KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY,

KUMASI

QUANTITATIVE MICROBIAL RISK ASSESSMENT: AN INTEGRATED PROBABILISTIC MODELLING OF HUMAN EXPOSURE TO *NOROVIRUS* IN WASTEWATER

By

EMMANUEL de-GRAFT JOHNSON OWUSU-ANSAH

(MPHIL APPLIED MATHEMATICS)

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE

OF

DOCTOR OF PHILOSOPHY IN MATHEMATICAL STATISTICS

SEPTEMBER, 2015

DECLARATION

I hereby declare that this submission is my own work toward the PhD and that, to the best of my knowledge, it contains no 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 acknowledgement has been made in the text.

Emmanuel. D.J. Owusu-Ansah ---------- ----------

Signature Date

Certified by: ---------- ----------

Prof. S.K. Amponsah Signature Date

(Supervisor)

Certified by: ---------- ----------

Prof. Tine Hald (Co-supervisor) Signature Date

Certified by: ---------- ----------

Prof. R.C. Abaidoo (Co-supervisor) Signature Date

Certified by: ---------- ----------

Prof. S.K Amponsah Signature Date

Head of Department

DEDICATION

This thesis is dedicated to the nameless tens of thousands of innocent children, women and men, witnesses, reporters, volunteers and ordinary men and women who work hard to make a living and making this planet a better place to live.

ACKNOWLEDGEMENT

Prima facea, I am grateful to the Almighty God for the good health and wellbeing that were necessary to complete this thesis and His unlimited knowledge and strength which saw me through.

I wish to express my sincere thanks to my supervisors Prof. S.K. Amponsah, Dr. Tine Hald, and Prof. R.C Abaidoo for their continuous feedback, guidance, patients and their encouragement as well as relevant suggestions and opinions through all the stages of this thesis. I am also grateful to my faculty members at the section of Risk Assessment and Epidemiology, National Food Institute, Technical University of Denmark(DTU), not forgetting my co-postgraduate fellows at University of Copenhagen (Ofred, Li Kang and Tra-Min Phu).

I am also grateful to my colleagues on the various work packages of the SaWaFo Project, especially Angelina Sampson for her diligent work as well as managing our social media group. Special thanks go to Dr. Darko (Head of Department, Chemistry, KNUST) for his constant effort of encouragement and putting smiles on our faces through whatsapp images, which kept me smiling in times of difficulties.

This research was fully funded by DANIDA project Safe Water for Food (SaWaFo). My sincere thanks goes to Prof Anders Permin (Deputy Director, National Food Inst., DTU) and to Dr. Torben Wilde Schou, Loni Skov Madsen and Enrico Ulisse Remigi all of DHI, Denmark.

I am greatly indebted to Rejoice Ametepeh, my wonderful SaWaFo project coordinator for all of her support and timely provision of all logistics for the numerous travels during the study. Lastly, to my family, friends and love ones who continuously pray and encourage me. May the good Lord richly bless you all.

TABLE OF CONTENTS

Table of Contents

LIST OF TABLES

LIST OF FIGURES

SYMBOLS

B	Diarrheal Burden of Diseases in Ghana	$person^{-1}year^{-1}$
C_Z	Salad consumption	$gday^{-1}$
n	Frequency of consumption	$day(year^{-1})$
V_P	Volume of irrigation water caught by product	ml/g
	Lettuce	ml/g
	Cabbage	ml/g
k	Pathogen kinetic decay constant	day^{-1}
t	Time for withholding or irrigation cessation	$days$
w	Post-harvest/Food Preparation Washing for Virus reduction	$\log_{10} units$
r	Cryptosporidium spp response parameter for infectivity	
α	Loss of full immunity	$year^{-1}$
γ	Loss of partial immunity	$year^{-1}$
μ_s	Duration of incubation for *Norovirus*	$days$
ρ	Duration of asymptomatic infection for *Norovirus*	$days$
μ_a	Duration of symptoms for *Norovirus*	$days$
$\mu(a)$	Parametric mean dose for *Norovirus*	
ω,η	Dose response parameters for illness given infection for *Norovirus*	
A	Life Expectancy	$year$

P	The infection probability for subjects with disaggregated dose
P_{ill}	Probability of illness
$P_{\inf} = P_I$	Probability of infection
$P_{ill\|\inf}$	Probability of illness given infection
N	Population
E	Total Exposure
λ	Force of infection
τ	Inflation factor

ABSTRACT

The call for models applying quantitative data of pathogens that are of interest to replace the otherwise commonly applied models using fecal indicator conversion ratio has gained prominence, challenges of analytical studies on virus enumeration (genome copies or particles) have contributed further to the low availability of data in Quantitative Microbial Risk Assessment (QMRA) modelling. In this thesis, a probabilistic stochastic model was developed to respond to the call for virus of interest based models. Quantitative data on genome copies of *Norovirus* and oocyste of *Cryptosporidium* spp. were applied in a QMRA model. The model was extended to include an induced immunity for Dose Response Incidence (DRI) of illness reduction in individual and population exposures, five different scenarios were modelled for *Norovirus* based on the epidemiological understanding of immunity within an individual and *Norovirus* transmission dynamics. A third model was developed to measure the uncertainty of compliance and reliability of wastewater effluent with integrated policy standards. The probabilistic QMRA model revealed fecal indicator ratio conversion method underestimated the Disability Adjusted Life Years (DALYs) with more than two (2) orders of magnitude and were confirmed using the *Cryptosporidium* spp. data. For immunity extended DRI models, results shows, illness incidence is much reduced when both dose-dependent and immunity are integrated into risk assessment models. Integration of immunity only into DRI model also performed better than dose-DRI model only. It was also revealed that, irrespective of the epidemiology transmission dynamics within the population, DRI models predictions were similar and dose-immunity DRI model was better predictor. Finally, the analyses of compliance and reliability of wastewater effluent measurements revealed that results from wastewater effluents which met the policy standard values, in some cases could not meet the compliance level needed for effluent discharge. A chart was developed for the various wastewater treatment effluent discharge parameters for easy comparison with effluent discharges.

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

To secure and increase crop production, irrigation has become a principal water use in many countries, which otherwise traditionally have been depending on rain water(Anonymous, 2014; Jiménez, 2006). This drive of irrigating farm products for food production has resulted in the use of wastewater in areas where water is scarce. Wastewater is undoubtedly a major source of human pathogens as well as a key source of nutrients for plants growth. Wastewater is known to serve as both an ingredient of foodstuffs and an independent vehicle for human exposure to microbiological hazards (Food and Agricultural Organization, 2003). The practice of using wastewater for irrigation is dated as far back as before the 14th century where farmers in China used it in combination with human excreta and animal excrement as fertilizers (Drechsel et al., 2009).

It has been estimated that, wastewater irrigation alone covers 4-6 million hectares (Jimenez and Asano, 2008) and in the near future, for 4 out of 5 cities, urban agriculture workers will use wastewater for peri-urban farming activities (Lundqvist and Raschid-Sally, 2013). Middle East countries are known for high usage of treated wastewater. More than 60% of treated wastewater is used for agricultural irrigation in

Israel alone (Lawhon and Schwartz, 2006). Other countries included are Egypt, Jordan and countries along the Nile (Hegazy, 2013; Drechsel and Scott, 2009). In Ghana, several studies indicate the abundant use of wastewater for irrigation (Amoah et al., 2007; Seidu et al., 2008; Silverman et al., 2013; Amoah et al., 2005; State, 2010; Mok and Hamilton, 2014; Drechsel and Scott, 2009; Karavarsamis and Hamilton, 2010; Haas, et al., 2014; Scheierling et al., 2010)

As a global phenomenon, wastewater usage has both advantages and disadvantages. It is noted that 1000 cubic meters of municipal wastewater used to irrigate one hectare can supply approximately 16 – 62 kg total nitrogen and 4 – 24 kg phosphorus of soil nutrient (Qadir et al., 2010) making it a substitute for artificial fertilizers and essential for chemical-free farming. However, among the disadvantages, the most important is the risk of pathogen occurrence (WHO, 2006), which may results human disease and hence makes it a concern for public health. In dealing with such public health issues, a more holistic method of assessing the risk of pathogen occurrence instead of diseases has gained ground among scientists(Drechsel et al., 2007).

1.1.1 Quantitative Modelling of Hazards from Wastewater

Quantitative Risk Assessment modelling is a growing area of research attracting a great number of academicians at both regional and global levels, due to the mathematical and computational intensity required to describe the physical,

chemical, financial, economic and biological phenomena, etc. Over the last three decades, one of the powerful branches of quantitative risk assessment modelling that has seen enormous growth in theoretical background and applications for quantification of health risk is Quantitative Microbial Risk Assessment (QMRA), which primarily focuses on identification and quantification of the potential adverse health risk (probability of infection/illness) associated with exposure of individuals or populations to hazardous materials (Heidinger, 2009). Technically, it provides a framework for modelling a food chain process and hence estimates an associated probable risk especially for food borne infection from consumption of contaminated food (Duarte and Nauta, 2015). Fundamentally, risk assessment combines the knowledge and nature of the hazards, and characterises the risk through robust models based on theoretical statistics (Haas et al., 2014). After the first publications of human health risks associated with exposures (Dudley et al., 1976; Fuhs, 1975), the field has grown exponentially into all areas (van Lieverloo et al., 2007; An et al., 2007; Labite et al., 2010; Armstrong, 2005) and has been making inroads into other field of studies such as environmental assessment (Ward, 1993; Verdonck, 2001), drug delivery (Heylings, 2011) etc.

Despite the great advancements of estimating risk and in particular QMRA, risk estimate of many waterborne and food borne diseases are based on the ratio conversion method (Silverman et al., 2013), i.e. where the occurrence of common fecal bacteria like E. coli are used to estimate the pathogen contamination (e.g.

Norovirus) level of water and so the human health risk, by assuming a fixed ratio between the occurrence of indicator bacteria and the pathogen of interest. Likewise, there has also been an apparently rare interest in accounting for induced immunity when estimating the risk of illness due to infection (Havelaar and Swart, 2014). These limitations calls for a more unrelenting and resolute research to dive into modelling approaches where the pathogen of interest is in focus and the translation of risk estimates to illness accounts for induced immunity with a parsimonious model such as the fractional Poison dose response, where probabilistic modelling offers a way.

1.2 Problem Statement

Traditionally, modelling of illness or diseases has to do with mathematical epidemiology; nevertheless, such a method is unable to always provide sufficient sensitivity to measure health risk directly with the availability of human health data as well as not pathogenic specific. Distinctively, prediction of relative risk of diseases for future scenarios, in order to evaluate efficacy and efficiency of alternative mitigation processes in diseases modelling rest entirety on the use of QMRA. However, in QMRA modelling for estimating water safety, studies typically use fecal indicator ratio conversion for describing virus concentration to express the relationship between the occurrence of fecal indicators (typical *E. coli*) and the pathogen of interest (bacteria or virus). This approach was used in the WHO guideline for the Safe Use of Wastewater, Grey water and Excreta(WHO, 2006), and has been adopted by all

subsequent studies (Barker et al., 2014; Fiona Barker et al., 2013; Mara et al., 2010; Mara and Sleigh, 2010; Travis et al., 2010; Mok and Hamilton, 2014; Ackerson and Awuah, 2012) . This method has raised concern from practitioners as well. Silverman et al., (2013) noted "while the ratio of *Norovirus* (NoV) GII to *E. coli* or thermotolerant coliform is likely to differ over place and time and may include animal fecal sources as well as environmental sources and reservoirs, it is an important finding that the current assumption of 0.1 – 1 *Norovirus* particles per 10^5 *E. coli* would underestimate virus dose with exposure to wastewater and surface water sample". This call of concern was supported by Mok and Hamilton (2014) who remarked "if standard pathogen concentrations are to be used effectively, there should be a move away from indicator species such as *E. coli* toward the pathogens of interest such as viruses''. Moreover, risk assessment basically ends with either predicting annual risk of illness or the annual Diseases Adjusted Life Years (DALYs), but most research studies applying the illness dose-dependent of infection fail to include and characterize the various dynamics of immunity response of individuals within the population.

It is against this background that this study is carried out to model through all the stages from using pathogen of interest to predict the incidence of illness based on epidemiological inclusion of induced immunity for humans' exposure to the use of wastewater.

1.3 Research Objectives

The overall aim of this study was to develop improved risk models for humans exposed to hazardous substances from wastewater used for unrestricted irrigation of vegetables. The specific objectives were:

(i) To develop a probabilistic quantitative risk assessment model with genome copies of Norovirus and fecal indicator ratio conversion for dose estimation.

(ii) To develop an integrated induced immunity dose response model for *Norovirus* for modelling probability of illness incidence reduction.

(iii) To develop statistical measurements modelling for quantification of uncertainty and compliance level of low quality water effluent for measuring wastewater discharge.

1.4 Methodology

The study presents and discusses statistical approaches for probabilistic modelling of gastroenteritis following human's exposure to wastewater through the consumption of vegetables subjected to unrestricted irrigation. Probabilistic modelling formulation with an exposure to a single infectious pathogen are constructed, then induced immunity will be incorporated as a results of continuous exposure to pathogens and combine epidemiological data into the modelling to predict the incidence of illness either in an individual or in the population. Additionally, incidence of illness models with induced immunity are formulated, introducing transmission dynamics within the

age structured epidemiological modelling of *Norovirus*. Relevant epidemiological data are incorporated into the Dose Response (DR) models to provide simulation results for estimating risks and incidence of illness among the population. Finally, based on data from two wastewater ponds in Kumasi, Ghana, statistical measurements for assessing reliability and compliance given policy standard values for such measurements are also formulated.

1.5 Scope of the Study

The study is limited to quantitative microbial risk assessment of *Norovirus* and *Cryptosporidium* (for validation) in Ghana specifically on wastewater used for unrestricted irrigation. Ghana falls among countries with a high volume of wastewater usage for peri-urban and urban agriculture and hence exposed consumers to the risk associated with wastewater(Amoah et al., 2005). Though other pathogens are presents in wastewater, the study takes special interest in Norovirus and in some instances *Cryptosporidium* spp., as *Norovirus* has seen some special interest as a better replacement for *rotavirus* which was used in WHO policy guideline for use of Wastewater and Greywater for farming. *Cryptosporidium* spp. are interesting because they are quite robust parasites (protozoa) that can survive for a considerable time in the environment.

1.6 Significance of the Study

This study will rally round the need to move towards a more realistic model approach and minimize the assumptions under which conditions pertaining to the risk estimates are made by applying data on the pathogen of interest to minimize and if possible eliminate the fecal indicator ratio conversion method. It will facilitate the move to combine epidemiological data into risk assessment by shedding more light on the transmission of probability of infection to illness stage. The study will further provide insight into probabilistic modelling with quantification of experts' opinion in formulating probability distribution inclusion in risk assessment. Finally it will add more knowledge to the existing literature on risk assessment estimate procedures particularly, towards the move to pathogen of interest modelling, risk estimate of illness with induced immunity modelling and measurements of wastewater discharges with policy standards as well as provide the platform for researchers to extend the frontiers of knowledge on risk estimates of diseases.

1.7 Organization of the Thesis

The thesis is organized into different chapters as follows: Chapter 1 deals with the background of the study, problem statement and objectives of the study, the methodology, scope and significance of the study are also put forward. Chapter 2 presents a literature review of the QMRA procedure, the deterministic and stochastic

approach in QMRA probabilistic modelling, detection methods of hazards as well as quantification of uncertainty in QMRA. In chapter 3 put forward various statistical backgrounds in probabilistic modelling. The chapter presents the background of probability distributions, estimation of parameters of the various distributions used in the models and the modelling of pathogen densities. Chapter 4 presents the probabilistic modelling of using quantitative measures of the virus of interest amidst insufficient data and inclusion of expert's judgment as well as experts opinions for estimating does, as opposed to the conventional method of using the conversion ratio based the occurrence of fecal indicators. Simulations for predicting risk of illness is carried out, sensitivity results also examined and discussed. Chapter 5 is devoted to estimating probability of illness accounting for temporary acquired induced immunity by including epidemiological data. In Chapter 6, measuring parameters that characterise effluent discharge of treated and untreated wastewater, which form the basis of hazardous (pathogenic) substances, are put forward. Chapter 7, the final chapter, presents the summary, conclusions and recommendations of the study.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this chapter, the study put forward relevant literature on wastewater use and the quantitative microbial risk assessment method used to estimate risk associated either directly or indirectly exposure to wastewater. This chapter also focuses on water scarcities which lead to the use of wastewater, the merits and demerits of wastewater usage for food production as well as the extent of wastewater usage in Ghana. The quantitative microbial risk assessment approach, which is being used for estimation of risk in addition to uncertainty quantification associated with illness given infection as a result of exposure to wastewater are also captured. The chapter ends with the description of epidemiological models estimating induced immunity for quantifying illness.

2.2 Water Scarcity

One of the most challenging natural gifts of nature affecting the existence of mankind in the 21st century and beyond is shortage of freshwater. Freshwater is estimated to make up a very small fraction of all water on this planet earth. Nearly seventy percent (70%) of the world is covered by water, while only 2.5 percent of it is fresh for human and animal usage. Yet, just one percent (1%) of freshwater is easily accessible, and most of it is trapped in glaciers and snowfields. Only 0.007% of the

planet's water is available to fuel and feed its seven billion people (National Geographic Report, 2015).

Observation from National Geographic Report (2015) stated, :"Due to geography, climate, engineering, regulation, and competition for resources, some regions seem relatively flush with freshwater, while others face drought and debilitating pollution''. In most of the developing countries, clean water is either hard to come by or a commodity that requires laborious work or significant currency to obtain. Since the debate of freshwater scarcity is not settling (Rooijen and Rooijen, 2008), this still remains as an abstract concept to many and a stark reality to others

World Wildlife Fund Report (2015) estimations suggest that, globally, 1.2 billion people, live in areas of physical scarcity of water, an additional 500 million people are approaching this situation, whereas another 1.6 billion people face economic water shortage [1](Fig.2.1). Surprisingly, up to 70% of the worlds freshwater are used for agricultural purposes, of which 60% is specifically for irrigation of crops alone. These have all contributed to increasing freshwater withdrawals from aquifers and hence increasing the demand for water for all sphere of life.

[1] where countries lack the necessary infrastructure to take water from rivers and aquifers

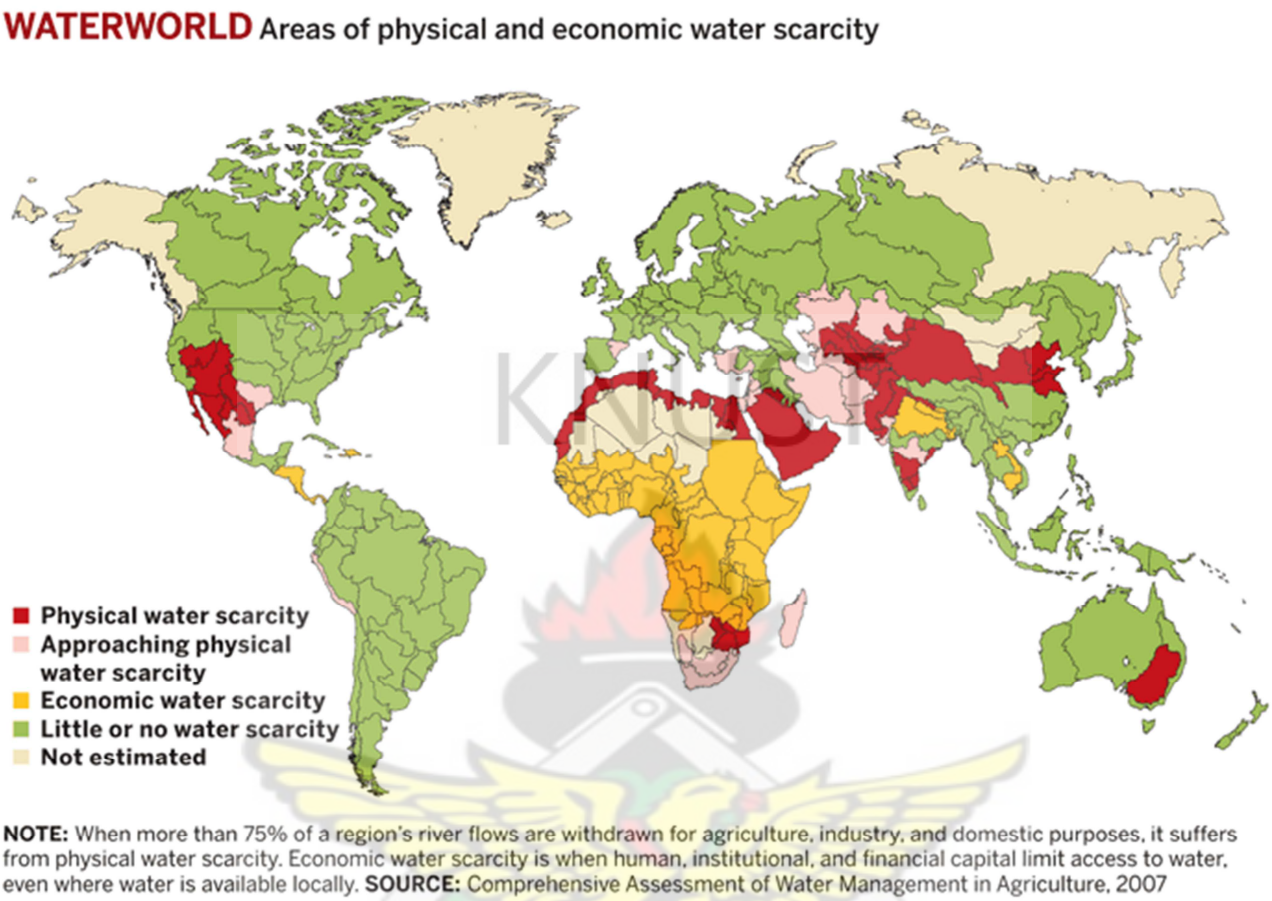


Figure 2.1: Areas of Physical and Economical Water Scarcity, (Sources WMA, 2007)

2.3 Irrigation Practices and Wastewater Usage

There is no complete global inventory on the extent to which wastewater is used to irrigate land, mostly due to lack of heterogeneous data and the fear that countries have about disclosing information; economic penalties can be imposed if produce is found to have been irrigated with low-quality water (Jiménez, 2006). Farm irrigation as a practice has been going on for millennia, however, from the 20th century and

beyond, irrigation has become a principal water use feature in many countries to serve as a guarantee for increased crop production (Anonymous, 2014).

Due to the high dependence of water for food production, wastewater has been used as an alternative to irrigate many farms (Jiménez, 2006). In developing countries, the increase in wastewater usage has been productive, as millions of small-scale farmers in urban and peri-urban areas depend on wastewater or polluted water sources to irrigate high-value edible crops for urban markets (Qadir et al., 2010).

Wastewater also serves as a source of both water and nutrients needed for plant growth and reduces the cost of using fertilizer on plants(Jimenez and Asano, 2008; Jiménez, 2006). It permits higher crop yields, year-round production, and enlarges the range of crops that can be irrigated, particularly in (but not limited to) arid and semi-arid areas. Wastewater recycles organic matter and other nutrients to soils. Again, it improves soil properties (soil fertility and texture) and offers additional benefits such as greater income generation from cultivation and marketing of high-value crops, which contribute to improved nutrition and guarantee employment opportunities for farmers.

However, several studies point out (Ackerson and Awuah, 2012; Amoah et al., 2005; An et al., 2007; Barker, 2014; Barker et al., 2014; Crabtree et al., 1997; Scheierling and Mara, 2010; Qadir et al., 2010; Mara et al., 2007; Shuval et al., 1997; Petterson, 2002; Seidu et al., 2008; Mara and Sleigh, 2010; Petterson, 2001; Silverman et al., 2013; Lundqvist and Raschid-Sally, 2013; Mara et al., 2010), that as much as wastewater has

merits of its usage, there is also associated public health concern with it; the presence of pollutants in wastewater can reduce soil productivity, toxic to plants or humans consuming crops and the pathogens contained in wastewater can cause health problems for humans as well as animals.

