastava R. N. (1983). “Research Article: A stochastic model of malaria transition rates from longitudinal data: considering the risk of ‘lost to follow-up’”. J Epidemiol Community Health, Vol. 37, pp.153-156.

Webb J. L. A. (2008).“Humanity’s Burden: A Global History of Malaria Excerpt”, Cambridge University Press, pp. 3-9.

World Health Organization (2010). “Guidelines for the treatment of malaria”, second edition. Technical report, Geneva, pp. 1- 194.

World Health Organization (2011).“World Malaria Report: Roll Back Malaria”, Technical report, Geneva.

World Health Organization (2006). “World Malaria Report: Roll Back Malaria”, Technical report, Geneva.

Zipkin E. F., Jennelle C. S. and Cooch E. G. (2010). "Methods in Ecology and Evolution", British Ecological Society, Vol. 1 pp. 192–198.

APPENDICES

Appendix 1CALCULATION OF DISEASE METRICS

The following script computes for the disease metrics
%% 1 Create the matrix P that contains the transition probabilities for susceptible (state 0),

infected (state 1), and dead (state 2)

%%Define the transition probabilities

p01 = 0.70; %MOHG

p02 = 0.00039; % who average

p00 = 1-p01-p02

p10 = 0.47; %Who

p12 = 0.0004;

p11 = 1-p10-p12

p20 = 0

p21 = 0

p22 = 1

P = [p00,p01,p02; p10,p11,p12; p20,p21,p22];

%2. Calculate steady states

syms x0 x1 x2

_etn2='x1=x0*p01+x1*p11+x2*p21';

_etn3='x2=x0*p02+x1*p12+x2*p22';

```
etn4= 'x0+x1+x2=1';

[a,b,c]=solve(etn1,etn3,etn4,x0,x1,x2);

a=subs(a)

b=subs(b)

c=subs(c)

%3. Calculate metrics of interest

  %Probability of moving from susceptible (0) to infected (1) over course of

%study period and expected time to first transition

Pr01 = p01 / (1-p00);

Ex01 = 1 / (1-p00);

%%Probability of moving from infected (1) to susceptible (0) over course of

%%study period and expected time to first transition

Pr10 = p10 / (1-p11);

Ex10 = 1 / (1-p11);

%%Calculate the probability of first infection at each time step for a susceptible individual

%%Create the number of time steps, n

n=1:30;

%%Create a vector with zero values (as place holders) that
```

```
%%is the length of n for f01(n) as defined in eqn 4

f01n = zeros(length(n));

 %%Write a loop that uses eqn 15 to calculate the probability

%% of first infection at each time step

for i=1:length(f01n)

    f01n(i) = p01 * p00^(i-1);

end

%%print the values

f01=f01n(:,1)

 f10n = zeros(length(n));

%% of first recovery at each time step

for i=1:length(f01n)

    f10n(i) = p10 * p11^(i-1);

end

f10=f10n(:,1)

 pinf=cumsum(f01)

prec=cumsum(f10)
```

%%4. Find the life expectancy (expected time to absorption) for susceptible and infected states.

```
%%Create the matrix Q with the transition probabilities

Q = P(1:2,1:2)

 %%Define an identity matrix called Identity

Id = eye(size(Q));

%%Find the inverse of the Identity matrix minus Q using

IminusQ = Id - Q;

Fmatrix= inv(IminusQ);

 %% Create a 2 by 1 vector with the value 1 in each element

one = ones(length(Q),1);

 %% Use eqn 9 to find life expectancies

Lifeexpectancy =inv(IminusQ)*one

 b=Lifeexpectancy*8/366
```

Appendix 2: TABLES

Table A2.1: The incidence rates in Ghana

Year	1997	1998	1999	2000	2001	2002	2003
incidence rate per 100000	11615	11191	12284	13590	15667	15726	16363
incidence rate	0.11615	0.11191	0.12284	0.1359	0.15667	0.15726	0.16363

Year	2004	2005	2006	2007	2008	Average
incidence rate per 100000	16015	17513	18032	15833	21376	
incidence rate	0.16015	0.17513	0.18032	0.15833	0.21376	0.1667*

Source: GHS. NB * is the average from 2000- 2008

Table A2.2: Cases of Admission for various age groups from 2000 to 2008

	All ages		<5 years		All ages	<5 years
	All-cause	**Malaria**	**All-cause**	**Malaria**	**All-ages**	**Malaria**
YEAR	**admissions**	**Admissions**	**Admissions**	**admissions**	**proportion**	**proportion**
2000	263269	84091	98507	27478	0.319411	0.278945
2001	268598	87236	102397	38911	0.324783	0.380001
2002	310793	116600	100895	38340	0.375169	0.379999
2003	517566	115401	120126	45648	0.222969	0.380001
2004	844091	132566	123384	46886	0.157052	0.380001
2005	483038	118449	174522	31644	0.245217	0.181318
2006	356000	122928	97860	51407	0.345303	0.525312
2007	556036	157628	113952	22019	0.283485	0.19323
2008	900242	272802	181427	99217	0.303032	0.54687
Average	499959.2	134189	123674.4	44616.67	0.286269	0.360631

Source: World Health Organization(WHO).