2.4 Wastewater Usage in Ghana

Urbanization and population growth come along with increasing demand for sanitation infrastructure. In Ghana, and other parts of Sub-Saharan Africa, sanitation infrastructure within the urban areas is inadequate lagging behind to the population growth rate. According to (Keraita et al., 2002) less than 5 percent of city dwellers are linked to infrequently functional sewage systems and sewage treatment plants for wastewater treatment is also limited, less than 8 percent of wastewater generated are being treated and discharge. Most untreated wastewater end up in streams and other water bodies which serve as sources for irrigation water in many urban and peri-urban areas and constitute the only available surface water for irrigation which guarantees all year access to water for farming(Drechsel et al., 2009; Keraita et al., 2002).

In Ghana, the use of polluted water for vegetable farming is more widespread in the more populated cities where safe water is scarce (Seidu et al., 2008). This makes it among countries with largest volume of raw wastewater usage worldwide (Fig. 2.2), and represents one of the centres used mainly in the study of wastewater used for

agriculture in developing countries (Fung, 2011; Ackerson and Awuah, 2012; Labite et al., 2010; Amoah et al., 2006; Amoah et al., 2007; Amoah et al., 2005; Hall et al., 2009; Tiimub et al., 2012; Mok and Hamilton, 2014; Mara and Sleigh, 2009; WHO, 2006; Drechsel and Scott, 2009; Lundqvist and Raschid-Sally, 2013; Jiménez, 2006; Raschid-sally and Jayakody, 2008; Mara and Sleigh, 2010). This has also led to the studies investigating wastewater associated health risks in Ghana for several pathogenic hazards (Ackerson and Awuah, 2012; Barker et al., 2014; Barker, 2014; Drechsel et al., 2009; Karavarsamis and Hamilton, 2010; Labite et al., 2010; Mok, Barker, and Hamilton, 2014; Seidu et al., 2008).

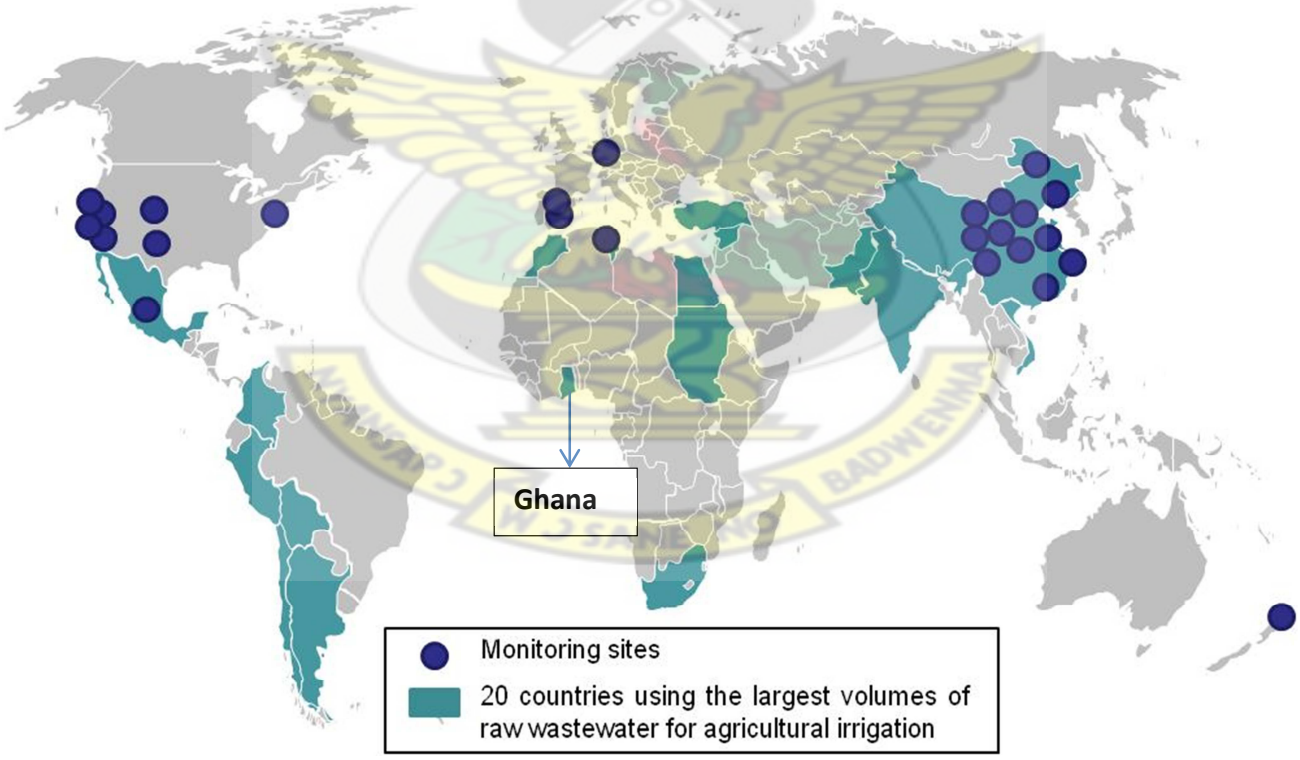


Figure 2.2: Countries using large volumes of raw wastewater for agricultural irrigation (Sources: intechopen.com)

2.5 Probabilistic Model Applications and Statistics to the Study of Quantitative Risk Assessment (QRA) of Exposure to Wastewater

It must certainly be noted that, it has taken too long for probabilistic modelling to use for biological phenomena such as microbial risk assessment as compared to physical phenomena. Most researches in quantification of uncertainty modelling of physical situations were statistical and probabilistic (Rabiner, 1989). It has been used to mimic physical phenomena by providing the necessary uncertainty surrounding most physical situation and has helped in understanding models in projecting reality in an experimental setting. Probabilistic model has been used in wide range of applications due to its flexibility and underlying assumptions (Furman and Pivi, 2002; Heath et al., 2008; Soize et al., 2008; Palmer et al., 2005; Boomsma et al., 2008; Sparck et al., 2000).

Risk assessment has also seen a great application of probabilistic modelling approach in the use of wastewater over the years (Haas, 2014; Mota et al., 2009; van der Voet et al., 2009; Haas et al., 2014; Hamilton et al., 2006; Mofarrah and Husain, 2010; Hamilton et al., 2006; van der Voet et al., 2007; Duarte and Nauta, 2015; Labite et al., 2010; Ackerson and Awuah, 2012; Machdar et al., 2013) mostly to make sound predictions and estimates. One of the key features of QRA is that it attempts to look at whole systems and not at isolated parts. Each possible adverse event is followed through to its consequences, and the consequences of different adverse events can

be combined. This is possible by using a quantitative approach, which provides a common basis for the evaluation of risks and harms (Parsons et al., 2005).

2.6 Quantitative Microbial Risk Assessment (QMRA) Framework

Quantitative Microbiological Risk Assessment (QMRA) is the process of estimating the risk from exposure to microorganisms. It is a process that involves measuring known microbial pathogens or indicators and subjects them to uncertainty quantification procedure through simulation to estimate the risk of transfer. Basically QMRA falls within the frame work of risk analysis, which basically deals with three different phases Figure 2.3 depicts the risk analysis framework(Charles et al., 2014)

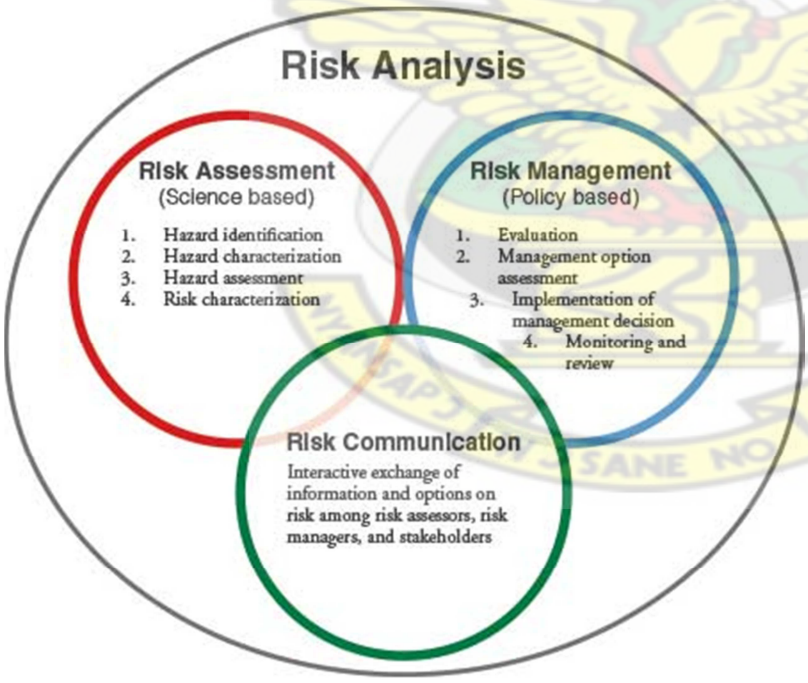


Figure 2.3: Risk Analysis Framework

QMRA in itself is made up four steps, namely;

- Hazard Identification
- Hazard Characterization
- Exposure Assessment
- Risk Characterization

2.6.1 Hazard Identification

Hazard Identification is the process of determining whether exposure to a stressor can cause an increase in the incidence of specific adverse health effects (e.g., microbial pathogens) and whether the adverse health effect is possible to occur in humans(Haas et al., 2014). Hazard identification examines scientific data available for a given pathogen of interest (or chemicals) and develops evidence to describe the link between the negative effects and the pathogen. Wastewater is known to contain numerous pathogenic microbes and serve as a source of pathogenic concern for humans, most of these pathogenic microbes are reference pathogens which includes *Norovirus* and *Cryptosporidium* spp (U S EPA, 2006).

Norovirus

Norovirus (*NoV*) is responsible for numerous cases of waterborne and food borne gastroenteritis every year which makes it a leading cause of both endemic and

epidemic gastroenteritis in the world (Hassine-Zaafrane et al., 2014; Atmar, 2010). Unlike the case of rotavirus, Norovirus cause illness among both children and adults (Glass et al., 2009) and hence makes it an ultimate target for hazard identification. After the use of rotavirus for generalization of gastroenteritis in WHO (2006), Norovirus risk assessment has gained prominence to be a replacement for rotavirus, since the latter affects children under 5 years.

Cryptosporidium spp

Cryptosporidium is also an important contaminant found in drinking water and is associated with a high risk of waterborne disease particularly for the immune compromised (Rose, 1997). The parasites can infect a significant proportion of the exposed population at low doses. The characteristics of *Cryptosporidium*, however, may vary among isolates, *Cryptosporidium parvum* (*C. parvum)* and *Cryptosporidium hominis* (*C. hominis)* are the two species of primary importance in human infections ((Reinoso and Bécares, 2008)). *Cryptosporidium* is frequently isolated from publicly owned treatment works (POTW) effluent, storm water, and livestock manure, and their respective oocyst can survive for extended periods of time in the environment. The high environmental loading of potentially human infectious *Cryptosporidium* in calves makes *Cryptosporidium* of particular interest in estimating risk related to livestock sources of fecal pollution (Reinoso and Bécares, 2008).

2.6.2 Hazard Characterization

Hazard characterisation describes the properties of the hazards and the vulnerability of consumers exposed to hazards, ending up with an expression of the dose response relationship i.e. the relationship between the quantitative occurrence of the hazard and the human health outcome (Haas et al., 2014; Medema and Ashbolt, 2006). Specifically, it is the evaluation of the nature of adverse effects of physical, chemical or biological agent which may be present in the wastewater (WHO, 2006)

2.6.3 Exposure Assessment

It is the process of estimating or measuring the magnitude, frequency and duration of exposure to a hazard, along with the number and characteristics of the population exposed. Supremely, it defines the sources, pathways, routes, and the uncertainties in the assessment (Haas et al., 2014; Medema and Ashbolt, 2006).

2.6.4 Risk Characterisation

A risk characterization is the estimation of the associate risk given the nature and presence or absence of risks, along with information about how the risk was assessed, where assumptions and uncertainties still exist, and where policy choices can be made (Haas et al., 2014; Medema and Ashbolt, 2006) . The risk characterisation is the integration of the hazard characterization (i.e. the dose-response relationship) and exposure assessment (i.e. the estimated dose). It quantifies the relationship among the integrated factors to estimate a probable adverse effect to occur within a

population or an individual with attended uncertainties, and also to indicate places or points, where mitigation for hazard control could be implemented. However, such decision of activating a mitigation process is solely risk management or political decision.

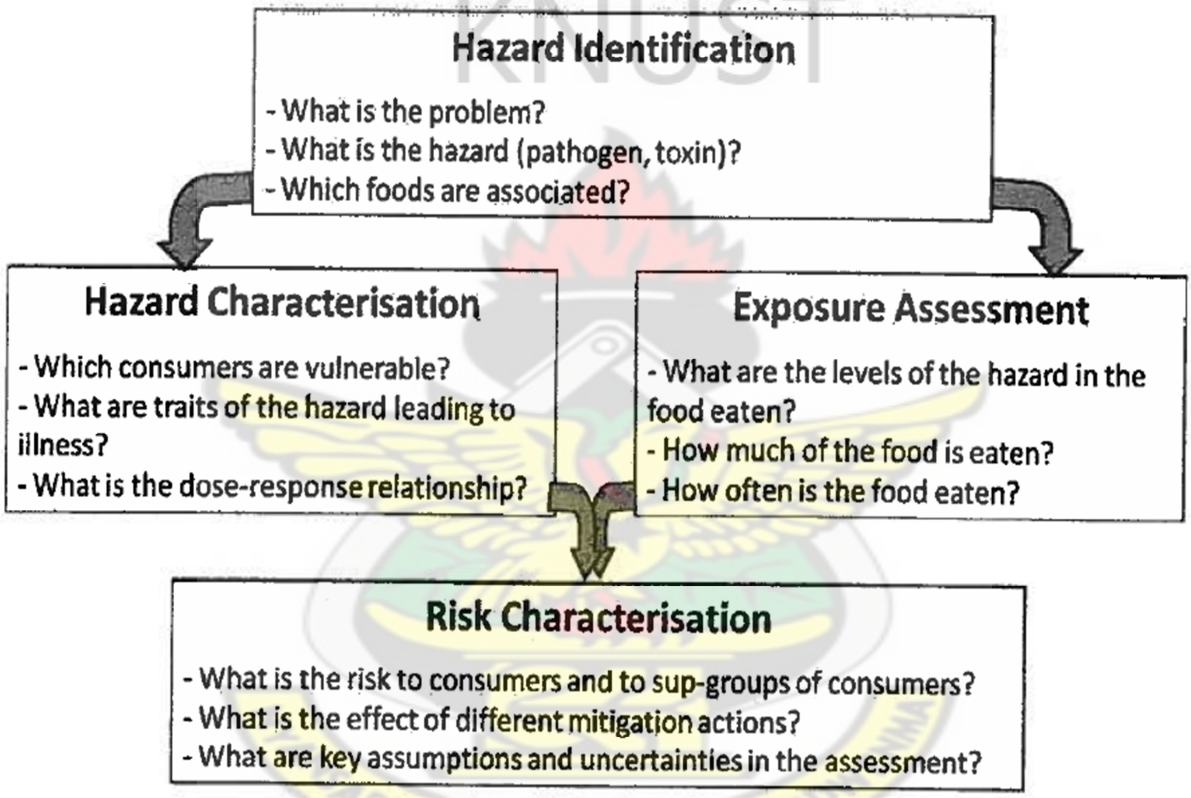


Figure 2.4: Risk QMRA Framework (sources: Brul et al., 2012)

2.7 Microbial Data Estimation in QMRA

Microbial data estimation is setting the basis of exposure to hazards that might results in infection, in as much as there is no one acceptable standard for microbial data estimation, several studies have used different methods of estimating dose per

exposure (Barker et al., 2013; Mok et al., 2014; Karavarsamis and Hamilton, 2010; Hamilton et al., 2006b; Soller, 2008; Mota et al., 2009; Teunis et al., 2002). However, one thing that is certain is the use of ratio conversion for fecal indicator organism to represent pathogen of interest, modelling after the adoption of conversion ratio used by WHO (2006) in QMRA. In the quest to improve the process, several studies (Sofia and Duarte, 2013; Yusoff et al., 2011; Duarte et al., 2013)have contributed in the advancement of improving counts of microbial data and effects of making sense of false zeroes. Nevertheless, microbial enumeration methods such as plate counts, most probable number and other alternatives are all limited in one way or the other for making a complete enumeration, hence enumeration is relied on probabilistic approach in quantifying the uncertainty surrounding it such as the recovery methods (Petterson et al., 2007).

2.8 Uncertainty and Variability Risk Estimate of Illness given infections

Inputs in QMRA may result in uncertainty if ignorance of input parameters are results of incertitude arise from due to limits on empirical studies or mensurational precision (Ferson and Ginzburg, 1996), this can be diminished by gathering additional information about the parameter, on the other hand, variability results in an intrinsic heterogeneity in input values for a parameter (Hass et al., 2014). Uncertainty results from several sources in modelling approach, which includes, the parameter, model and the scenario, whiles variability results from mainly identifiable characteristics which results in differential exposure or dose response characteristics (Haas, 2014). It is

essential to quantify uncertainty either in parameter, scenario or model source in order to prioritize factors that must be necessary to be assessed to arrive at a fairly account estimates, all different aspects that can contribute to a specific kind of uncertainty or variability must be accounted for, and their relative importance assessed (Duarte et al., 2013). Such a phenomenon is well suited for probability distribution characterization, though a consensus as to the cauterization method whether through frequentist approach or Bayesian approach is yet to settle (Rigaux et al., 2013; Ntzoufras, 2009; Greiner et al., 2013; Albert et al., 2008; Parsons et al., 2005).

2.9 Population Risks Estimation in QMRA

The focus of QMRA on hazard exposure has all been centralized in risk characterization for a single person or single exposure as an endpoint (Hass et al., 2014). The focus of characterizing exposure has mainly ended with annual risk of illness or Diseases Adjusted Life Years (DALYs) (Hamilton et al., 2006; Mok et al., 2014; Barker, 2014; Pavione et al., 2013; Barker et al., 2014; Dawber et al., 2009; Amoah et al., 2005). However, where detailed peculiarities of dynamics of illness incidence is desirable, it is better to look at the population in its entirety rather than individual exposures (Hass et al., 2014; Messner et al. 2014). Community level risk estimate is essentially based on three (3) criteria:1) the duration of disease state; 2) the carrier state; and 3) the rate at which secondary cases occur from direct or indirect contact with primary cases or individuals in the asymptomatic state. These three basic criteria

hinged on mathematical epidemiology studies of the illness and hence can be better analysed with a combination of QMRA and mathematical epidemiology of illness/diseases. Recently, a couple of studies (Swart et al., 2012; Tribble et al., 2010; Messner et al. 2014; Teunis et al., 2002) have provide the way for such inclusion and extension in quantifying risk assessment for a population.

CHAPTER 3

METHODOLOGY

Modelling with probability distribution delves much deeper into mathematical statistics or statistical theory and draws assumptions from distribution to make sense out of data and makes useful relationship among physical parameters. This chapter presents the building block of the modelling; it begins with review of some statistical parameters, probability and ends with quantification of uncertainty in probability modelling.

3.2 Statistics Measurement

Statistics is in essence closely related to probability theory; however, the two fields have entirely different goals. A typical probability problem starts with some assumptions about the distribution of a random variable (e.g., that it's binomial), and the objective is to derive some properties (probabilities, expected values, etc.) of said random variable based on the stated assumptions. The statistics problem goes almost completely the other way around. Indeed, in statistics, a sample from a given population is observed, and the goal is to learn something about that population based on the sample (Wackerly et al., 2007). In other words, the goal in statistics is to reason from sample to population, rather than from population to sample as in the case of probability.

3.2.1 Statistics

Definition 3.1. *Let* $X_1, X_2, ..., X_n$ *be a sample whose distribution may or may not depend on an unknown parameter* θ. *Then any measureable function* $T = T(X_1, X_2, ..., X_n)$ *include; the sample mean* (Hogg, McKean and Craig, 2005)

$$T = \widehat{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad 3.1$$

The sample variance

$$T = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \widehat{X}\right)^2 \qquad 3.2$$

The sample median

$$T = M_n \qquad 3.3$$

In general, $T = I_{A(x)}$ where $I_{A(x)}$ denotes the indicator function describe as (Rice, 2001)

$$I_{A(x)} = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases} \qquad 3.4$$

3.2.2 Probability

A random experiment is any process whose outcome is unpredictable. The sample space contains all the possible outcome of the experiment (Spanos, 1999).

Example (*Sample Space)*

Consider a single coin-toss, and assume that the coin will either land heads (H) or tail (T) but not both, we may define: $\Omega = \{H, T\}$, $F = \{\varnothing, \{H\}, \{T\}, \{H, T\}\}$ and hence $P\{\varnothing\} = 0$

, $P\{H\}=P\{T\}=\frac{1}{2}$, $P\{H,T\}=1$. Probability is the measure of likeliness that an event will occur and quantifies as a number between 0 and 1. Probability density functions in one, *discrete* or *continuous,* variable are denoted by $p(x)$ and $f(x)$ respectively. They are assumed to be properly normalized such that

$$\sum_x p(x)=1 \text{ and } \int_{-\infty}^{\infty} f(x)dx=1 \qquad 3.5$$

for discrete and continuous cases respectively.

Where the sum or integral are taken over all relevant values for which the probability density function is defined. The distribution function or cumulative function is also defined as (Rice, 2001)

$$p(x)=\sum_{i=-\infty}^{x} p(i) \text{ and } F(x)=\int_{-\infty}^{x} f(t)dt \qquad 3.6$$

for discrete and continuous cases respectively.

Axioms of Probability

1. The probability of an event E must be between 0 and 1 inclusive: $0 \le P(E) \le 1$
2. $P(\Omega)=1$
3. For two given events M and N, the law of total probability is presented as

$$P(M \cup N)=P(M)+P(N)-P(M \cap N)$$

4. For mutually exclusive events $E_1, E_2, ..., E_n$ (Rice, 2001),

$$P(E_1 \cup E_2 \cup ... \cup E_n) = P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$$

3.3 Expectation and Variance

The Mathematical expectation of a random variable X denotes the average of all possible values of $X_1, X_2, ..., X_n$ whiles the variance of a random variable χ measures the spread of value around the expectation. These two characteristics of a random variable are used in the description of distribution, reliability, sufficiency, consistency and performance of random variables as well as hypothesis testing procedures. In practices of all spheres, scientists rely heavily on the use of these two parameters in describing phenomena, procedures and processes to give a summary of an observed phenomenon in order to help make sound judgment from physical and experimental circumstances (Jaynes, 2003).

In the description of expected value of a random variable X, the probability mass function or density function is used for discrete and continuous random values respectively.

Definition3.2

If X is a discrete random variable with frequency probability mass function $p(x)$, the expected value of X, denoted $E(X)$ is

$$\mu = E(X) = \sum_i x_i p(x_i) \qquad 3.7$$

Provided that $\sum_i |x_i| p(x_i) < \infty$, *if the sum diverges, then expectation is undefined* (Jaynes, 2003; Wackerly et al., 2007).

For the case of X continuous random variable with density $f(x)$

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad 3.8$$

Provided that $\int |x| f(x) dx < \infty$, if the integral diverges, the expectation is undefined(Rice, 2001).

Theorem 3.1

If X *is a random variable with* $P(X \geq 0) = 1$ *and for which* $E(X)$ *exists, then* $P(X \geq t) \leq E(X)/t$ (Wackerly et al., 2007).

Proof

For the discrete case

$$E(X) = \sum_x x p(x)$$

$$= \sum_{x<t} x p(x) + \sum_{x \geq t} x p(x)$$

Therefore,

$$\Rightarrow E(X) \geq \sum_{x \geq t} xp(x) \qquad 3.9$$

All terms in the sums are non-negative due to random variable X taking only non-negative values(Rice, 2001)

$$E(X) \geq \sum_{x \geq t} tp(x) = tP(X \geq t) \qquad 3.10$$

For measure of spread or dispersion, variance is measured as follow

Definition 3.3

If X is a random variable with expected value $E(X)$, the variance of X is

$$Var(X) = E\left\{\left[X - E(X)\right]^2\right\} = \int (x-\mu)^2 f(x)dx \qquad 3.11$$

This can lead to a simplified form as(Hogg et al., 2005)

$$Var(X) = \sigma^2 = E\left[X^2\right] - \mu^2 = E\left[X^2\right] - \left(E[X]\right)^2 \qquad 3.12$$

Proof

Provided that, the expectation exist

$$\left.\begin{aligned} Var(X) &= E\left\{\left[X - E(X)\right]^2\right\} \\ &= E\left(X^2 - 2\mu X + \mu^2\right) \end{aligned}\right\} \quad 3.13$$

By the linearity of expectation

$$Var(X) = E\left(X^2\right) - 2\mu E(X) + \mu^2$$

$$= E\left(X^2\right) - 2\mu^2 + \mu^2$$

$$= E\left(X^2\right) - \mu^2$$

$$= E\left(X^2\right) - \left(\left[E(X)\right]\right)^2 \quad 3.14$$

3.3.1 Moments

The moments are the expectations of powers of the of the random variable(Hogg et al., 2005), in general the algebraic moment of order r is defined on the expectation value as

$$E\left(X^r\right) = \mu_r^{'} = \sum_k k^r p(k) = \int_{-\infty}^{\infty} x^r f(x)\,dx \quad 3.15$$

Form moments running from order $0,1,...,n$, clearly $\mu_0^{'} = 1$ and from normalization condition $\mu_1^{'} = \mu = E(X)$, thus the mean or the expectation.

Central moment of order r is also defined as

$$\mu_r^{'} = E\left[\left(k - E(k)\right)^r\right] = E\left[\left(x - E(x)\right)^r\right] \qquad 3.16$$

The most common used of the central moment is order 2 (Hogg et al., 2005), thus $\mu_2^{'}$ which is the variance of the distribution. In using the third and fourth central moments, the third central moment is used to measure the asymmetry or skewness in the distribution whiles the fourth central moment is also used to measure the degree of peakness. We often define the coefficient of skewness η_1 and kurtosis η_2 by

$$\eta_1 = \frac{\mu_3}{\mu_2^{3/2}} \text{ And } \eta_2 = \frac{\mu_4}{\mu_2^2} \qquad 3.17$$

The coefficient of excess kurtosis is given by

$$\eta_2^{'} = \frac{\mu_4}{\mu_2^2} - 3 \qquad 3.18$$

Where the shift by 3 units assures that both measures are zero for a normal distribution. Distribution with positive kurtosis is called *leptokurtic*, those with kurtosis around zero are also called *mesokurtic* and those with negative kurtosis are called *platykurtic*. Leptokurtic distributions are normally more peaked than normal distribution while platykurtic distributions are more flat topped (Jaynes, 2003; Rice, 2001).

3.4 Joint Distribution and Marginal Density

The joint behaviour of two random variables X and Y is determined by the multiplication of their density functions or the cumulative distribution function,

irrespective of the nature of the density function (discrete or continuous)(Hogg et al., 2005; Wackerly et al., 2007).

Definition 3.4: *For any two random variables X and Y, the joint cumulative probability distribution function of X and Y is*

$$F(x,y)=P(X\le x,Y\le y),-\infty<x,y<\infty \tag{3.19}$$

$$P(x_1\le X\le x_2,y_1\le Y\le y_2)=F(x_2,y_2)-F(x_2,y\ )-F(x_1,y_2)+F(x_1,y_1) \tag{3.20}$$

For a discrete random variable X and Y defined on the same sample space, the probability joint mass function is defined as

$$P(x_i,y_i)=P(X=x_i,Y=y_i)$$

$$P_X(x)=P\{X=x\}=P(X=x,Y\in\mathbb{R})=\sum_{y:p(x,y)>0}p(x,y) \tag{3.21}$$

And

$$P_Y(y)=P\{Y=y\}=P(Y=y,X\in\mathbb{R})=\sum_{x:p(x,y)>0}p(x,y) \tag{3.22}$$

For a continuous random variable. Let X and Y be joint continuous, if there exists a function $f(x,y)$ defined for all x,y such that for any $C\subset\mathbb{R}^2$ the joint probability density function defined as

$$P\big((X,Y)\in A\big)=\iint\limits_{x,y\in A} f(x,y)\,dxdy \qquad 3.23$$

Thus $P\{X\in A,Y\in B\}=\int\limits_{B}\int\limits_{A} f(x,y)\,dxdy$, hence $\dfrac{\partial^2}{\partial a\partial b}F(a,b)=\dfrac{\partial^2}{\partial a\partial b}\int\limits_{-\infty}^{b}\int\limits_{-\infty}^{a} f(x,y)\,d = f(a,b)$

Definition 3.5: *Two random variables are said to be independent if for any two sets of real numbers* A *and* B*, then*

$$P\{X\in A,Y\in B\}=P\{X\in A\}\,P\{Y\in B\} \qquad 3.24$$

Then $\forall a,b$*, hence two random variables* X *and* Y *are independent if and only if the function is the product of the marginal distribution functions*

$$F(a,b)=F_X(a)\,F_Y(b) \qquad 3.25$$

Similarly, in the jointly continuous case independence is equivalent to

$$f(x,y)=f_X(x)\,f_Y(y) \qquad 3.26$$

For more than two random variables $X_1,...,X_n$ *are independent if for all sets of real number* $A_1,...,A_n$ *then*

$$P\{X_1\in A_1,X_2\in A_2,...,X_n\in A_n\}=\prod_{i=1}^{n}P\{X_i\in A_i\} \qquad 3.27$$

Hence the joint probability mass functions and joint probability density functions hold, that is in both cases independence is equivalent to being equal to their respective product of the marginal.

3.5 Some Common Probability Distributions

This section presents some common probability distributions which feature mostly in this study, with its characteristics. They were chosen mainly based on its relevant to characterize the conditions under study.

3.5.1Bernoulli Distribution

This can be used to model a single experiment which has two possible outcomes; It is a random variable probability distribution which takes value 1 with success probability p and 0 with failure probability $q = 1 - p$. Both p and $q = 1 - p$ are limited to the interval from zero to one. The distribution has the simple form (Hogg et al., 2005);

$$p(x, p) = \begin{cases} 1 - p, \ if \ \ x = 0 \ (failure) \\ p, \ if \ \ x = 1 \ (\text{Success}) \end{cases} \qquad 3.28$$

With expectation $E(X) = p$ and variance $Var(X) = p(1-p)$

3.5.2 Beta Distribution

It is essential for the modelling of behaviour of random variables limited to intervals of finite length. It is a family of continuous distribution defined on the interval $[0,1]$. The probability density function is given as:

$$f(x:\alpha,\beta)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1} \qquad 3.29$$

Where $\alpha,\beta>0$ and $0\le x\le 1$. The quantity $B(\alpha,\beta)$ is the beta function in terms of the common Gamma function defined as

$$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \qquad 3.30$$

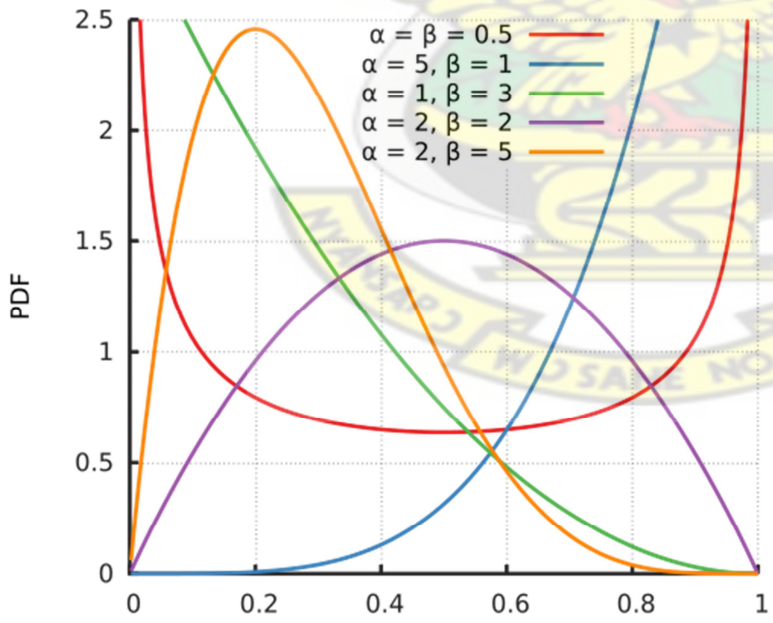


Figure 3.1: Beta Family of Functions for selected parametric values of α and β

The expectation and variance are given as

$$E(X)=\frac{\alpha}{\alpha+\beta} \text{ and } Var(X)=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad 3.31$$

The third central moment given as $\mu_3=\frac{2\alpha\beta(\beta-\alpha)}{(\alpha+\beta)^3(\alpha+\beta+1)(\alpha+\beta+2)}$ and fourth central moment are also given as $\mu_4=\frac{3\alpha\beta\left(2(\alpha+\beta)^2+\alpha\beta(\beta+\alpha-6)\right)}{(\alpha+\beta)^4(\alpha+\beta+1)(\alpha+\beta+2)(\alpha+\beta+3)}$

3.5.3 Binomial Distribution

The distribution describes the probability of exactly x successes in N trials if the probability of a success in a single trial is p and a failure is $q=1-p$. It has the distribution.

$$p(x,N,p)=\binom{N}{x}p^x(1-p)^{N-x} \quad 3.32$$

Where $0\le x\le N$, $0\le p\le 1$, with expectation $E(X)=Np$, variance $Var(X)=Np(1-p)=Npq$, third central moment $\mu_3=Np(1-p)(1-2p)=Npq(q-p)$ and fourth central moment $\mu_4=Np(1-p)\left[1+3p(1-p)(N-2)\right]=Npq\left[1+3pq(N-2)\right]$

3.5.4 Exponential Distribution

It is a random variable distribution use to model processes where events happen at a constant rate. The exponential distribution is given by

$$f(x,\lambda)=\frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right) \tag{3.33}$$

Where $\lambda, x>0$. The distribution has the expectation $E(X)=\lambda$, variance $Var(X)=\lambda^2$, third central moment $\mu_3=2\lambda^3$ and fourth central moment $\mu_4=9\lambda^4$

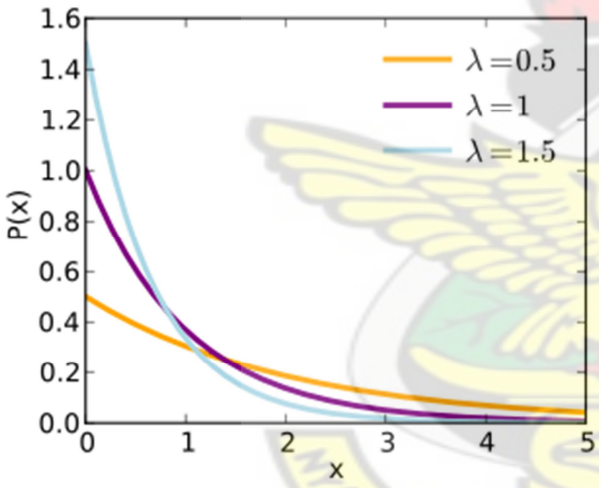


Figure 3.2: Exponential distribution with different expectation values.

3.5.5 Double Exponential Distribution

The Double exponential distribution popular known as the *Laplace distribution* is given by:

$$f(x;\mu,\lambda)=\frac{\lambda}{2}\exp\left(-\lambda|x-\mu|\right) \quad 3.34$$

Where $x,\mu\in\mathbb{R}$,and $\lambda>0\in\mathbb{R}$, it has the expectation $E(x)=\mu$

3.5.6 F-Distribution

This is a random variable continuous distribution widely used in test statistics and most notably in the analysis of variance, it is given by

$$f(F;d_1,d_2)=\frac{d_1^{\frac{d_1}{2}}d_2^{\frac{d_2}{2}}\Gamma\left(\frac{d_1+d_2}{2}\right)}{\Gamma\left(\frac{d_1}{2}\right)\Gamma\left(\frac{d_2}{2}\right)}*\frac{F^{\frac{d_1}{2}-1}}{\left(d_1F+d_2\right)^{\frac{d_1+d_2}{2}}}$$

$$=\frac{d_1^{\frac{d_1}{2}}d_2^{\frac{d_2}{2}}}{B\left(\frac{d_1}{2},\frac{d_2}{2}\right)}*\frac{F^{\frac{d_1}{2}-1}}{\left(d_1F+d_2\right)^{\frac{d_1+d_2}{2}}} \quad 3.35$$

The expectation and variance are given by $\frac{d_2}{d_2-2};d_2>2$ and $\frac{2d_2^{\,2}(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}$ respectively. Figure 3.3 depicts the F-distribution values of d_1 and d_2.

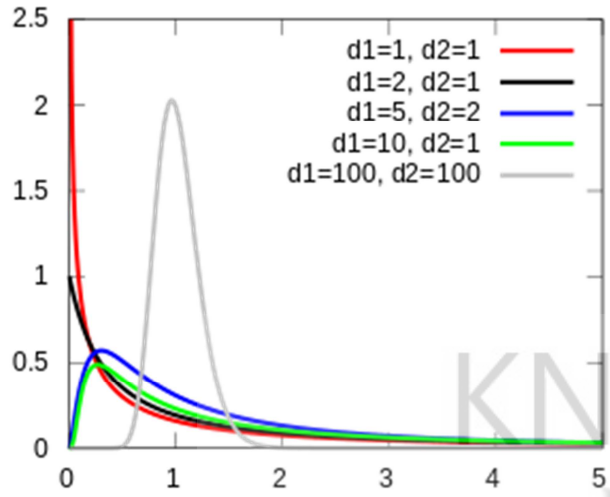


Figure 3.3: The F-distribution for different values of d_1 and d_2

3.5.7 Gamma Distribution

The probability density function of the Gamma distribution is given by

$$f(x,\theta,k)=\frac{\theta(\theta x)^{k-1}}{\Gamma(k)}\exp(-\theta x)$$

3.36

For $k>0$, $\theta>0$, $x\geq 0$ with expectation $E(X)=\frac{k}{\theta}$ and variance $Var(X)=\frac{k}{\theta^2}$, the k is a shape parameter and is θ a scale parameter. The gamma function is defined as:

$$\Gamma(k)=\int_0^{\infty} e^{-y}y^{k-1}dy \qquad 3.37$$

Figure 3.4 shows the Gamma probability density function with different scale and shape parameter values

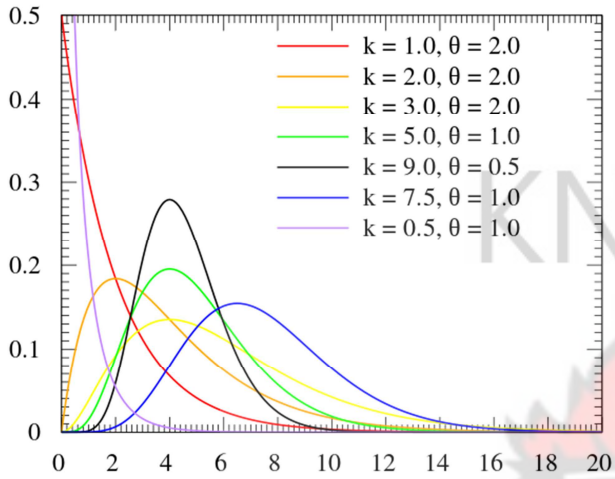


Figure 3.4: Gamma probability density function with different scale and shape parameter values

3.5.8 Chi-square Distribution

It's a distribution of random variable that describes the ratio of two independent standard normal variable, it is a special form of Gamma distribution with $\theta = 0.5$ and $k = k/2$. It is useful for the derivation distribution of sample variance and the goodness of fit test. The probability density function with k degrees of freedom is given by

$$f(x;k) = \frac{\left(\frac{x}{2}\right)^{\frac{k}{2}-1} e^{\frac{x}{2}}}{2\Gamma\left(\frac{k}{2}\right)} \qquad 3.38$$

Where $x \geq 0$. The figure below shows some chi-square family of graphs for k values of 1,2,3,4,6 and 9

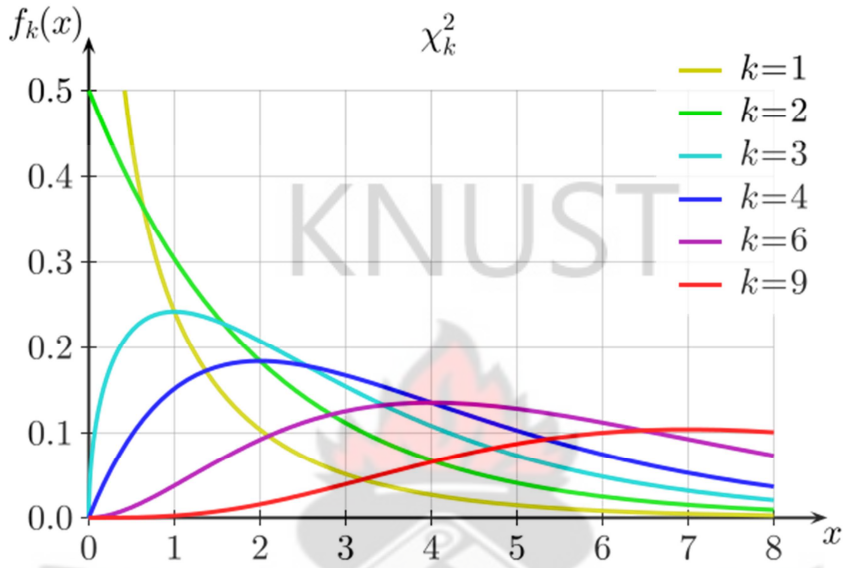


Figure 3.5: Graph of chi-square distribution for some values of k

3.5.9 The Generalized Gamma Distribution

The gamma distribution is a random variable distribution bounded on one side, a generalized gamma distribution is obtained by adding a third parameter giving it a more flexible version of the distribution, the distribution is given by

$$f(x;a,b,c) = \frac{ac(ax)^{bc-1}}{\Gamma(b)} \exp\left[-(ac)^c\right] \qquad 3.39$$

Where a (a scale parameter) and b are real positive parameters as is used for the gamma distribution but a third parameter c has been added to control the distribution into different family of distribution (Rice, 2001)

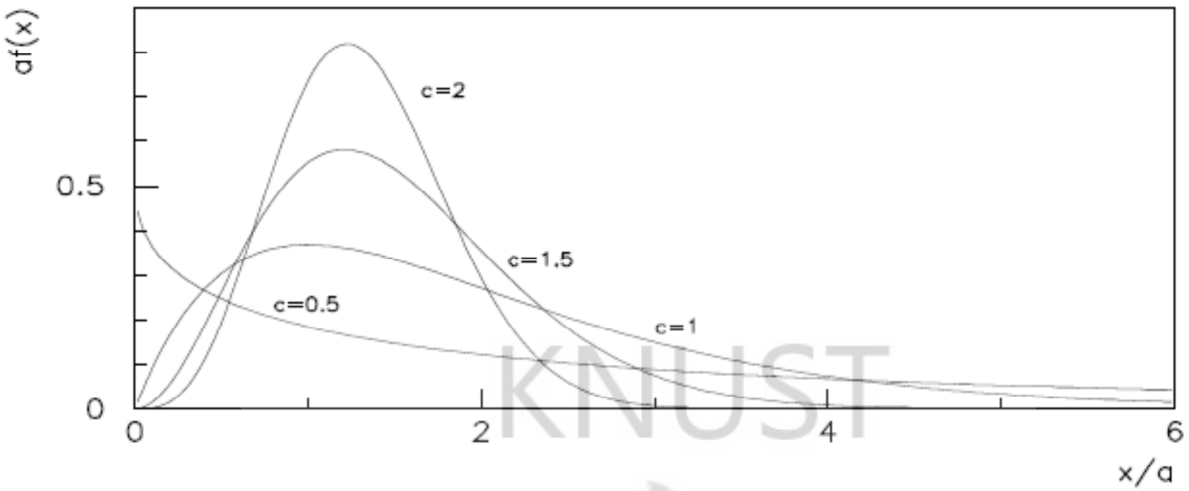


Figure 3.6: A generalized Gamma distribution with different values of c $[a=1, b=2]$

Table 3.1: Generalized Gamma Distribution to other related distributions

Distribution	**a**	**B**	**C**
Generalized Gamma	a	B	C
Gamma	a	B	1
Chi-squared	$1/2$	$n/2$	1
Exponential	$1/\alpha$	1	1
Weibul	$1/\sigma$	1	η
Rayleigh	$1/\alpha\sqrt{2}$	1	2
Maxwell	$1/\alpha\sqrt{2}$	$3/2$	2
Standard normal (folded)	$1/\sqrt{2}$	$1/2$	2

3.5.10 Geometric Distribution

In practice, the geometric distribution is use to express the probability of having to wait exactly r trials before the first successful event, if the probability of a success in a single trial is p (and its failure is $q = 1 - p$). The distribution is given by

$$p(r; p) = p(1-p)^{r-1} \qquad 3.40$$

With expectation $E(r)=\frac{1}{p}$, variance $Var(r)=\frac{1-p}{p^2}$, the third central moment $\mu_3=\frac{(1-p)(2-p)}{p^3}$ and the fourth central moment $\mu_4=\frac{(1-p)\left(p^2-9p+9\right)}{p^4}$

3.5.11 Hyper-geometric Distribution

In practice the random variable distribution of the hyper-geometric distribution describes the experiments where elements are picked at random without replacement. It is given by

$$p(r;n,N,M)=\frac{\binom{M}{r}\binom{N-M}{n-r}}{\binom{N}{n}}$$

3.41

Where the random variable r has limits from $\max(0,n-N+M)$ to $\min(n,M)$, $p=\frac{M}{N}$, $q=(1-p)$. it has the expectation $E(r)=np$, the variance $Var(r)=npq\left(\frac{N-n}{N-1}\right)$ the third central moment $\mu_3=npq(q-p)\frac{(N-n)(N-2n)}{(N-1)(N-2)}$ and the fourth central moment $\mu_4=npq(N-n)\frac{N(N+1)-6n(N-n)+3pq\left(N^2(n-2)-Nn^2+6n(N-n)\right)}{(N-1)(N-2)(N-3)}$

3.5.12 Log-normal Distribution

The log-normal distribution is given by

$$f(x;\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right] \qquad 3.42$$

Where the variable $x>0$ and the parameters $\mu>0$ and $\sigma>0$ are all real numbers, in this case denoted in the same spirit as normal distribution but $\mu=\ln x$. It has the expectation $E(x)=\exp\left(\mu+\frac{\sigma^2}{2}\right)$ and variance $Var(x)=\exp\left(2\mu+\sigma^2\right)\left[\exp\left(\sigma^2\right)-1\right]$. In practices it is used to characterize random variables which have non-negative values.