Table A2.3: Mortality summaries for children below 5 from 2000 to 2008

Year	All-cause	Malaria	Mortality rates	p_{02} per sampling time	proportion of malaria deaths	Yearly Case fatality rates (%)
2000	8872	3952	0.09007	0.00197	0.21569	-
2001	6265	1717	0.06118	0.00134	0.21998	-
2002	5913	2376	0.05861	0.00128	0.27266	-
2003	5983	2103	0.04981	0.00109	0.27540	-
2004	5887	1575	0.04771	0.00104	0.27501	2.7
2005	4532	2037	0.02597	0.00057	0.30817	2.1
2006	4988	3125	0.05097	0.00111	0.20692	2.7
2007	5263	1241	0.04619	0.00101	0.06746	2.4
2008	4907	1697	0.02705	0.00059	0.07987	1.6
Average	5845.6	2202.6	0.05084	0.00111	0.21347	1.9**

Source: WHO, **GHS (with current treatment)

Table A2.3b

Year	under 5 mortality rates per 1000 alive		mortality rate at sampling time	p_{02}
1988	155.00	0.155	0.00068	0.00061
1993	119.00	0.119	0.00052	0.00047
1998	107.00	0.107	0.00047	0.00042
2003	111.00	0.111	0.00049	0.00044
2008	80.00	0.080	0.00035	0.00032
average from 2003-2008	99.33	0.099	0.00043	0.00039

Source: GHS

Table A2.4: Recovery cases of various age groups from 2000 to 2008 in Ghana

RECOVERED	<5 years	Proportion of recovered	Recovery rates
	Malaria		
YEAR	RECOVERY		
2000	23526	0.856176	0.0693
2001	37194	0.955874	0.1114
2002	35964	0.938028	0.0993
2003	43545	0.95393	0.1099
2004	45311	0.966408	0.1212
2005	29607	0.935628	0.0980
2006	48282	0.939211	0.1000
2007	20778	0.94364	0.1027
2008	97520	0.982896	0.1453
Average	42414.11	0.94131	0.1063

Source: World Health Organization(WHO).

NB. For probability of infection during sampling period we made an assumption of equal proportion over the year.

Table A2.5: The Estimated first transition probabilities and overall probabilities for infection and recovery in Ghana

Sampling time	probability first of infection, f01	probability of first recovery, f10	total probability of infection, pinf	total probability of infection, prec
1	0.7000	0.4700	0.7000	0.4700
2	0.2097	0.2489	0.9097	0.7189
3	0.0628	0.1318	0.9726	0.8507
4	0.0188	0.0698	0.9914	0.9205
5	0.0056	0.0370	0.9970	0.9575
6	0.0017	0.0196	0.9987	0.9771
7	0.0005	0.0104	0.9992	0.9875
8	0.0002	0.0055	0.9994	0.9930
9	0.0000	0.0029	0.9994	0.9959
10	0.0000	0.0015	0.9994	0.9974
11	0.0000	0.0008	0.9994	0.9982
12	0.0000	0.0004	0.9994	0.9987
13	0.0000	0.0002	0.9994	0.9989
14	0.0000	0.0001	0.9994	0.9990
15	0.0000	0.0001	0.9994	0.9991

Table A2.6: List of Governmental and external financing in malaria control.

	2006	2007	2008	2009
Other*	0	0	1000000	0
USAID/PMI	0	5000000	16900000	17300000
UNITED NATIONS CHILDREN'S FUND	0	1200000	1200000	939300
WHO	0	100000	200000	290000
World Bank	0	5000000	4000000	1283389
Global Fund	21762030	9269310	10544980	18363180
Government**	1229000	2980000	3235000	8700000

Source WHO

NB:*Bilaterial: DFI, JICA; and EU, UN agencies, etc. Government expenditure may not include cost at sub-national level and cost related to health systems, human resources, etc.**

Appendix 3: Diagrams showing risk and distribution of malaria

FigureA3.1: Map of malaria risk areas. Source: World Health Organization, 2005.

Figure A3.2: The distribution of confirmed cases (per 1000 population) in Ghana

Source: WHO (World Malaria Report 2011).