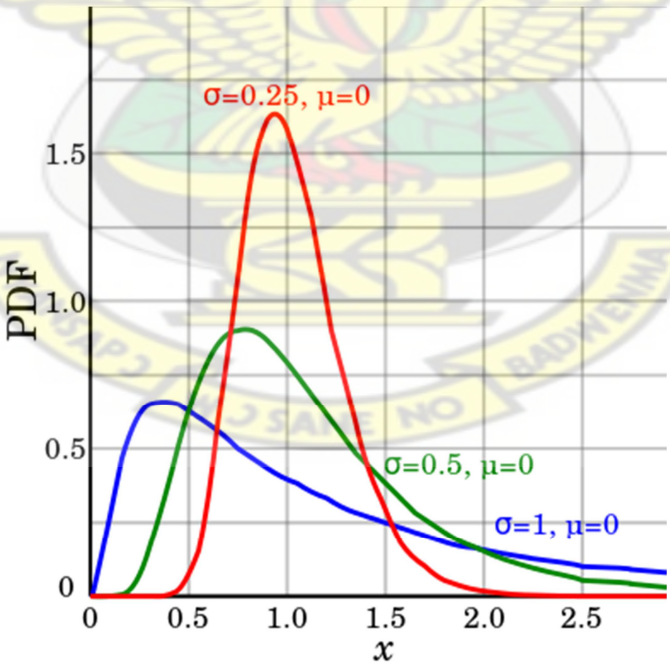


Figure 3.7: The log-normal distribution

3.5.13 Negative Binomial Distribution

In practice, the negative binomial distribution expresses the probability of having to wait exactly r trials until successes have occurred if the probability of a success in a single trial is p and its failure as $q = 1 - p$. The distribution is given by

$$p(r;k,p) = \binom{r-1}{k-1} p^k (1-p)^{r-k} \qquad 3.43$$

Where the variable $r \geq k$ and the parameter $k > 0$ are integers and the parameter $0 \leq p \leq 1$ is a real number, it has the expectation $E(r) = \frac{k}{p}$, Variance $Var(r) = \frac{kq}{p^2}$, third central moment $\mu_3 = \frac{kq(2-p)}{p^3}$ and fourth central moment $\mu_4 = \frac{kq(p^2 - 6p + 6 + 3kq)}{p^4}$

3.5.14 Normal Distribution

The normal distribution often call Gauss distribution is the most important distribution used in statistics, it has many application due to its underlying assumption. The distribution is given by

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma} \right)^2 \right] \qquad 3.44$$

It has the expectation $E(x) = \mu$ and variance $Var(x) = \sigma^2$

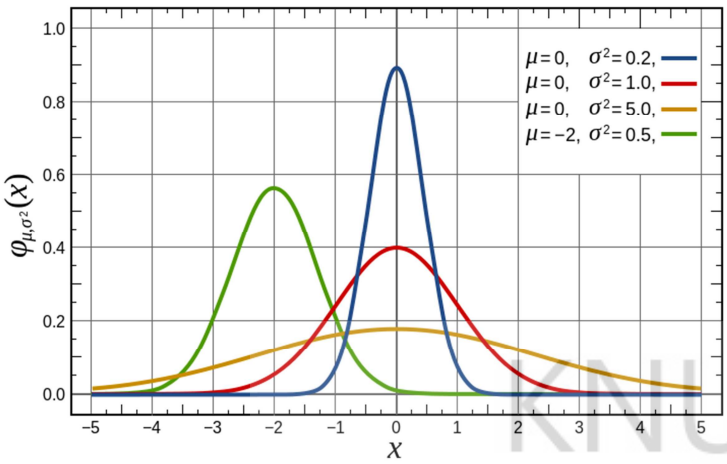


Figure 3.8: Normal distribution

3.5.15 Poisson Distribution

Poisson distribution describes the probability to find exactly x events in a given length of time if the events occur independently at a constant rate μ. The distribution is given by

$$p(x;\mu)=\frac{\mu^{x}e^{-\mu}}{x!} \qquad 3.45$$

The distribution has expectation $E(x)=\mu$, variance $Var(x)=\mu$, third central moment $\mu_3=\mu$ and fourth central moment $\mu_4=\mu(1+3\mu)$

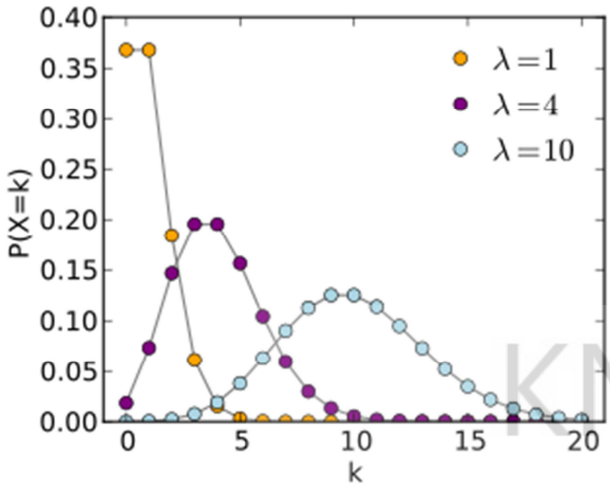


Figure 3.9: Poisson distribution

3.5.16 Student's t-distribution

The student's t-distribution is given by

$$f(x;v)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{x^2}{v}\right)^{-\frac{v+1}{2}} \qquad 3.46$$

Where $v>0$ and $x\in\mathbb{R}$

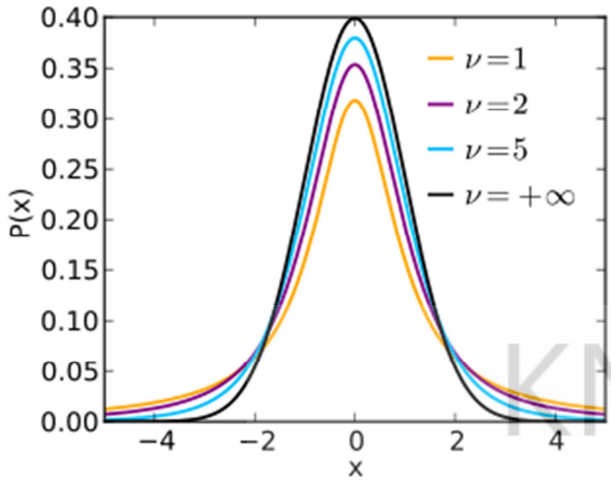


Figure 3.10: The student t distribution

3.5.17 Triangular Distribution/Pert Distribution

The triangular distribution is given by

$$f(x;\mu,\theta)=\frac{-|x-\mu|}{\theta^2}+\frac{1}{\theta} \qquad 3.47$$

Where the variable x is bounded to the interval $\mu-\theta\le x\le\mu+\theta$ and the location and scale parameters μ and $\theta(\theta>0)$ all are real numbers. It has expectation $E(x)=\mu$, due to its symmetry of the distribution odd central moments vanishes while even moments are given by $\mu_n=\frac{2\theta^n}{(n+1)(n+2)}$

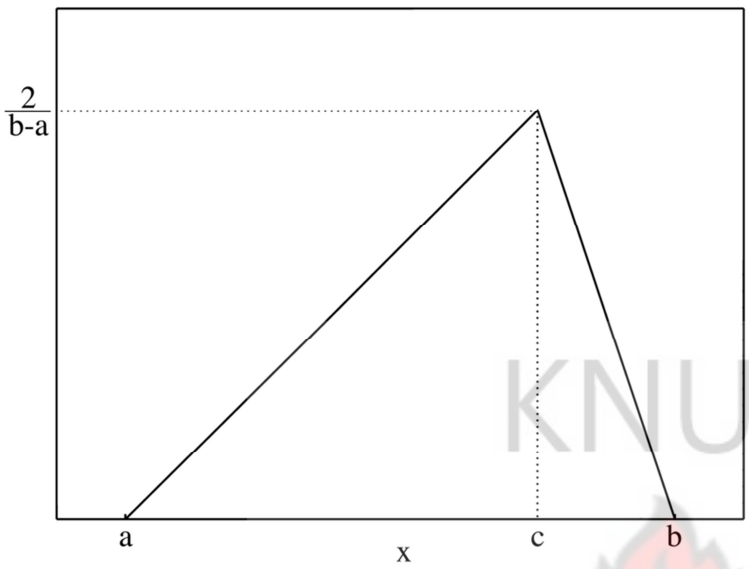


Figure 3.11: The Triangular Distribution

Practically, the pert distribution is used exclusively in modelling expert's estimates, where one is given expert's minimum, most likely and maximum values or guesses, and forms an alternative to the triangle distribution. It is given by $Pert(a,b,c) = Beta4(\alpha_1, \alpha_2, a, c)$, $\alpha_1 = \frac{(\mu - a)(2b - a - c)}{(b - \mu)(c - a)}$, $\alpha_2 = \frac{\alpha_1 (c - \mu)}{(\mu - a)}$ and mean $\mu = \frac{a + 4b + c}{6}$.

3.5.18 Uniform Distribution

The uniform distribution is a simple case with the distribution given by

$$f(x;a,b)=\begin{cases}0 & for\ x<a\\ \dfrac{1}{b-a}, & for\ \ a\le x\le b\\ 0 & for\ x>b\end{cases} \qquad 3.48$$

It has the expectation $E(x)=\frac{a+b}{2}$, variance $Var(x)=\frac{(b-a)^2}{12}$, the third central moment $\mu_3=0$ and the fourth central moment $\mu_4=\frac{(b-a)^4}{80}$

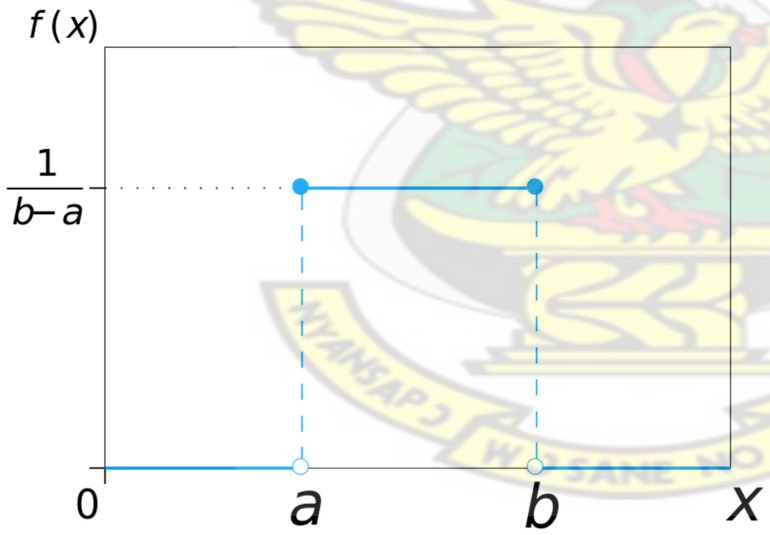


Figure 3.12: The probability density function of Uniform Distribution

3.5.19 Inverse-Gaussian Distribution

Also known as *Wald distribution* is a continuous probability density function with two parameter of support, with *pdf* given as

$$f(x,\mu,\lambda)=\left[\frac{\lambda}{2\pi x^3}\right]^{1/2}\exp\frac{-\lambda(x-\mu)^2}{2\mu^2 x} \qquad 3.49$$

Where $x>0,\mu>0$ is the mean and $\lambda>0$ is the shape of the parameter, with expectation $E(x)=\mu, E\left(\frac{1}{x}\right)=\frac{1}{\mu}+\frac{1}{\lambda}$ and variance $Var(x)=\frac{\mu^3}{\lambda}, Var\left(\frac{1}{x}\right)=\frac{1}{\mu\lambda}+\frac{2}{\lambda^2}$

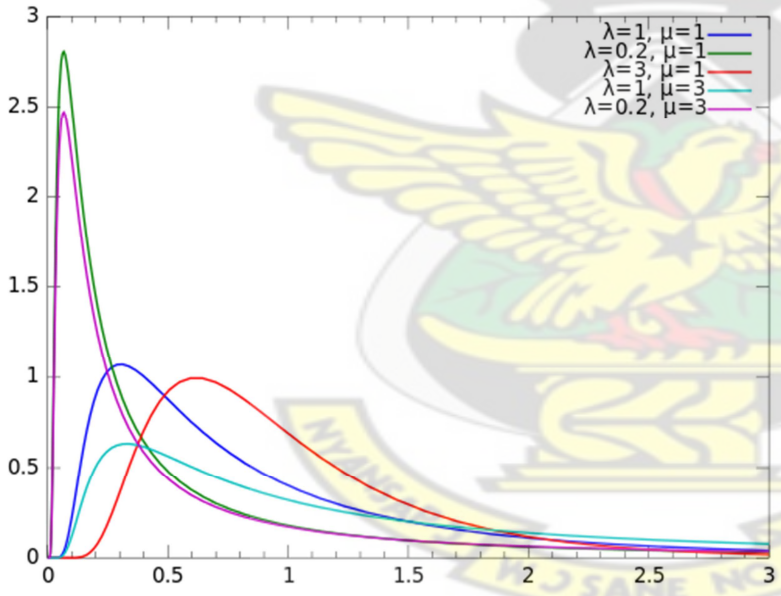


Figure 3.13: Graph of Inverse Gaussian Distribution

3.6 Estimation of Parameters in Distribution

Let $X_1, X_2, ..., X_n$ be a sample from a distribution (section 3.5.1-3.5.19) with a Cumulative Density Function (CDF) F_θ, depending on the parameter θ which is unknown (Jaynes, 2003). In most cases it is important also to know the parameter space Θ

Definition 3.6

Let $X_1, X_2, ..., X_n$ *be a sample distribution* F_θ *with* $\theta \in \Theta$. *A parameter estimate of* θ *is a function* $\hat{\theta} = \hat{\theta}(X_1, X_2, ..., X_n)$ *taking values in* Θ. *When data is observed but we don't know which of the models fit* F_θ *assuming* $\theta \in \{X_1, ..., X_n\} \approx (f_\theta)$ *then it is up to the determination of the best model.*

3.6.1 Maximum Likelihood Estimators (MLEs)

Supposed $X_1, ..., X_n \sim f_\theta$ where θ is unknown. Then by definition, the likelihood function for the independent means multiply is given as (Johansen and Juselius, 2009)

$$L(\theta) = f_\theta(X_1, ..., X_n) = \prod_{i=1}^{n} f_\theta(X_i) \qquad 3.50$$

Definition 3.7: *Given* $X_1, ..., X_n \overset{iid}{\sim} f_\theta$, *let* $L(\theta)$ *and* $l(\theta)$ *be the likelihood and log-likelihood functions, respectively, then the maximum likelihood estimator of* θ *is defined as* (Kotz et al., 2004)

$$\hat{\theta} = \arg\max_{\theta \in \Theta} L(\theta) = \arg\max_{\theta \in \Theta} l(\theta) \qquad 3.51$$

The estimation problem reduces by solving the likelihood equation $\frac{\partial}{\partial\theta}l(\theta)=0$.

Example with normal distribution

If $X_1,...,X_n \overset{iid}{\sim} N(\mu,\sigma^2)$ defined in section 3.2, their joint density is the product of their marginal densities

$$f(x_1,...,x_n|\mu,\sigma)=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right] \qquad 3.52$$

The log likelihood function is thus

$$l(\mu,\sigma)=-n\log\sigma-\frac{n}{2}\log 2\pi-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \qquad 3.53$$

With partial derivatives with respect μ to and σ yields

$$\left.\begin{aligned}\frac{\partial l}{\partial\mu}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)\\ \frac{\partial l}{\partial\sigma}&=-\frac{n}{\sigma}+\sigma^{-3}\sum_{i=1}^{n}(X_i-\mu)^2\end{aligned}\right\} \qquad 3.54$$

Solving simultaneously yields

$$\left.\begin{aligned} \hat{\mu} &= \overline{X} \\ \hat{\sigma} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2} \end{aligned}\right\} \qquad 3.55$$

3.6.2 Uncertainty Quantification and Sampling

Monte Carlo Method and Latin Hypercube Sampling

Monte Carlo simulation relies on repeated random sampling procedure and statistical analysis to compute test statistics (Mason et al., 2008). Let a test statistics $T_n = T\left(X_1,...,X_n\right)$, where T_n is a random variable from a function of random variables $X_1,...,X_n$. Hence sampling distribution T_n is

$$\left.\begin{aligned} X_1^{(1)},...,X_n^{(1)} &\rightarrow T_n^{(1)} = T\left(X_1^{(1)},...,X_n^{(1)}\right) \\ X_1^{(2)},...,X_n^{(2)} &\rightarrow T_n^{(2)} = T\left(X_1^{(2)},...,X_n^{(2)}\right) \\ X_1^{(s)},...,X_n^{(s)} &\Longrightarrow T_n^{(s)} = T\left(X_1^{(s)},...,X_n^{(s)}\right) \end{aligned}\right\} \qquad 3.56$$

That is for each sample $X_1^{(s)},...,X_n^{(s)}$ of size n, there is a corresponding $T_n^{(s)}$ obtained by applying the function $T(.)$ to that particular sample. This method of simulation is very closely related to a random experiment for which specific results are not known in advanced. Latin Hypercube Sampling (LHS) provides an efficient sampling method in

place or a random sampling, it partition the range into N intervals of equal probability and sample within each range with equal probability as explore in Keramat and Kielbasa (1999).

Gibbs Sampling

Suppose a joint probability $f(x, y_1, ..., y_n)$, the variable of interest characteristics of the marginal density $\int ... \int f(x, y_1, ..., y_n) dy_1, ..., dy_n$. Rather than computing the direct approximate of $f(x)$, the Gibbs sampler effectively generation of a sample $X_1, ..., X_m \sim f(x)$ without requiring $f(x)$ by simulating a large enough sample, the desired characteristics can be calculated to a desired degree of accuracy (Casella and George I, 1992).

3.7 Fitting Distribution to Pathogen Concentration for Exposure Assessment

Practically, the purpose of exposure assessment is to determine the amount of pathogen organism corresponding to a single exposure which comprises the expected dose and its distribution. The expected dose d concentration of microorganisms given the mean concentration μ and consumption per exposure m,

$$d = E(\mu m) \qquad 3.57$$

Assuming independency of μ and m, then by the expected dose computation (section 3.3) is

$$\bar{d} = \overline{\mu m} \tag{3.58}$$

3.7.1 Poisson Random Distribution of Organism

Measuring micro-organism distribution with Poisson distribution as described (section 3.6.1). If organism is distributed randomly in volume V, hence

$$P(x=n) = \frac{\left(\bar{\mu}V\right)^n}{n!}\exp\left(-\bar{\mu}V\right) \tag{3.59}$$

Where $\bar{\mu}$ is the mean density and $P(x=n)$ is the probability of x samples containing n organisms. The distribution is completely known when the parameter $\bar{\mu}$ is known employing the maximum likelihood method (section 3.6.1.), the average dose expected for set of samples of each volume is equal to $\bar{\mu}V$. Assume $x \sim \exp(\mu)$ for $N_L \le x \le N_U$, then

$$P(n_L \le x \le n_U) = \sum_{n=n_L}^{n=n_U} \frac{\left(\bar{\mu}V\right)^n}{n!}\exp\left(-\bar{\mu}V\right) \tag{3.60}$$

For infinite upper limit of the concentration

$$\left.\begin{aligned} P\left(n_L \le x \le \infty\right) &= 1 - P\left(0 \le x \le \left(n_L - 1\right)\right) \\ &= 1 - \sum_{n=0}^{n=n_L-1} \frac{\left(\bar{\mu}V\right)^n}{n!} \exp\left(-\bar{\mu}V\right) \end{aligned}\right\} \quad 3.61$$

At different volumes, and measuring the number of organisms in each samples, assuming independence of samples, the likelihood function for obtaining the unknown parameters (Section 3.6.1)

$$\left.\begin{aligned} L &= \prod_{i=1}^{k} \frac{\left(\bar{\mu}V_i\right)^{n_i}}{n_i!} \exp\left(-\bar{\mu}V_i\right) \\ -\ln\left(L'\right) &= \bar{\mu}\sum_{i=1}^{k} V_i - \sum_{i=1}^{k} n_i \ln\left(\bar{\mu}V_i\right) \\ &= \bar{\mu}\sum_{i=1}^{k} V_i - \ln\left(\bar{\mu}\right)\sum_{i=1}^{k} n_i - \sum_{i=1}^{k} n_i \ln\left(\bar{\mu}\right) \end{aligned}\right\} \quad 3.62$$

Estimating the Poisson mean count assay with either a constant and variable volume, the above equation can directly be applied to body of data and expected mean estimated as:

$$\left.\begin{aligned}\frac{d\left(-\ln\left(L'\right)\right)}{d\left(\bar{\mu}\right)}&=\sum_{i=1}^{k}V_i-\frac{1}{\mu}\sum_{i=1}^{k}n_i=0\\ \bar{\mu}_{ML}&=\frac{\sum_{i=1}^{k}n_i}{\sum_{i=1}^{k}V_i}\end{aligned}\right\} \quad 3.63$$

Since $\dfrac{d^2\left(-\ln\left(L'\right)\right)}{d\left(\bar{\mu}\right)^2}>0$

For count assays with upper limits where detection are termed 'too numerous to count' (TNTC)

$$\left.\begin{aligned}-\ln\left(L'\right)&=\bar{\mu}\sum_{i=1}^{k}V_i-\sum_{i=1}^{k}n_i\ln\left(\bar{\mu}V_i\right)-\sum_{i=j+1}^{j+k}\ln\left[1-\Gamma\left(\left(n_{L,i}-1\right),-\bar{\mu}V\right)\right]\\ L&=\left[\prod_{i=1}^{k}\frac{\left(\bar{\mu}V_i\right)^{n_i}}{n_i!}\exp\left(-\bar{\mu}V_i\right)\right]\left\{\prod_{i=j+1}^{k+j}\left[1-\Gamma\left(\left(n_{L,i}-1\right),-\bar{\mu}V\right)\right]\right\}\end{aligned}\right\} \quad 3.64$$

Where $\left[1-\Gamma\left(\left(n_{L,i}-1\right),-\bar{\mu}V\right)\right]$ is the incomplete gamma distribution.

3.7.2 Non-Poisson Distributions

For a discrete distribution, the probability that a random variable assumes some value less than or equal to x is related by

$$F(x) = \sum_{i=x_L}^{x} p(i) \qquad 3.65$$

For a continuous distribution and its relation to probability density function (pdf)

$$F(x) = \int_{x_L}^{x} f(z)dz \qquad 3.66$$

Where x_L is the lower limit of support for the distribution. Alternative distributions are mainly based on Poisson distribution; it's a mixture distribution. A mixture distribution can be derived as follows from a Poisson distribution which provides much more flexibility and a greater variability in the expected count among replicates:

$$P_M(x;V,\beta) = \int_0^{\infty} P_P(x;\mu V)\, h(\mu;\beta)\, d\mu \qquad 3.67$$

Where $P_P(x;\mu V)$ derivation from the Poisson distribution is, μV is mean density in a sample and $h(\mu;\beta)$ is the mixing distribution describing the variability in means density with distribution parameter. $h(\mu;\beta)$ can take any form for the description of variability as described in (Haas et al., 2014).

Negative Binomial: (Greenwood et al., 1920) derivation of negative binomial as a gamma mixture of Poisson distribution, deriving the negative binomial with mean density in a sample as a mixture of gamma distribution

$$P_{NB}(x)=\int_0^{\infty}\frac{\left(\bar{\mu}V\right)^x}{x!}\exp\left(-\bar{\mu}V\right)\left[\frac{\theta(\theta x)^{k-1}}{\Gamma(k)}\exp(-\theta x)\right]d\mu \qquad 3.68$$

Evaluated analytically as

$$P_{NB}=\frac{\Gamma(x+k)}{\Gamma(k)x!}\left(\frac{\theta V}{1+\theta V}\right)^x(1+\theta V)^{-k}=\frac{\Gamma(x+k)}{\Gamma(k)x!}\left(\frac{\bar{\mu}V}{k+\bar{\mu}V}\right)^x\left(\frac{k+\bar{\mu}V}{k}\right)^{-k} \qquad 3.69$$

Poisson Lognormal: Describing the variability with log-normal distribution (Reid, 2012) which has been used to fit species-abundance and bibliometric data (section 3.7.1) will lead to

$$P_{LN}(x)=\int_0^{\infty}\frac{\left(\bar{\mu}V\right)^x}{x!}\exp\left(-\bar{\mu}V\right)\left[\frac{1}{x\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]\right]d\mu \qquad 3.70$$

Solving it numerically with the Gauss-Hermite quadrature (William et al., 1989)will lead to

$$\left.\begin{aligned}P_{LN}(x)&=\frac{1}{x!\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left(-q^2\right)\exp\left[-\varphi(q)V\right]\left[\varphi(q)V\right]^x dq\\ \varphi(q)&=\exp\left(\mu+sq\sqrt{2}\right)\end{aligned}\right\} \qquad 3.71$$

Poisson-Inverse Gaussian: the inability of analytically expressing the integral of the Poisson log-normal makes data fitting somewhat difficult, inverse Gaussian is a potential replacement for the lognormal due to its properties and its being positively

skewed, when the Inverse-Gaussian is used for mixing distribution for a Poisson the results follows:

$$P_{PIG}(x)=\int_0^{\infty}\frac{\left(\bar{\mu}V\right)^x}{x!}\exp\left(-\bar{\mu}V\right)\left[\left[\frac{\lambda}{2\pi x^3}\right]^{1/2}\exp\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right]d\mu \qquad 3.72$$

Integrated to yield the following (54)

$$P_{PIG}(x)=\frac{e^{\phi}}{x!}\left(\bar{\mu}V\right)^x\left[\phi\left(\phi+2\bar{\mu}V\right)\right]^{1/4}\left(\frac{\phi}{\phi+2\bar{\mu}V}\right)^{x/2}K_{x-1/2}\sqrt{\phi\left(\phi+2\bar{\mu}V\right)} \qquad 3.73$$

Where $K(x)$ is a modified Bessel function of the third kind.

3.7.3 Empirical Distribution Data Fitting

Let $X_1,...,X_n$ be a random sample following the ordered statistics $X_1<X_2<...<X_n$. The empirical cumulative distribution function (CDF) $F_n(x)$ is given as;

$$F_n(x)=\begin{cases}0, & x<X_{(1)}\\ \frac{1}{n}, & X_{(i)}\le x\le X_{(n)}\\ 1, & x\ge X_{(n)}\end{cases} \qquad 3.74$$

This represents a positively skew continuous step function, the statistics measuring the differences of $F(x)$ and $F_n(x)$ gives the Empirical Distribution function (EDF).

Kolmogorov-Smirnov statistics (KS) compares the EDF with the fitted distribution function defined by;

$$KS = \begin{cases} D^{+} = \sup_x \left\{ F_n(x) - F(x) \right\} \\ D^{-} = \sup_x \left\{ F(x) - F_n(x) \right\} \end{cases} \qquad 3.75$$

Hence the KS statistics D is defined as $D = \sup_x \left| F_n(x) - F(x) \right| = \max\left(D^{+}, D^{-} \right)$. The quadratic statistics is defined by $Q = n \int_{-\infty}^{\infty} \left\{ F_n(x) - F(x) \right\}^2 \psi(x)\, dF(x)$. It is noted that, when $\psi(x) = 1$, then the function $Q = n \int_{-\infty}^{\infty} \left\{ F_n(x) - F(x) \right\}^2 dF(x)$ is ***the Cramer-von Mises Statistics*** and when $\psi(x) = \left[\left\{ F(x) \right\} \left\{ 1 - F(x) \right\} \right]^{-1}$, then the function turns to be ***Anderson-Darling Statistics*** (AD).

3.8 Handling Scarce Data and the Principle of Maximum Entropy

Practically, in risk assessment modelling, statistics known about the data is either scarce, insufficient and in most cases of detection of pathogen methods applied makes it undetectable. Handling such data sources also call for the use of maximum entropy in characterizing the uncertainty of such dataset.

The measure of uniformity of a distribution is by its entropy, thus, the higher the entropy, the higher the uniformity (Harremoës and Topsøe, 2001). The maximum entropy consists of selecting the most uniform distribution of a set of possible distributions, that's the one with maximum entropy. Given a set of observation

$$\Omega = \left\{ x_i \middle| i = 1, 2, \ldots N \right\}$$

For a distribution p on χ, the set of all distributions on χ by Δ is defined as

$$\Delta = \left\{ p : \chi \to R^{+} \middle| \sum_{x \in \chi} p(x) = 1 \right\} \qquad 3.76$$

The empirical distribution of sample is denoted by $\bar{p}$ namely $\bar{p} = \frac{\sum_{i=1}^{N} \chi_x(x_i)}{N}$ for $x \in \chi$

Defining $\chi_x(x) = \begin{cases} 1 \text{ if } x=1 \\ 0 \text{ otherwise} \end{cases}$. The concept of feature function (thus expectation). It is a non-negative value function on χ, hence the expectation of the feature function f respect to the distribution p by

$$E_p[f] = \sum_{x \in \chi} f(x) p(x) \qquad 3.77$$

Hence having the set of feature function s $\{ f_i | i = 1, 2, \ldots N \}$. Defining the distribution p which is a subset c of Δ as $c = \left\{ p \in P \middle| E_p[f_i] = E_{\bar{p}}[f_i] \text{ for } i = 1, \ldots, n \right\}$, this is called the

constraints for these equations (Singh and Setup, 2003). The maximum entropy dictates that, there should be a selection of distribution p^* of C that maximize entropy. Hence as an optimization problem, we are left with

$$p^* = \arg\max_{p \in c} \left(-\sum_{x \in \chi} p(x) \log p(x) \right) = \arg\max_{p \in c} H(p)$$

As seen earlier, the entropy is bounded below by zero and from above by $\log|\chi|$ with the uniform distribution on χ. Hence $H(p)$ is continuous, strictly convex, bounded function in Δ, moreover, C is bounded, closed, convex and non-empty subset of $R^{|\chi|}$ since $\bar{p} \in C$.

3.8.1 Relation of Maximum Entropy to Maximum Likelihood

We defined the log-likelihood of a model p with respect to the empirical distribution $\tilde{p}$ by the function(Coughlan, 2010)

$$L_{\tilde{p}}(p) = \log \prod_{x \in \chi} p(x)^{\tilde{p}} = \sum_{x \in \chi} \tilde{p}(x) \log p(x) \qquad 3.78$$

Given $\lambda \in R^n$

$$\left.\begin{aligned} L_{\tilde{p}}p\lambda &= \sum_{x\in\chi}\tilde{p}(x)\log p\lambda(x) \\ &= \sum_{x\in\chi}\tilde{p}(x)\log\left(\frac{1}{Z_\lambda}\exp^{\sum_{i=1}^{n}\lambda_i f_i(x)}\right) \\ &= \sum_{x\in\chi}\tilde{p}(x)\left(-\log Z_\lambda + \sum_{i=1}^{n}\lambda_i f_i(x)\right) \end{aligned}\right\} \quad 3.79$$

$$\left.\begin{aligned} &= -\log Z_\lambda + \sum_{i=1}^{n}\lambda_i E_{\tilde{p}}[f_i] \\ &= \psi(\lambda) \end{aligned}\right\} \quad 3.80$$

The maximum entropy principle is sometimes regarded as ***an ideal learning method*** that makes ***minimal assumptions*** in arriving at an estimate of a distribution learned from data. Finally, note that maximum likelihood is sometimes regarded as non-Bayesian because there is no explicit prior given.

3.8.3 Normalization and partition functions

The general approach for assigning probabilities where normalization is absorbed into the denominator is given as:

$$\exp^{-\lambda_0}\sum_{k=1}^{n}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]=1 \tag{3.81}$$

$$p(x_i)=\frac{\exp^{-\lambda_0}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]}{1} \tag{3.82}$$

$$=\frac{\exp^{-\lambda_0}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]}{\exp^{-\lambda_0}\sum_{k=1}^{n}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]} \tag{3.83}$$

$$=\frac{\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]}{\sum_{k=1}^{n}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right]} \tag{3.84}$$

$$=\frac{k(x_i)}{z(\lambda_1,\ldots,\lambda_m)} \tag{3.85}$$

Where $f_j(x_i)$ is a function of the random variable x_i reflecting what we know

$$k(x_i)=\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right] \tag{3.86}$$

Is the kernel and

$$z(\lambda_1,\ldots,\lambda_m)=\sum_{k=1}^{n}\exp\left[-\sum_{j=1}^{m}\lambda_j f_j(x_i)\right] \tag{3.87}$$

Is the normalizing facto, called the partition function (Harremoës and Topsøe, 2001). Probability assignment is completed by determining the Lagrange multipliers $\lambda_{j,} j = 1,...,m$ from m constraints which are function of the random variables

Examples of Maximum Entropy Given Constraints

1 Range Constraints

For a discrete case $R \in \{0,1,...,r_{\max}\}$

Lagrangian

$$L = -\sum p_i(x) \log p_i(x) + \lambda\left(\sum p_i(x) - 1\right)$$

Critical points

$$\frac{\partial L}{\partial p_i} = -\log p_i(x) - 1 + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum p_i(x) - 1 = 0$$

Solution $p_i = e^{\lambda - 1} = \frac{1}{N}$ ***Uniform distribution***

Practical example, suppose we know only three possible empirical values and .The maximum entropy

$$L = \max_{p_i}\left[-\sum_{i=1}^{3} p_i(x)\log p_i(x) + \lambda\left(\sum_{i=1}^{3} p_i(x) - 1\right)\right]$$

For first order conditions yield

$p_i = \exp[\lambda_1 - 1]$ for i=1,2,3 . $\lambda_1 = \log 3$

As expected, the maximum entropy probability assignment is a discrete uniform distribution with

$$p_i = \frac{1}{3}$$

2. Mean constraints $\mathbb{E}[R] = \sum_{x=0}^{\infty} xp(x) = \mu$

Lagrangian: $L = -\sum p_i(x)\log p_i(x) + \lambda_1\left(\sum p_i(x) - 1\right) + \lambda_2\left(\sum xp_i(x) - \mu\right)$

Critical points

$$\frac{\partial L}{\partial p_i} = -\log p_i(x) - 1 + \lambda + x\lambda_2 = 0$$

Solution

$$p(x) = e^{\lambda_2 x + \lambda_1 - 1}$$

$$= \mu e^{-x\mu}$$

Exponential distribution

Partition function approach

With knowing the support and the mean, the kernel is given as

$$z(\lambda_1) = \sum_{i=1}^{n} \exp[-\lambda_1 x_i]$$

And

$$p_i = \frac{k(x_i)}{z(\lambda_1)}$$

$$= \frac{\exp[-\lambda_1 x_i]}{\sum_{j=1}^{m} \exp\left[-\lambda_1 x_j\right]}$$

Where xi are the random empirical data, solving the constraints

$$\sum_{i=1}^{n} x_i p_i(x) - \mu = 0$$

$$\sum_{i=1}^{n} \frac{\exp[-\lambda_1 x_i]}{\sum_{j=1}^{m} \exp\left[-\lambda_1 x_j\right]} x_i - \mu = 0$$

This produces λ_1 to identify the probability constraints

Example: suppose we know a little more on the empirical values thus the support and the mean, n=3 and mean is 2.5

$$\mathbb{E}[R] = \sum_{x=1}^{3} xp(x) = 2.5$$

$$\sum_{i=1}^{3} \frac{\exp[-\lambda_1 x_i]}{\sum_{j=1}^{3} \exp[-\lambda_1 x_j]} x_i - 2.5 = 0$$

$$\sum_{x=1}^{3} xp(x) - 2.5 = 0$$

3. Mean and variance $\mathbb{E}[R] = \sum_{x=0}^{\infty} xp(x) = \mu$

$$\mathbb{E}\left[(R-\mu)^2\right] = \sum_{x=0}^{\infty} (x-\mu)^2 \, p(x) = \sigma^2$$

We find that

$$\frac{\partial \log z}{\partial \lambda_1} = \frac{3 + 2\exp[\lambda_1] + \exp[2\lambda_1]}{1 + \exp[\lambda_1] + \exp[2\lambda_1]} = 2.5$$

Solving gives

$\lambda_1 = -0.834 \quad p_1 = 0.116, p_2 = 0.268, p_3 = 0.616$

Lagrangian

$$L = -\sum p_i(x) \log p_i(x) + \lambda_1\left(\sum p_i(x) - 1\right) + \lambda_2\left(\sum x p_i(x) - \mu\right) + \lambda_3\left(\sum_{x=0}^{\infty} (x-\mu)^2 \, p(x) - \sigma^2\right)$$

$$p(x) = \exp^{\lambda 2x + \lambda 1 - \lambda 3(x-\mu)^2}$$

With density function

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)}{2\sigma^2}\right]$$ (in continuous distribution form for approximation)

The partition function approach

$$p(x) = \frac{\exp\left[-\lambda_2 (x-\mu)^2\right]}{\sum \exp\left[-\lambda_2 (x-\mu)^2\right]}$$

And the average empirical constraint is

$$\sum (x-\mu)^2 f(x) - \mu = 0$$

$$\sum (x-\mu)^2 \frac{\exp\left[-\lambda_2 (x-\mu)^2\right]}{\sum \exp\left[-\lambda_2 (x-\mu)^2\right]} - \mu = 0$$

So that $\lambda_2 = \frac{1}{2\sigma^2}$

Hence the density function for is $= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(x-\mu)}{2\sigma^2}\right]$ Gaussian distribution(Elod, 2013)

3.9 Modelling Population and Community Illness with Epidemiological Framework

Modelling illness incidence with epidemiological framework as described (Charles et al., 1999) and other mathematical epidemiological model techniques (Chiyaka et al., 2010; Li et al., 2011; Medlock, 2009; Andersson and Britton, 2000), describing the incubation distribution as a fraction of persons who first become ill as a result of exposure, defining the instantaneous rate of infection as $\beta(t)$ within a population N (Sartwell, 1995), hence the instantaneous rate of a person entering a pool of persons becoming infected is $\beta(t)N$ resulting in $\frac{dN}{dt} = -\beta(t)N$ (Williams, 1965). The general incubation distribution could be defined as a differential equation based on the cumulative fraction of persons who become ill before or on t days as an instantaneous rate of illness as a convolution (Haas et al., 2014) $Q(t) = \int_0^t \lambda\beta(\tau)N(\tau)\, f(t-\tau)\, d\tau$ where λ is the fraction of infected persons who becomes ill which is based on asymptotic forms for incubation time distribution (Williams, 1965).

Figure 3.14 illustrates a simple epidemiological model for diseases transmission to infected subjects

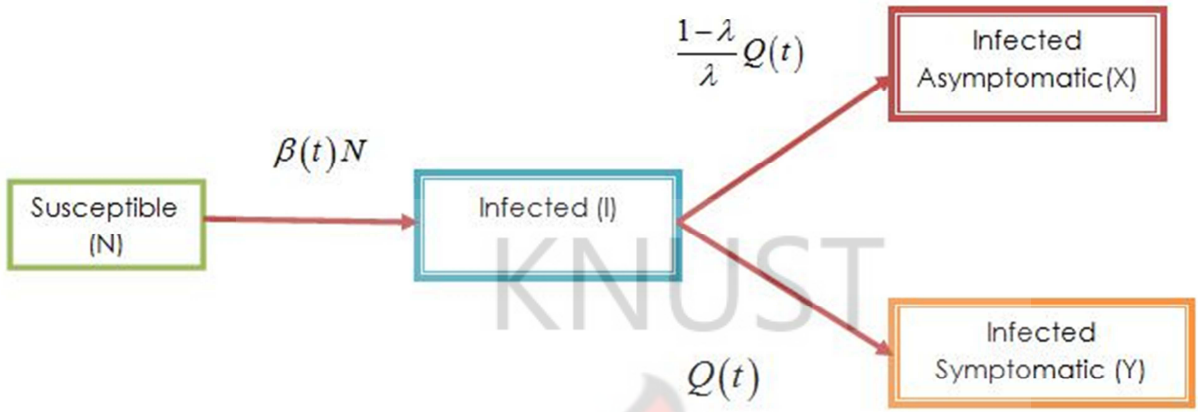


Figure 3.14: A simple epidemiological model for diseases transmission to infected subjects

From this concept, the simple model becomes (Fig. 3.14)

$$\left.\begin{aligned} \frac{dN}{dt} &= -\beta(t)N \\ \frac{dI}{dt} &= \beta(t)N - \frac{Q(t)}{\lambda} \\ \frac{dX}{dt} &= \frac{1-\lambda}{\lambda}Q(t) \\ \frac{dY}{dt} &= Q(t) \end{aligned}\right\} \quad 3.88$$

This allows the estimation of the rate at which cases appear as a function of instantaneous rate of infection. It must be noted that, the mode assumed that the

underlying incubation distribution for the conversion to the symptomatic state is the same as for the asymptomatic state. For a more complex modelling with the conversion of individuals to the post-infected state described by two different parameters, thus the number of symptomatic infected individuals per unit time and the number of asymptomatic individuals per unit time who enters the post-infected rate (Fig. 3.15) leading to the model.

Figure 3.15 shows a schematic epidemiological model for diseases transmission to post-infected subjects

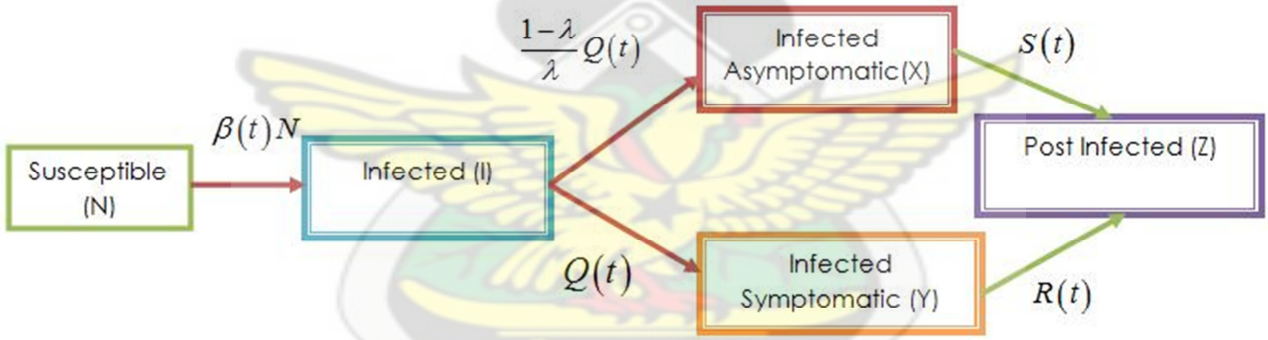


Figure 3.15: A schematic epidemiological model for diseases transmission to post-infected subjects

$$\left.\begin{aligned}
\frac{dN}{dt} &= -\beta(t)N \\
\frac{dI}{dt} &= \beta(t)N - \frac{Q(t)}{\lambda} \\
\frac{dX}{dt} &= \frac{1-\lambda}{\lambda}Q(t) - S(t) \\
\frac{dY}{dt} &= Q(t) - R(t) \\
\frac{dZ}{dt} &= S(t) + R(t)
\end{aligned}\right\} \quad 3.89$$

Hence defining the cumulative distribution $G(t)$ as a function of individuals existing the diseases state before or at time t with a density function $g(t)$ to account for duration of illness, hence

$$R(t) = \int_0^t \left[Q(t)\right] g(t-\tau)\, d\tau \quad 3.90$$

Again, defining the cumulative distribution $H(t)$ as a fraction of asymptomatic person from post-infected state to infected state and $h(t)$ is its derivative, hence

$$S(t) = \int_0^t \left[\frac{1-\lambda}{\lambda}Q(t)\right] h(t-\tau)\, d\tau \quad 3.91$$

(White et al., 1986) documentation of Norwalk virus in food borne outbreak shows that evidence exist of some organism illness may occur for both pre-infection as well as post-infection, others of such post-infection has also recorded (Ozawa et al., 2007;

Nordgren et al., 2010; Vega et al., 2011; Teunis et al., 2014; Mattison, 2011; Sukhrie et al., 2012) such a secondary occurrence is made by modification of the epidemiological framework, Kermack and McKendrick (1927) describe the occurrence as a product of susceptible and infected individuals, this indicates both symptomatic and asymptomatic serves as a point for infection within the time they are in their states, by extension of the model will lead to:

$$\left.\begin{aligned}
\frac{dN}{dt} &= -\beta(t)N - \gamma(t)N(X+Y) \\
\frac{dI}{dt} &= \beta(t)N + \gamma(t)N(X+Y) - \frac{Q'(t)}{\lambda} \\
\frac{dX}{dt} &= \frac{1-\lambda}{\lambda}Q'(t) - S(t) \\
\frac{dY}{dt} &= Q'(t) - R(t) \\
\frac{dZ}{dt} &= S(t) + R(t)
\end{aligned}\right\} \qquad 3.92$$

$$Q'(t) = \int_0^t \lambda\left\{\beta(\tau)N(\tau) + \gamma(\tau)N(\tau)\left[X(\tau)+Y(\tau)\right]\right\} f(t-\tau)\,d\tau \qquad 3.93$$

To account for immunity impact, it's worth making the assumption that post-infection does not occur, thus persons are no longer susceptible, defining the distribution $i(t)$ as a density function for residency in the post-infection state (immune), hence transition back for effect of waning immunity to susceptible state is by the convolution:

$$T(t)=\int_0^t\left[R(\tau)+S(\tau)\right]i(1-\tau)d\tau \tag{3.94}$$

$$\left.\begin{aligned}
\frac{dN}{dt}&=-\beta(t)N-\gamma(t)N(X+Y)+T(t)\\
\frac{dI}{dt}&=\beta(t)N+\gamma(t)N(X+Y)-\frac{Q'(t)}{\lambda}\\
\frac{dX}{dt}&=\frac{1-\lambda}{\lambda}Q'(t)-S(t)\\
\frac{dY}{dt}&=Q'(t)-R(t)\\
\frac{dZ}{dt}&=S(t)+R(t)-T(t)
\end{aligned}\right\} \tag{3.95}$$

3.9.1 Transmission of Pathogens in the Population

Pathogen transmission among individuals within the population depends on the shedding and transmission rate of exposed and infected individuals. Three different transmission modes are described to cater for all forms of pathogen transfer within the population.

- **Susceptible Population**

All individuals within the population are susceptible and hence becomes infectious, individuals within the susceptible compartment moves into the exposed compartment with rate (λ), from exposed an individual can then move either to the

infected asymptomatic compartment with rate (U_D) or move to infected symptomatic compartment with rate (U_s) and then move into the infected asymptomatic compartment with rate (U_A). From asymptomatic compartment an individual moves to recovery compartment with rate (λ) and hence can move back into the asymptomatic compartment again with rate (p), or waning immunity can make the individual becomes susceptible again and move to susceptible compartment with rate (θ), all compartments have a natural death rate of d, the differential equation for the transmission is given below:

$$\left.\begin{aligned}
\frac{dS}{dt} &= B+\theta R-(\lambda+d)S \\
\frac{dE}{dt} &= \lambda S-(U_s+U_D+d)E \\
\frac{dI_s}{dt} &= U_s E-(U_A+d)I_s \\
\frac{dI_A}{dt} &= U_A I_s+\lambda R-(p+d)I_A+U_A E \\
\frac{dR}{dt} &= pI_A-(\lambda+\theta+d)R
\end{aligned}\right\} \qquad 3.96$$

- **No Immune Boosting Transmission mode**

Individuals do not move back into the asymptomatic compartment, the only path is recovery and hence by waning immunity becomes susceptible again.

$$\left.\begin{aligned}
\frac{dS}{dt} &= B+\theta R-(\lambda+d)S \\
\frac{dE}{dt} &= \lambda S-(U_s+U_D+d)E \\
\frac{dI_s}{dt} &= U_sE-(U_A+d)I_s \\
\frac{dI_A}{dt} &= U_AI_s-(p+d)I_A+U_AE \\
\frac{dR}{dt} &= pI_A-(\theta+d)R
\end{aligned}\right\} \qquad 3.97$$

- **Genetic Resistance Individuals within the Population**

Not all exposed individuals within the population are susceptible, proportion (υ) of the individuals inherit a genetic resistance which makes them resistance to the pathogens, but they forms part of the population for transmission.

$$\left.\begin{aligned}
\frac{dG}{dt} &= (B-d)\upsilon \\
\frac{dS}{dt} &= B(1-\upsilon)+\theta R-(\lambda+d)S \\
\frac{dE}{dt} &= \lambda S-(U_s+U_D+d) \\
\frac{dI_s}{dt} &= U_s E-(U_A+d)I_s \\
\frac{dI_A}{dt} &= U_A I_s+\lambda R-(p+d)I_A+U_A E \\
\frac{dR}{dt} &= PI_A-(\lambda+\theta+d)R
\end{aligned}\right\} \quad 3.98$$

3.10 Summary

This chapter presented the different statistical theory principle ideas of the methodology used in this study, various probability distributions which forms the models under studies were presented as well as their usage in the determination of quantify pathogen concentration. The chapter also presented the handling of various forms of dataset that emerged as part of the studies which does not forms part of the quantitative nature of probability distributions and not enough to fit parametric distributions on them. This chapter has presented the various theoretical aspects of the distributions and data handling quantification used in the rest of the study as well as the epidemiological aspect of illness incidence within a population

with immunity. In the next chapter, probabilistic quantitative risk assessment model with genome copies and fecal indicator ratio conversion for dose estimation is presented.

CHAPTER 4

PROBABILISTIC QUANTITATIVE RISK ASSESSMENT MODEL WITH GENOMIC COPIES FOR DOSE ESTIMATION

The main objective of this chapter is to develop a probabilistic model for the risk assessment with both genome copies of virus/oocyst particles and the use of fecal indicator ratio conversion method to meet the objectives in Section 1.3. In this chapter a probabilistic model approach of modelling with genome/oocyst particles is presented in addition to ratio conversion method to assess the impact of the use of pathogen of interest in microbial risk estimation procedure. The daily probability of infection are determined as well as the annual estimation of risk. Daily Adjusted Life Years (Diseases Burden) has also been determined for consumer exposure to pathogens in wastewater. Various parameter estimations and sensitivity analysis are also carried out.

4.2 The Model and Its Analysis

This section presents the models formulation for quantifying a probable risk of infection or illness by accidental ingestion of *Norovirus* through consumption.

The models include dose ingestion, build-up of probabilistic approach of pathogen ingestion and its corresponding dose-response models, survival and hazard function of illness and characterization of probable risk to annual risk and the estimation of Daily Adjusted Life Years (DALYs).

4.2.1 Model formulation and Pathogen Concentration

The description of transmission of pathogen from water unto the vegetables then onto consumption is modelled as a farm to fork approach, irrigation with wastewater on farms for all crops including vegetables and salad crop eaten uncooked is termed as unrestricted irrigation, and poses a health challenge to both farmers and consumers as well. These stakeholders are directly exposed to the hazards in the wastewater and create a public health concern. The pathogen path for contact with its host is through a complex interaction method as shown in Figure 4.1.

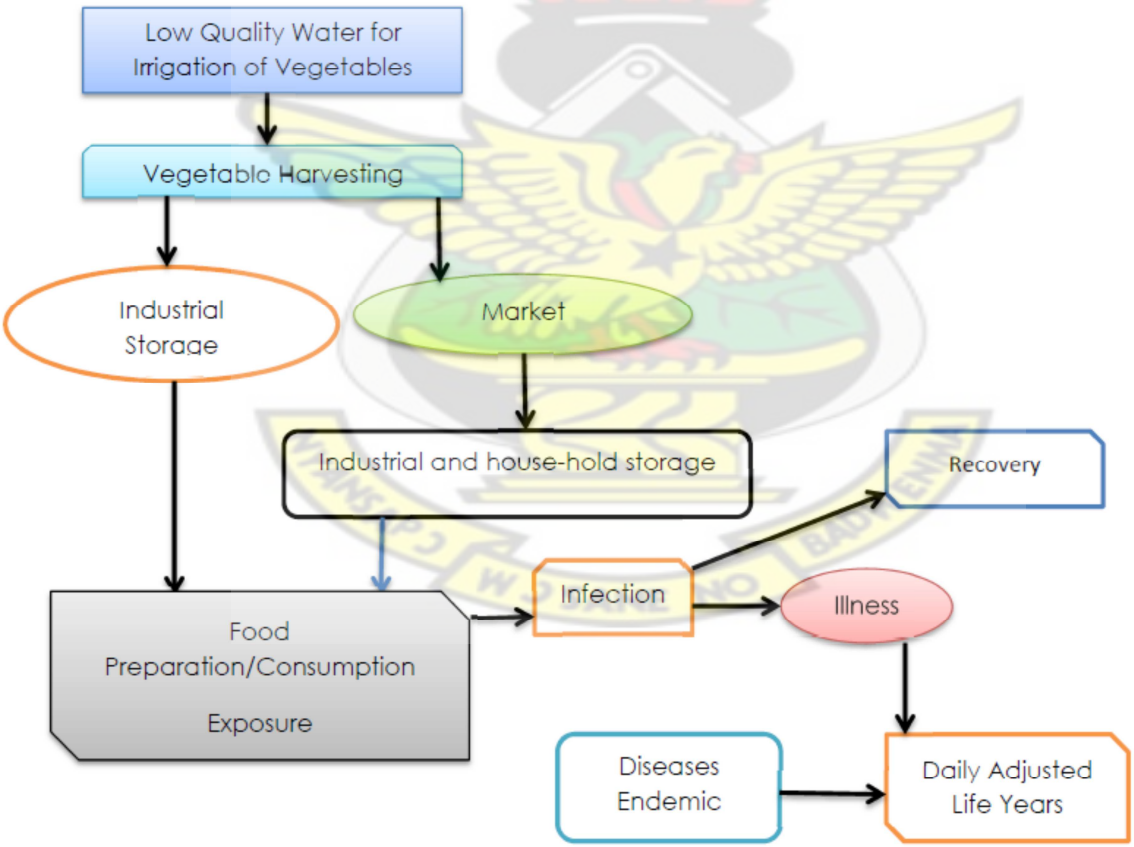


Figure 4.1: Schematic Exposure to wastewater through to estimating illness

The pathogen dose (virus genomic or particles/oocyst) that the consumer is exposed to on the kth day (d: number of virus/oocyst particles ingested per person per exposure) resulting from consumption of either salad (cabbage, lettuce or a salad crop) irrigated with wastewater. The water/produce model for modelling ingestion of pathogen from wastewater directly or indirectly (salad consumption) is given as

$$d = C_z * C_q * I * T_R * R_c * V_p * 10^{-w} * \exp(-kt) = \begin{cases} V_p \in \mathbb{R}, \text{ Water model} \\ V_p = 1, \text{ Produce model} \end{cases} \quad 4.1$$

Hence dose (d)is a joint probability distribution given as

$$d = P\left(\left(C_z, C_q, R_C, I, T, V_P, W, k, t\right) \in A\right) = \begin{cases} \sum_A f\left(C_z, C_q, I, T, R_C, V_P, W, k, t\right), \text{ discrete case} \\ \int_A \left(C_z, C_q, I, T, R_C, V_P, W, k, t\right) dA, \text{ continuos case} \end{cases} \quad 4.2$$

Where $C_z = f\left(x; a, b\right) = \frac{1}{b-a}$, *for* $a \le x \le b$ (section 3.5.18) is the daily consumption of vegetable per person ($gperson^{-1}day^{-1}$)(Fung, 2011), $C_q = P_M\left(x; V, \beta\right) = \int_0^\infty P_P\left(x; \mu V\right) h\left(\mu; \beta\right) d\mu$ is the mixture distribution of concentration of pathogen in irrigation water/on vegetable produce (no/ml)(Rice, 2001), $R_c = p\left(r; k, p\right) = \binom{r-1}{k-1} p^k \left(1-p\right)^{r-k}$ (section

3.5.13) is the recovery methodology of pathogen concentration (no/ml)(Petterson et al., 2007), I is the percentage of infection of virus, T_R is the transfer rate of virus from the irrigation water to the produce in the case of water model, we assume in this work, for water model all virus genome copies/oocyst detected in the irrigation water were transfer to the produce for the worst case approach, and half of the Norovirus genome copies and all oocyst are infectious.

$V_p = \begin{cases} \frac{1}{b-a}, \; for \;\; a \le x \le b, \; cabbage \\ f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma} \right)^2 \right], \;\; lettuce \end{cases}$ is the volume of irrigation water caught by product (mg^{-1}) (Mok et al., 2014; Mok and Hamilton, 2014; Shuval et al., 1997), $w = Pert(a,b,c) = Beta(\alpha_1, \alpha_2, a, c)$ (Section 3.5.17) is the pathogen reduction by pre-consumption preparation ($\log_{10} unit$, $k = f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma} \right)^2 \right]$ (section 3.5.14) is the pathogen kinetic decay constant (per day) and $t = \frac{1}{b-a}, \; for \;\; a \le x \le b$ (3.5.18) is time between last waste water irrigation event and harvesting vegetable/storage (days).

4.2.2. The Dose Response Model Formulation

Building a mathematical dose response relationship is establishing the relations between level of pathogen, exposure and the measure of likelihood occurrence of

adverse effects. The two sequential sub processes for estimating the level of infection is as namely(Haas et al., 2014);

i. One or more organism or virus genomic copies (particles) ingested is (are) capable of causing diseases
ii. Ingested organism/virus/oocyst particle undergo decay or inactivated to multiply to cause infection/disease by host susceptible responses, and only a fraction of the ingested organism reach a site where infection can begin by breaking all barriers within the body immune system.

The two measures are to ensure that, for an infection to occur within a susceptible host organism or individual, there is at least a surviving organism breaking all the mechanism of the defensive nature of the an individual, be it fully protected, partially protected or immune-compromised susceptible. Infection can only occur when there is a surviving pathogen to cause infectious foci within a cell of an individual(Furumoto and Mickey, 1970; Haas et al., 2014b; Mickey and Furumoto, 1970).

Hence, ingestion precisely j organism from exposure to wastewater contaminated with a pathogen of a mean dose d is expressed as j given d(Mickey and Furumoto, 1970):

$$P_1(j|d) \qquad 4.3$$

Continually, the probability of k surviving organism or pathogen particles within the ingestion j organism to initiate an infection process can also be expressed as;

$$P_2(k|j) \tag{4.4}$$

Assuming independency for the two processes (thus the ingestion of precisely j organism and the surviving pathogens remain to initiate infection process), the probability of k organisms surviving to initiate infection by breaking all defense mechanism within the body is given by the independent event (Furumoto and Mickey, 1970; Furumoto and Ray, 1967; Haas et al., 2014b):

$$P(k) = \sum_{j=1}^{\infty} P_1(j|d) P_2(k|j) \tag{4.5}$$

The least number of organism ($k_{\min}$) surviving to initiate an infection leads to a probability of infection (Furumoto and Mickey, 1970; Furumoto and Ray, 1967; Haas et al., 2014b).

$$P_I(d) = \sum_{k=k\min}^{\infty} \sum_{j=k} P_1(j|d) P_2(k|j) \tag{4.6}$$

Where ($k_{\min}$) is *not* the minimal infection dose or threshold needed to be reached to cause an infection, however the average inoculate dose required to cause half of

the subjects to experience a response of infection(Furumoto and Mickey, 1970a; Mickey and Furumoto, 1970b).

4.2.2.1 Exponential Dose Response Model

Characterising the distribution of organism between each dose as random and assuming independency for each ingested organism, this has an identical survival probability[2] r and that $k_{\min}=1$ (thus for a single hit assumption). Hence for Poisson distribution of organism(Furumoto and Mickey, 1970a; Mickey and Furumoto, 1970b).

$$P_1(j|d)=\frac{d^{j}}{j!}e^{-d} \qquad 4.7$$

Modelling survival means of organism to cause an infection with binomial distribution leads to

$$P_2(k|j)=\frac{j!}{k!(j|k)!}(1-r)^{j-k} \qquad 4.8$$

Hence, substituting equation 4.7 and 4.8 into equation 4.5 leads to equation 4.9 (Furumoto and Mickey, 1970a; Haas et al., 2014; Mickey and Furumoto, 1970b)

[2] This is a probability that an organism survives all barriers of defense mechanisms and initiate an infectious focus within cell

$$\left.\begin{aligned} P_I(d) &= \sum_{k=k\min}^{\infty} \sum_{j=k} \left[\frac{d^j}{j!} e^{-d}\right]\left[\frac{j!}{k!(j|k)!}(1-r)^{j-k}\right] \\ &= \sum_{k=k\min}^{\infty} \frac{(dr)^k e^{-dr}}{k!} \sum_{j=k}^{\infty} \frac{\left[d(1-r)\right]^{[j-k]} e^{(-d[1-r])}}{(j-k)!} \end{aligned}\right\} \quad 4.9$$

But $\sum_{k=k\min}^{\infty} \frac{(dr)^k e^{-dr}}{k!} = 1$, hence

$$P_I(d) = 1 - \left[\sum_{k=0}^{k\min-1} \frac{(dr)^k e^{-dr}}{k!}\right] \quad 4.10$$

With the earlier single hit assumption (thus one organism survived is capable to cause an infection) $k_{\min} = 1$ yields (Mickey and Furumoto, 1970)

$$P_I(d) = 1 - e^{-rd} \quad 4.11$$

Where d is the dose subjected to individuals and r is the infectivity rate of the pathogen of interest or under study. Hence given a mean dose of d from ingestion precisely j organism with k organism surviving to initiate an infection, the exponential dose response model is as given (equation 4.11).

4.2.2.2 The Beta-Poisson Dose Response Model

Replacing equation 4.8 with a mixture distribution with respect to the parameter r to account for variability in the interaction probability yields equation 4.12 (Furumoto and Mickey 1967)

$$P_2(k|j) = \int_0^1 \left[\frac{j!}{k!(j-k)!} (1-r)^{j-k} r^k \right] f(r)dr \qquad 4.12$$

From equation 4.11 and applying a mixture operation directly, for assuming a variation in the dose to dose for the Poisson distribution, then

$$P_{\inf}(d) = \int_0^1 \left[1 - e^{-rd} \right] f(r)dr \qquad 4.13$$

$$\left.\begin{aligned} &= \int_0^1 f(r)dr - \int_0^1 e^{-rd} f(r)dr \\ &= 1 - \int_0^1 e^{-rd} f(r)dr \end{aligned}\right\} \qquad 4.14$$

Again, accounting for the variation between doses to dose, a great deal of flexibility is the use of beta distribution, hence incorporating the beta distribution into equation 4.10 yields (Furumoto and Ray, 1967)

$$P_{\inf}(d) = 1 - \int_0^1 \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta-1} \right] e^{-rd} dr$$

The integral can be expressed as confluent hyper-geometric written as a series expansion

$$\int_0^1 e^{-rd} dr = e^{-d} \int_0^1 e^{d(1-r)} f(r) \qquad 4.15$$

$$= e^{-d} \int_0^1 \sum_{j=0}^{\infty} \frac{a^j}{j!} (1-r)^j \, f(r) \, dr$$

$$= e^{-d} \sum_{j=0}^{\infty} \frac{a^j}{j!} \int_0^1 (1-r)^j \, f(r) \, dr$$

$$= e^{-d} \sum_{j=0}^{\infty} \frac{a^j}{j!} \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta+j-1} \, dr$$

$$= e^{-d} \sum_{j=0}^{\infty} \frac{a^j}{j!} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha)\Gamma(\beta+j)}{\Gamma(\alpha+\beta+j)} \qquad 4.16$$

$$= e^{-d} \left\{ 1 + \frac{\beta}{(\alpha+\beta)} d + \frac{\beta(\beta+1)}{(\alpha+\beta)(\alpha+\beta+1)} \frac{d^2}{4} + \ldots \right\}$$

$$= e^{-d} {}_1F_1(\beta, \alpha+\beta; d)$$

Therefore,

$$\int_0^1 \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} r^{\alpha-1}(1-r)^{\beta-1} \right] e^{-rd} dr = {}_1F_1(\beta, \alpha+\beta; d) \qquad 4.17$$

The ingestion of virion genome copies/oocyst is based on whether it is aggregated or disaggregated, as virion genome copies/oocyst particles may or may not be aggregated.

Hence in a confluent hyper-geometric function (Furumoto and Ray, 1967; Mickey and Furumoto, 1970)

$$P_{\inf}(d) = 1 - {}_1F_1(\alpha, \alpha+\beta, -d) \qquad 4.18$$

Furumoto and Mickey (1967) derived the following expression approximation to Hyper geometric function (equation 4.16) based on the certain valid parameter values, thus when $\beta \gg 1$ and $\alpha \ll \beta$ the simple relation holds.

$$P_I(d) = 1 - \left(1 + \frac{d}{\beta}\right)^{-\alpha} \qquad 4.19$$

4.2.2.3 Modelling Illness Resulting from Infection

Given the presence of a high pathogen, colonization of a host with ≥ 1 surviving organism does not necessarily lead to illness, however, there is an increase in the probability of the host defense measure to fight off the colonization of the surviving organisms, this leads to two ingredients for the risk of acute illness, thus the hazard of illness and the duration of infection as described by Teunis et al., (1999).

Definition 4.1

Suppose T is a non-negative random variable representing the time until an event of interest happens. Let assume T is a continuous variable unless otherwise specify, then the probability density function (pdf) and its cumulative distribution function (cdf) are used to characterise the random variable distribution and denoted as:

$$\left.\begin{array}{l} pdf: f(t) \\ cdf: F(t) = P(T \le t) \end{array}\right\} F(0) = P(T = 0)$$

Therefore, the survival function (the probability that the event of interest has not yet occurred by time t) is defined as;

$$S(t) \overset{def}{=} 1 - F(t) = P(T > t), \forall t > 0 \qquad 4.20$$

And the corresponding hazard function is also defined as;

$$h(t) \overset{def}{=} \lim_{h \downarrow 0} \frac{P\left[t \le T < t + h \middle| T \ge t\right]}{h} = \frac{f(t)}{S(t_-)}$$

Where $S(t_-) = \lim_{h \uparrow t} S(s)$, thus hazard function is the conditional density, given that the event of intereste has not yet occurred prior to time t. For continuous time T

$$h(t) = -\frac{d}{dt} \ln\left[1 - F(t)\right] = -\frac{d}{dt} \ln S(t)$$

Let the cumulative hazard function defined as;

$$H(t) \overset{def}{=} \int_0^t h(u)\,du, \quad t > 0$$

$$\begin{aligned} &= -\ln\left[1 - F(t)\right] \\ &= -\ln S(t) \end{aligned} \qquad 4.21$$

Therefore,

$$S(t) = \exp\left(-H(t)\right)$$
$$f(t) = h(t)\exp\left(-H(t)\right)$$

Let infection at initial stage $(0 \le t < \Lambda)$ and inactivation of pathogen within host at given a hazard function $h(t)$ where Λ is the time until infection occurs. Moreover, the commonest distribution for describing the survival function of illness be the exponential distribution function with scale parameter η, then

$$f(t) = \eta\exp\left[-\eta\Lambda\right] \text{ and } h(t) = \eta$$

$$H(t) = P\left(ill \middle| \inf : u\right) = 1 - \exp\left[-\int_0^t \eta\,dt\right], \quad 0 \le t < \Lambda$$

then the probability that infection results in illness can be written as the cumulative function

$$P\left(ill \middle| \inf; \Lambda\right) = F(t) = 1 - \exp\left(-\eta\Lambda\right) \qquad 4.22$$

Assuming the duration Λ is described by the gamma distribution to reflect individual differences in resistance of host colonization of pathogen, gamma density distribution is (Teunis et al., 1999)

$$g(\Lambda;\omega,d)=\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)$$

$$P(ill|\inf)=\int_{r=0}^{\infty}\left[1-\exp(-\eta\Lambda)\right]\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]d\Lambda=1-(1+\eta d)^{-\omega}$$

$$P(ill|\inf)=\int_{r=0}^{\infty}\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]d\Lambda-\int_{r=0}^{\infty}\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda(1+\eta)}{d}\right)\right]d\Lambda$$

$$=1-(1+\eta d)^{-\omega} \qquad 4.23$$

The probability that illness occurs is

$$P(ill|\inf)=\begin{cases}P(ill|\inf)=\int_{r=0}^{\infty}\left[1-\exp(-\eta\Lambda)\right]\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]d\Lambda=1-(1+\eta d)^{-\omega}, Norovirus\\ d(D,x:A,B,C)=AD-Bx+C=1-(1+AD-Bx+C)^{-\omega}, cryptosporidium\ oocyst\end{cases} \qquad 4.24$$

Hence from equation 4.24, the probability of illness per dose per person is calculated as

$$Pill = P_{ill|\inf}\,P_{I(d)} \qquad 4.25$$

4.2.2.3 Annual Risk of infection/illness and Diseases Burden

The annual risk of infection and illness are determined with the frequency of exposures (*n*)of an individual within a year and were modelled following the independent assumption of Karavarsamis and Hamilton (2010)· The annual risk of infection or illness *P was* estimated as

$$P = 1 - \prod_{k=1}^{n} [1 - P_k] \qquad 4.26$$

Where $P_k$is the k^{th} median probability of infection or illness per exposure event in n total exposures within a year. Estimating the annual disease burden using the Disability Adjusted Life Year (DALY, $person^{-1}year^{-1}$) metric (Barker, 2014; Barker et al., 2014; Barker et al., 2013; Mok and Hamilton, 2014; Mok et al., 2014), which is used to measure all disease burden expressed as the number of years lost due to disability, illness or premature death (Mara and Sleigh, 2010, 2009; Mara et al., 2010). Hence equation 4.26 is given as

$$DALY = P_{ill\sim k} BS \qquad 4.27$$

Where $P_{ill\sim k}$is the annual probability of illness per given dose per person and $B = \frac{1}{b-a}$, *for* $a \le x \le b$(section 3.5.18) is the disease burden (DALY per case of diarrhoea) and *S* is the proportion of population susceptible to the diseases.

4.3 Modelling Scarce, Imprecise Data and Expert's Opinion for Uncertainty Quantification in QMRA Model

Parameters input values for modelling in QMRA presents a challenge to the entire model in itself, accounting for most of the uncertainty pertaining to the either the model or the scenarios. Lack of data, imprecise data and interval assumptions are characterised by ambiguity and imprecise linguistic description of events rendering estimation of parameters based on combining experts' opinion and limited data available.

4.3.1 The Evidence Theory

The theory of evidence, popular referred to us Dempster-Shafer Theory (DST) is based on two principal elements developed by Dempster (1976) and Shafer (1976) (Benavoli et al., 2009), is solely based on belief and plausibility measures and are characterised by a function m, called *basic probability assignment* (credibility)(bpa).

A body of evidence induce credibility on a class $p(X)$ of all possible subset of X, with the assumption that, evidence resides upon an empty set.

Thus $m_i : p(X) \to [0,1]$ and

1. $m(\phi) \geq 0$

2. $\sum_{A \in p(X)} m_i(A) = 1$

The class of the focal subset of X corresponding to m_i is $F_i = \{A | A \subseteq X, m_i(A) > 0\}$

A belief measure and plausibility are defined on a universal set X as a function of mapping a power set to a range $[0,1]$(Salmona, 2014), where plausibility measure is the dual of the belief measure

$$\left.\begin{aligned} Bel \equiv m : P_X \rightarrow [0,1] \\ Pl(A) = 1 - Bel(A) \end{aligned}\right\} \quad 4.28$$

Given, 'A' a family of subset of X, then the belief, plausibility and ambiguity of 'A' by mi is defined as (Ayyub and Klir, 2006).

$$Bel_i(A) = \sum_{B \subseteq A, B \neq `0} m_i(B) \quad 4.29$$

$$Pl_i(A) = \sum_{B \cap A \neq `0} m_i(B) \quad 4.30$$

$$Amb_i(A) = \sum_{B \cap A \neq `0, B \not\subset A} m_i(B) \quad 4.31$$

4.3.2 Aggregation of Evidence and Conflict

Definition 4.2. (Simple Evidence)

A simple evidence denotes the case when the bodies of evidence are mutually exclusive, and it induce a probability distribution

Definition 4.3. (Mixed Evidence)

A pair of dependent bodies of evidence of experts dependently induce a joint probability distribution, thus $m_{ij}: P(X)P(X) \to [0,1]$. *If the body of evidence are independent, then* $m_{ij}(A,B) = m_i(A)m_j(B)$. *If* $m_j(B) > 0$, *then the conditional probability distribution on* $P(X)$ *given B is defined as* $m_{i|j}(A|B) = m_{ij}(A,B)/m_j(B)$. *The corresponding class of focal pairs of subsets is* $F_{i|j} = \{(A,B) | A \subseteq X, B \subseteq X, m_{ij}(A,B) > 0\}$

Let's m_1 and m_2 be basic probability assignments to the same element from two experts, then Dempster's rule of combination to obtain a combined opinion $(m_{1,2})$ as

$$m_{1,2}(A_i) = \frac{\sum_{allA_j \bigcap A_k = A_i} m_1(A_j) m_2(A_k)}{1 - \sum_{allA_j \bigcap A_k = \phi} m_1(A_j) m_2(A_k)} \qquad 4.32$$

Where A_i must be a nonempty set and $m_{1,2}(\phi) = 0$, the term $1 - \sum_{allA_j \bigcap A_k = \phi} m_1(A_j) m_2(A_k)$ becomes the normalizing factor for the contradiction or conflicts among the two experts (Sentz, 2002; Wierman, 2001; Zadeh et al., 1978). Generally, suppose we have evidences and we want to fuse this into forming a single body of evidence and assume all collected evidence concern the same universe X. Let evidence be $A_i = \langle F_i, m_i \rangle, i = 1,2,3,...$ and K be conflict among evidence, if $A_j \in F$ then $\bigcap_{\forall i} A_i = \phi$. Then

total conflict as described with the Dempster rule of combination is given as $K = \sum_{\bigcap_{\forall i} A_i = \phi} m_i(A_i)$. Hence focal set is $F = \left\{ \bigcap_{\forall i} A_i \middle| \bigcap_{\forall i} A_i, A_i \in F_i \right\}$. The basic probability for $C \in F$,

$$m(C) = \frac{\sum_{\bigcap_{\forall i} A_i = C} m_i(A_i)}{1 - K} \quad 4.33$$

Where C is the combination of the different expert's opinion, and K is the accounting for the conflict among the overlaps of the experts' opinion. Hence

$$m(C) = \frac{\sum_{\bigcap_{\forall i} A_i = C} m_i(A_i)}{1 - \sum_{\bigcap_{\forall i} A_i = \phi} m_i(A_i)} \quad 4.34$$

4.3.3 Weighting Evidence Assignment

A body of evidence induces a probability distribution on the class $m(A)$ of all possible opinion on the subject A. Assuming residence of evidence for combination is not equally credible and any contradiction are not taking into account expressed by (Bae et al., 2004).

$$m_{1,2,3,\ldots,n}(A_i) = \frac{1}{n} \sum_{k=1}^{n} w_k m_k(A_i) \quad 4.35$$

Definition 4.4

A family of weights is probabilistic if they satisfy the equalities $\sum_{C\in P(X)} w_i(C|A)=1\ \forall A\in F(X;m_i)$, *where both C and A are a collections of suspects. 'A' must be a focal set of body of evidence, but C is arbitrary.*

The weight factors are assigned based on the credibility of the evidence and its source. Hence relaxing the (Bae et al., 2004; H.-R. Bae et al., 2004) (equation 4.35) assumption and incorporating into Dempster rule (equation 4.34) leads to

$$m_{1,2,3,\ldots n}(C)=\frac{\sum_{allA_j\bigcap A_k=A_i} w_1m_1(A_j)\ldots w_nm_n(A_k)}{1-\sum_{allA_j\bigcap A_k=\phi} w_1m_1(A_j)\ldots w_nm_n(A_k)} \qquad 4.36$$

4.4 Model Implementation for Fecal Indicator Ratio Conversion and Virus Genomic copies for Dose Estimation through Vegetable Consumption in Ghana

The study includes the input parameters of different variables such as pathogen concentration in wastewater, which are predominantly the first hand sources of water available for irrigation in developing countries. Amoah et al., (2007) described the different sources of water primarily used for irrigation in developing countries, considered as wastewater mainly from streams, drains and partially treated wastewater from Waste Stabilization Ponds (WSP). This study uses data from these sources of water (streams, drains wastewater) and Waste Stabilization Pond (WSP)

effluent, however, the WSP on pond effluent involves the usage of both the influent and effluent discharge waters used for irrigation by farmers.

Lettuce and cabbage were selected as the main vegetables for the study to represent crops commonly eaten raw in developing countries such as Ghana and other African countries. These crops do not form part of the traditional cuisines in households, but are major component associated with street foods (Fung, 2011).

Information on actual volumes of consumed vegetables in Ghana is scarce and various QMRA studies (Ackerson and Awuah, 2012; Barker, 2014; Seidu et al., 2008) have therefore used estimates of salad consumption. Fung (Fung, 2011) reported that salad mainly consisted of lettuce and cabbage (> 75%) with a salad serving size of 20g per meal. This meal size is higher than the estimated value of 10 g – 12 g of lettuce per meal per day (Seidu et al., 2008), in Ghana there is lack of comprehensive study on salad servings contaminated from *Norovirus or Cryptosporidium spp*, hence all servings were assumed to be contaminated as a worst case scenario. The estimated value for consumption data (C_z)is combined through experts' opinion, a uniform distribution (Figure 4.2) was fitted to cater for the different portion sized found in earlier studies and again, a uniform distribution (Figure 4.3) was also fitted for all year round frequency (total exposure) of consumption of vegetable.

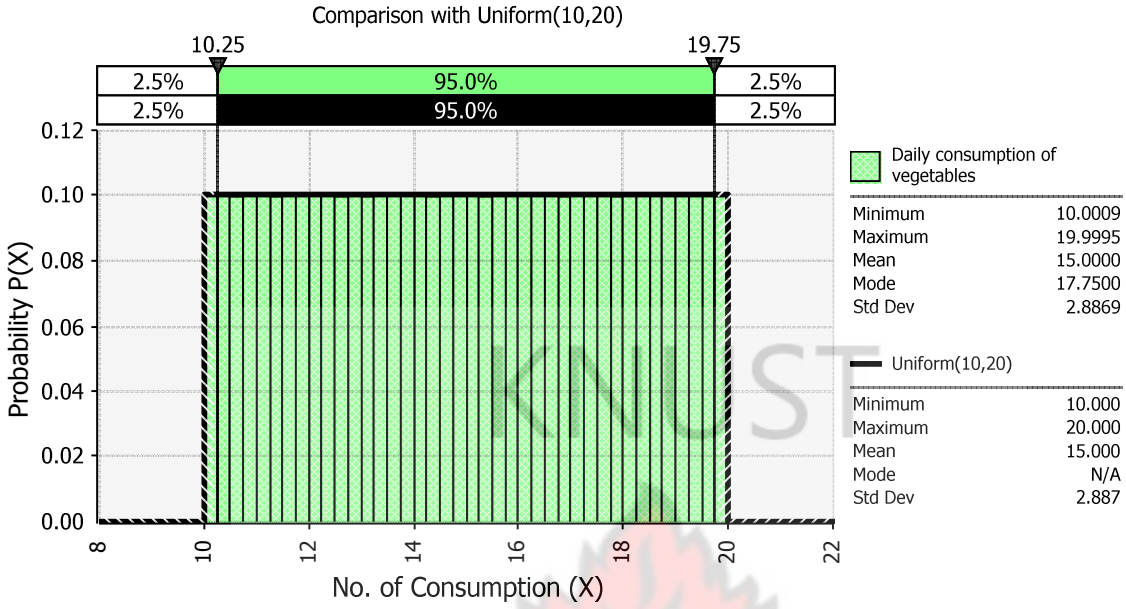


Figure 4.2: Fitted Distribution for Salad Consumption size

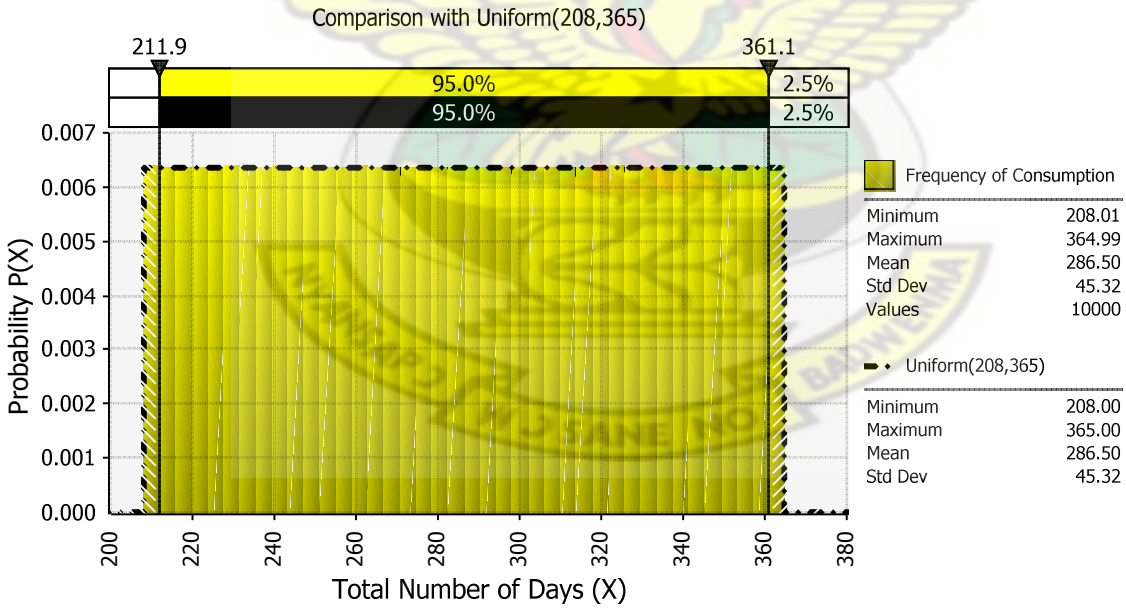


Figure 4.3: Fitted Distribution for Total Consumption Frequency

Volume of water (V_p) caught on the surface of Asian vegetables (Mok et al., 2014) was used as an approximate estimate for the study, as previous studies in Ghana have shown that such values are appropriate(Barker et al., 2014). Uniform distribution of water was used for cabbage (Figure 4.4) and normal distribution truncated at zero was used for lettuce (Figure 4.5) to characterize the volume of water detained by the two vegetables.

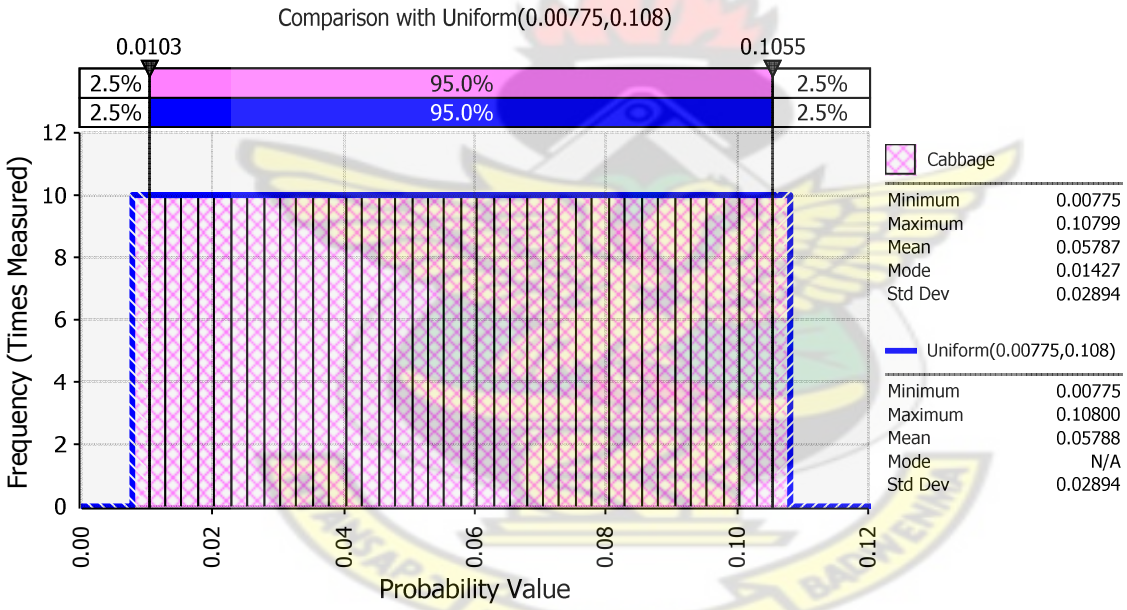


Figure 4. 4: Fitted Uniform Distribution for Wastewater caught on Cabbage

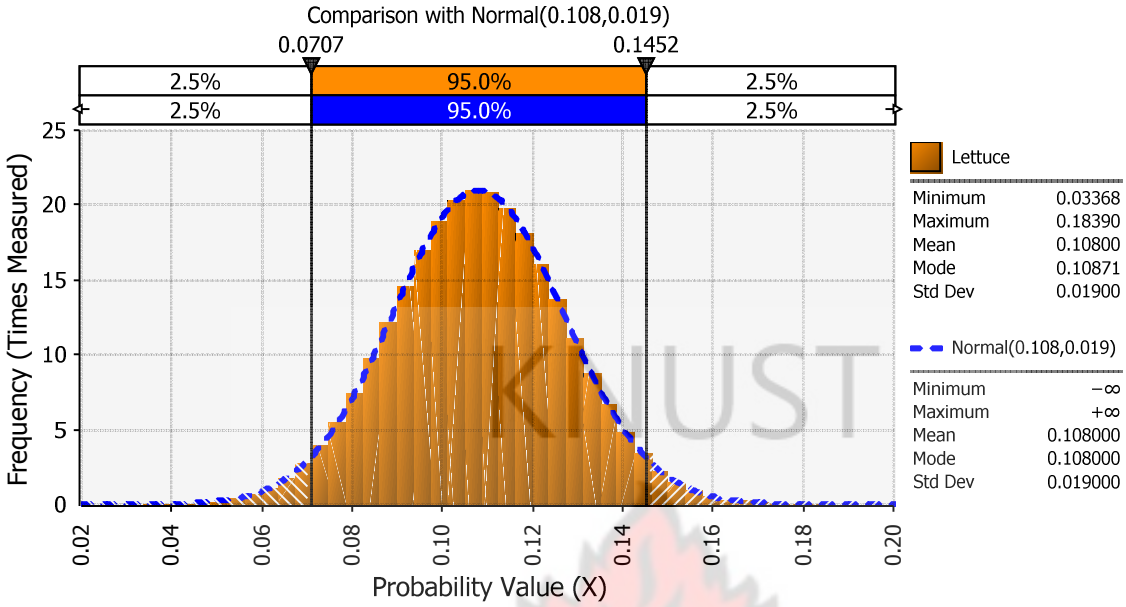


Figure 4.5: Fitted Normal Distribution for Wastewater Caught on Surface of Lettuce

The kinetic in-field decay constant (k) was fitted to normal distribution (Figure 4.6) with data used in previous studies (Barker et al., 2013; Mok et al., 2014; Hamilton et al., 2006a). However, cabbage and lettuce are perishable and consumption of these products is usually done soon after harvest. Hence, post-harvest virus decay beyond 48 hours was considered insignificant and was not included.

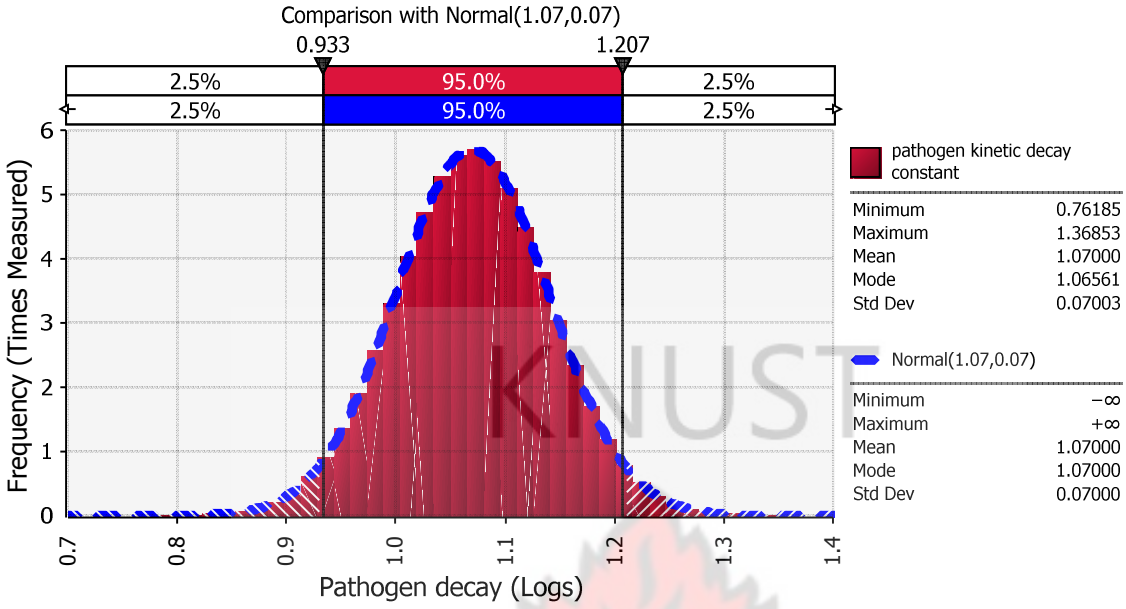


Figure 4.6: Fitted Distribution for Pathogenic Kinetic Decay

Time for withholding water (t) was assumed to be within 0 to 2 days after irrigation. A uniform distribution (Figure 4.7) was fitted to cover zero to a maximum of two days as vegetables in hot climatic conditions as in Ghana must be irrigated frequently, typically daily, to keep fresh(Seidu et al., 2008).

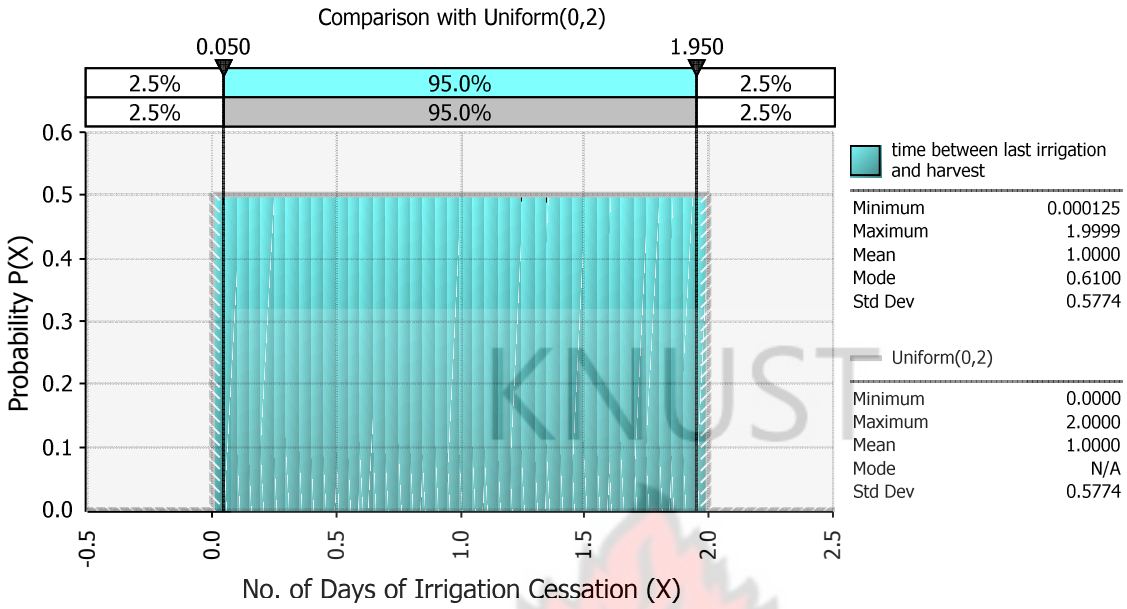


Figure 4.7: Fitted Distribution for Period of Irrigation Cessation before Harvest

Washing of vegetables during preparation is common practice in Ghanaian food stalls and households (Amoah et al., 2007; Fung, 2011; Seidu et al., 2008). Although reports on varying degree of efficiency of bacterial removal by washing and disinfection are available(Amoah and Drechsel, 2007; WHO, 2006b), similar information on reduction of viruses are scarce. Allwood and Malik, (2004) pointed out that viruses may be more resistant than bacteria during washing, and Norovirus is no exception (Mattison, 2011) as enteric viruses are known to be resistant to the environment as well. The distribution fitted by Barker et al., (2014) for washing vegetable was used as the modal value for the PERT distribution (Figure 4.8) in this study though it referred to bacterial reduction. Nevertheless, previous studies indicated that these values can be applied to virus as well (Ayuso-Gabella et al.,

2011; Barker et al., 2013; Mara. , Sleigh, 2010; Mok and Hamilton, 2014; Seidu et al., 2008) .

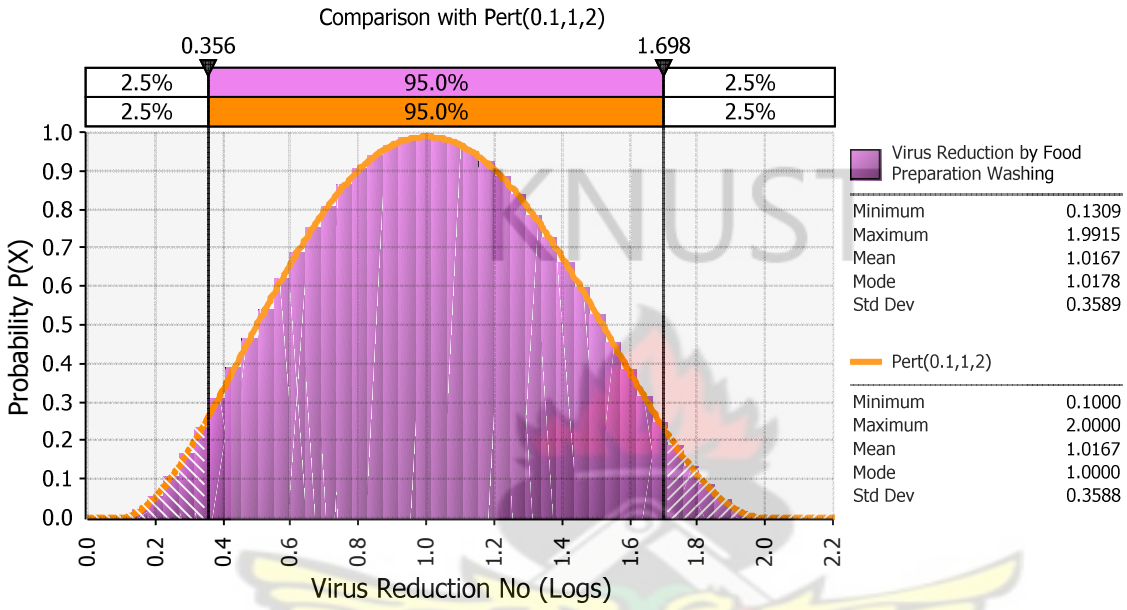


Figure 4.8: Fitted Distribution for Pre-Consumption Vegetable Preparation

All input parameters are reported in Table 4.I as well as Norovirus published data for modelling (Table 4.2).

Table 4.I: Distributions of input parameters

Parameter	Notation	Units	Distribution type (Value)~[mean]	Reference
Diarrheal Burden of Diseases in Ghana	B	$person^{-1}year^{-1}$	Uniform (1.06E-04-6.23E-03)~[3.16E-03]	Begg et al. 2007; Haagsma et al. 2008;
Salad consumption	C_Z	$gday^{-1}$	Uniform(10-20) ~[14.1]	Seidu et al. 2008; Fung 2011(Fung, 2011);
Frequency of consumption	n	$day(year^{-1})$	Uniform (208-365)	Seidu et al. 2008; Mok et al. 2014.
Volume of irrigation water caught by product	V_P	ml/g		
Cabbage		ml/g	Uniform(0.00775,0.108) ~[0.0580]	Mok et al. 2014; Barker et al. 2013; Hamilton et al. 2006; Shuval et al. 1997
Lettuce		ml/g	Normal(0.108,0.019)~truncated at zero~[0.108]	Mok et al. 2014; Barker et al. 2013; Hamilton et al. 2006; Shuval et al. 1997
Pathogen kinetic decay constant	k	day^{-1}	Normal(1.07,0.07)~truncated at zero~[1.07]	Barker et al. 2013; Hamilton et al. 2006; Petterson 2001; Petterson 2002

Time for withholding or irrigation cessation	t	*days*	Uniform (0,2)~[1.0]	Barker et al. 2013
Post-harvest/Food Preparation Washing for Virus reduction	w	$\log_{10}$ *units*	Pert(0.1,1.0,2.0)~[1.0]	Mok et al. 2014; Baert et al. 2009; Baert & Uyttendaele 2008; Ndiaye et al. 2011(Ndiaye et al., 2011); Croci et al. 2002(Croci et al., 2002); Mitakakis 2004(Mitakakis, 2004).

Table 4.2: Norovirus/Cryptosporidium spp and fecal indicator in raw sewage/Distributions and fit parameters used in models

Dose Response Parameters	**Units**	**Distribution type (Value)~[mean]**	**Reference**
Norovirus dose-response parameters for a+b inoculums[3]	Hyper geometric Beta Poisson	$\alpha = 0.040$, $\beta = 0.055$, $a = 0.997$, $\eta = 0.00255$, $\omega = 0.086$	(Teunis et al., 2008)(Teunis and Havelaar, 2000)
Cryptosporidium spp response parameters	$\omega = 1.10$, $A = B = 0$, $C = 0.77$, $r = 0.00419$		(Teunis et al., 2002)

Norovirus Data	Indicator Org.(CFU/mL)	Norovirus (gc/mL)	Virus Recovery	Ratio of Means	Min ratio	Max. Ratio
Drain	$10^2 - 10^6$	1.85×10^2	25%−50%	1.76×10^{-4}	1.85×10^{-5}	1.60×10^{-1}
Stream	$10^2 - 10^5$	1.03×10^2	25%−50%	7.13×10^{-4}	1.03×10^{-4}	9.90×10^{-2}
Pooled	$10^2 - 10^6$	1.64×10^2	25%−50%	1.56×10^{-4}	1.64×10^{-5}	1.60×10^{-1}

Reference:(Hassine-Zaafrane et al., 2014), (Katayama et al., 2008),(Haramoto et al., 2006), (Silverman et al., 2013), (La Rosa et al., 2010)(Flannery et al., 2012)

[3] Maximum likelihood estimates for the combined dose response models with and without virus aggregation and dose response model with no aggregation applied to susceptible subjects.

4.5 Simulation Results

4.5.1 Estimation of Annual Probability of infection and illness of gastroenteritis

The annual probability of infection for *Norovirus* ranged from 9.2×10^{-1} to 9.4×10^{-1} for all genome copies *Norovirus* while the ratio conversion method also ranged from 8.8×10^{-1} to 9.1×10^{-1}. Again, the annual probability of diseases[4] at a given infection ranged from 8.6×10^{-1} to 9.0×10^{-1} for the genome copies *Norovirus* and 8.1×10^{-1} to 8.3×10^{-1} for ratio conversion *Norovirus* (Table 4.3). Moreover, the ratio conversion for *cryptosporidium spp.* was found to be 2.3×10^{-3} and the oocyst *cryptosporidium spp.* data annual probability of infection was 4.9×10^{-1}, yet the annual probability of disease given infection were 1.5×10^{-3}, 2.7×10^{-1} for *E.coli* conversion and oocyst data respectively. Using the USEPA's threshold of 10^{-4} annual probability of infection and the recommended 10^{-6} risk of infection by Signor and Ashbolt (2009) daily risk target, all model scenario exceeded the thresholds, hence vegetables irrigated cannot be said to be safe for consumption.

[4] All servings of salad were assumed to be contaminated

Table 4.3: Annual probability of gastroenteritis infection an/illness and diseases burden (DALYs) per person per year

Model Scenarios (Norovirus Analysis)	Annual probabilities		Diseases Burden (DB: DALY pppy)		
	Infection	Diseases/infect	5thPercent'	Median	95thPercent'
Stream genome copies Norovirus	9.4×10^{-1}	8.6×10^{-1}	8.0×10^{-6}	4.1×10^{-5}	1.2×10^{-4}
Stream Norovirus Ratio	9.1×10^{-1}	8.1×10^{-1}	6.8×10^{-9}	$1.2\times10^{-7*}$	5.8×10^{-7}
Drains genome copies Norovirus	9.3×10^{-1}	9.0×10^{-1}	1.3×10^{-6}	1.8×10^{-5}	8.7×10^{-5}
Drains Norovirus Ratio	8.8×10^{-1}	8.2×10^{-1}	8.7×10^{-10}	$1.4\times10^{-8*}$	7.1×10^{-8}
Pooled genome copies Norovirus	9.2×10^{-1}	8.9×10^{-1}	8.2×10^{-6}	6.7×10^{-5}	1.7×10^{-4}
Pooled Norovirus Ratio	9.1×10^{-1}	8.3×10^{-1}	2.6×10^{-7}	3.7×10^{-6}	1.7×10^{-5}
Cryptosporidum spp Analysis					
E. coli conversion	2.3×10^{-3}	1.5×10^{-3}	1.5×10^{-7}	3.7×10^{-6}	3.7×10^{-5}
Oocyst Data	4.9×10^{-1}	2.7×10^{-1}	2.4×10^{-5}	6.6×10^{-4}	3.8×10^{-3}
Ratio Conversion distribution values					
Norovirus stream	$10^{[normal(4.45.0.86),truncatedat\ 3.3\ and\ 7.5]}/100-2.35x103b$				
Norovirus Drain	$10^{[normal(4.35.1.06),truncatedat\ 3.2\ and\ 7.0]}/100-2.49x103b$				
Norovirus Pooled	$10^{[normal(4.30.1.04),truncatedat\ 3.1\ and\ 7.2]}/100-2.40x103b$				
Cryptos oocyste	Lognormal(0.002,0.003)				

b is the mean from 3,650,000 iterations of pathogen concentration data

4.5.2 Annual Diseases Burden

The median annual diseases burden for ranged from 1.8×10^{-5} to 6.7×10^{-5} for all genome copies *Norovirus* (stream, drain and pooled data) concentration with 95th percentile values ranged from 8.7×10^{-5} to 1.7×10^{-4} while for ratio conversion *Norovirus* (stream, drain and pooled data), the median annual diseases burden ranged from $1.4\times10^{-8*}$ to 3.7×10^{-6} with 95th percentile ranges from 7.1×10^{-8} to 1.7×10^{-5}.

All scenarios (stream and drain data) using the genome copies *Norovirus* estimation for diseases burden (DALY) were ≥ 2 orders of magnitude higher than the use of ratio conversion *Norovirus* method of translating fecal indicator to *Norovirus*. It should be noted that only scenarios involving ratio conversion *Norovirus* achieved the health target of less than 1×10^{-6} DALY pppy, whereas scenarios involving the use of genome copies *Norovirus* data were ≤ 1order of magnitude less than the DALY health target of 1×10^{-4} DALY pppy (Figure 4.9). When pooled data were used for both genome copies and ratio conversion Norovirus, the median annual diseases burden for both achieved the DALY of 1×10^{-4} but not 1×10^{-6} (Figure 4.10).

On the part of *cryptosporidium spp.*, the median DALY diseases burden ranged from 1.5×10^{-7} to 3.7×10^{-5} for ratio conversion and 2.4×10^{-5} to 3.8×10^{-3} for oocyst data, representing the 5th and the 95th percentile for each respectively. Again all scenarios for the *cryptosporidium spp.* for the oocyst data DALY were close to ≥ 2 order of magnitude higher than the ratio conversion method (Figure 4.11).

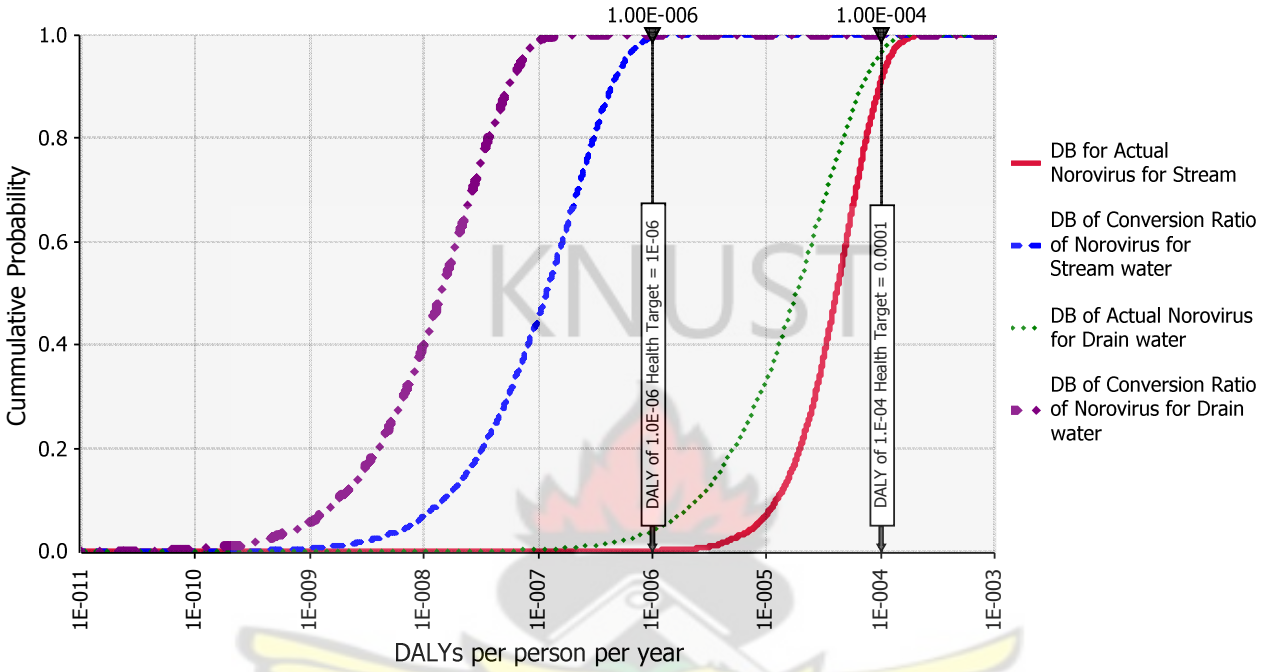


Figure 4.9: Cumulative Probability Curve of Daily Adjusted Life Years (Diseases Burden) for Stream and Drain wastewater for Actual *Norovirus*[5] and ratio conversion. Each cumulative probability represents the diseases burden for either the actual *norovirus* genome copies dose estimation or the conversion ratio dose estimation for stream water and drain water.

[5] Actual Norovirus: genome copies *norovirus* dose

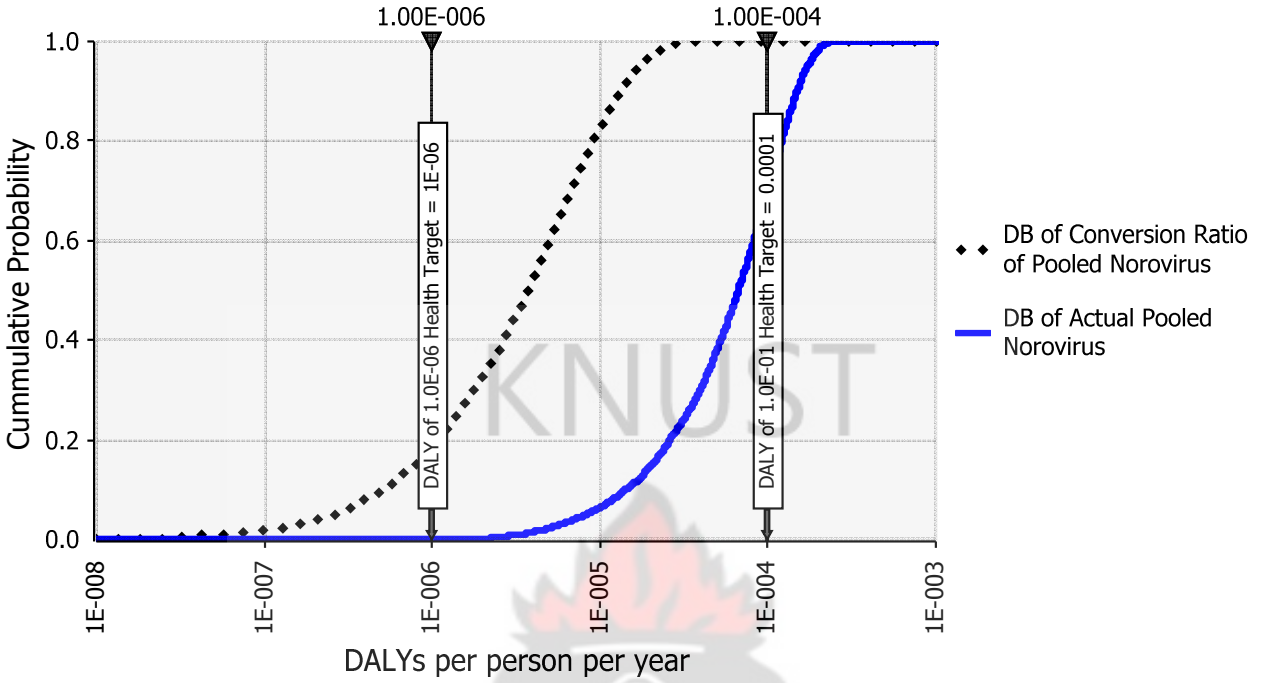


Figure 4.10: Cumulative Probability Curve of Daily Adjusted Life Years (Diseases Burden) for Pooled Data. This represents the combined pool data for stream and drain water, the cumulative probability graph shows the differences of using either the actual genome copies pool data or that of ratio conversion estimation for dose in modelling risk assessment.

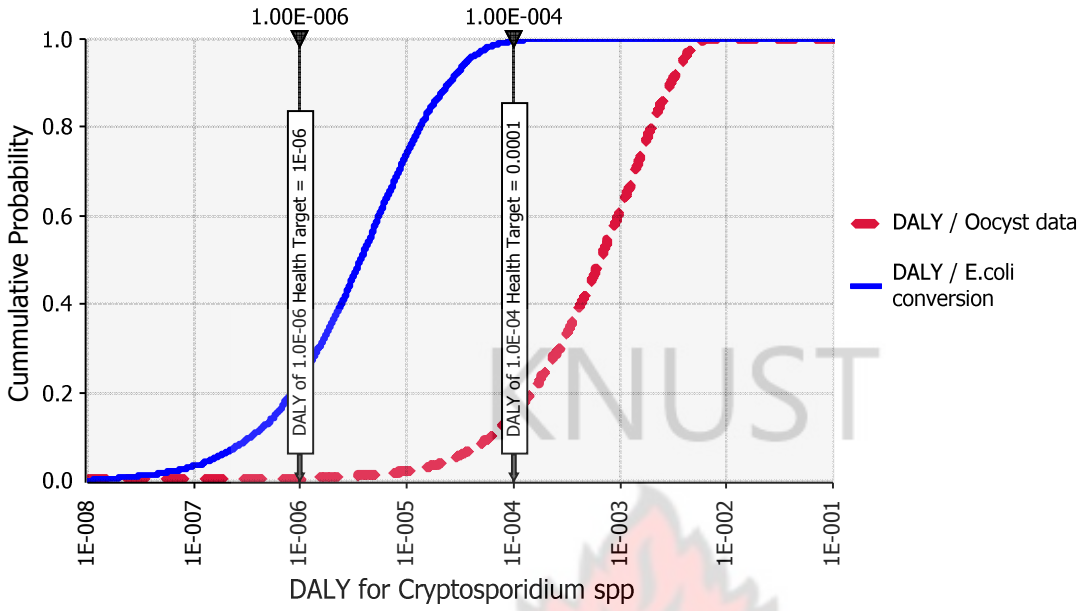


Figure 4.11: Cumulative Probability Curve of Daily Adjusted Life Years (Diseases Burden) for *Cryptosporidium*. This represents a comparison of E.coli conversion and that of oocyst data studies.

4.5.3 Sensitivity Analysis

In this section, sensitivity analysis was used to ascertain the model parameters that have strongest impact on the model output. This helps to account and to establish if one needed Hazard Analysis Critical Control Points (HACCP) to mitigate adverse effects of some essential input parameters. Different methods available for sensitivity include factorial design, sensitivity index, one-at-a-time, differential sensitivity analysis, the importance index, the spearmen, the Pearson etc. (Hamby, 1995, 1994)

In this study, Spearman's' was used to determining the relative change in various input parameters for estimation of probability of illness and for that matter the Daily Adjusted Life Years. As shown in Table 4.4 the probability of illness was very sensitive directly to water quality, volume of irrigation water caught by the vegetable, daily consumption of vegetable, it was somehow less sensitive to the kinetic decay constant. But it recorded an inversely sensitivity to the virus reduction due to food preparation and the time between the last irrigation and harvest (cessation of irrigation).

On the part of cryptosporidium spp., there is a strong relation to water quality, and diseases burden, kinetic decay constant and other weak positive correlation with volume of irrigation water caught on surface of vegetable as well as consumption frequency. Whiles there's a negative influence from pre-consumption food preparation and cessation of irrigation (Table 4.5).

Table 4.4: Spearman's rank order correlation coefficients for probability of illness (*Norovirus*)

Parameters	**Correlation Coefficient (Spearman Rank)**
Water Quality	0.62
Volume of irrigation water caught by vegetable	0.51
Kinetic decay Constant	0.04
Virus Reduction by Food Preparation	-0.24
Cessation of irrigation	-0.118

Table 4.5: Spearman rank order correlation coefficients for DALYs (*Cryptosporidium*)

Parameters	**Correlation Coefficient (Spearman Rank)**	
	Oocyst Data	**E.coli Conversion**
Water Quality	0.64	0.72
Volume of irrigation water caught by vegetable	0.09	0.38
Kinetic decay Constant	0.38	0.38
Virus Reduction by Food Preparation	-0.43	-0.55
Cessation of irrigation	-0.03	-0.05
Consumption Size	0.09	0.14
E. coli conversion	-	0.12
Diseases Burden	0.52	0.52
Consumption Frequency	0.08	0.10

4.6 SUMMARY

The probabilistic stochastic model presented to model exposure for dose estimation of *Norovirus* and *Cryptosporidium* with vegetable consumption irrigated with low quality water indicates the disease burden of the different model scenarios was found to be acceptable under different thresholds of DALY. Silverman et al., (2013) reported that sufficient Norovirus data in Ghana were not available to fit a distribution (11 quantifiable samples), however, it was indicated that the figures are conservative estimates and might be a few orders higher due to analytical challenges. Again the use of empirical model to characterize the genome copies data due to its insufficiency to fit a parametric distribution also contributes to uncertainty. This supports the result of this study to the effect that, the estimates for DALY in the case of stream and drain wastewater might be some orders higher than what is reported in this study and hence serve as conservative estimates.

In the scenarios presented here, none of the models using genome copies of *Norovirus* nor *cryptosporidium spp*. to predict the diseases burden found that it could establish the safety of consuming the produce i.e. the threshold of $\leq 10^{-6}$pppyDALY was not met. In contrast, the use of ratio conversion met the threshold for the same model in the case of *Norovirus.* The WHO guideline states that "if the overall burden of diseases from other exposures is very high, setting a less stringent level of acceptable risk of 10^{-5}DALY per person per year or 10^{-4}DALY per person per year may be more realistic as was argued by Mara and Hamilton (Mara and Hamilton, 2010) "This

assertion of WHO may guide the results of accepting the burden of diseases level for all model scenarios used in this study. On the other hand, the ratio conversion method currently applied to estimate the diseases burden produces significantly lower estimate of DALY with 2 or more orders of magnitude lower than the use of genome copies *Norovirus* concentration data or oocyst *cryptosporidium* data. It should be noted that, differences in diseases burden for stream and drain were significant for Norovirus, yet, both achieved the threshold of the health target of $10^{-4} DALYppp$, with the estimation of diseases burden in drain wastewater being less than that of stream water, whiles in case of cryptosporidium, the conversion ratio meets the threshold but the oocyst data does not.

With emphasis placed on the differences of order of magnitude in DALYs as a result of the use of fecal indicator ratio conversion in estimating health risk in various QMRA models, (Payment and Locas, 2011) argued that, the use of *E. coli* as indicator of fecal pollution does not represent well the presence of protozoa and other pathogen microorganisms. These Indicators are useful for monitoring hygiene such as in slaughter plants, but a high level fecal indicator does not necessarily mean a high level of pathogens, as this will depend on the infectivity level of the source. On the part of (Silverman et al., 2013), " while the ratio of NV GII to E. coli or thermotolerant coliform is likely to differ over place and time and may include animal fecal sources as well as environmental sources and reservoirs, it is an important finding that the current assumption of 0.1 – 1 Norovirus particles per 10^5 *E. coli* would underestimate

virus dose with exposure to wastewater and surface water sample". Again, "if standard pathogen concentrations are to be used effectively, there should be a move away from indicator species such as *E. coli* toward the pathogens of interest such as viruses as put forward by Mok and Hamilton (2014).

This study shows that, a move away from using fecal indicator conversion rates can lead to more realistic risk estimation as shown clearly with ≥ 1 order of magnitude higher when *cryptosporidium* oocyst or genome copies *Norovirus* particle concentration is used, though the values are considered conservative and by no means represent the total comprehensive *Norovirus* concentration in streams, drains and WSPs as reported by (Silverman et al., 2013) due to factors such as the technique applied for the quantification and the insufficient number of samples used to characterize the concentration. Moreover, the unavailability of aggregation data for quantification of risk in dose response model might contributes as a model uncertainty. Still, it gives a basis for a virus interest health risk assessment based on the concentration of genome copies of human Norovirus and a corresponding fecal ratio conversion in order to established specific health based targets.

CHAPTER 5

INTEGRATED IMMUNITY DOSE RESPONSE MODELING FOR ILLNESS REDUCTION INCIDENCE OF NOROVIRUS

In this chapter, dose-response models are formulated and an extension is made on the integration of immunity in risk assessment modelling with the use of Fractional-Poisson dose response function. A detailed derivation of various Dose Response Incidence(DRI) models are presented as a function of induced temporary immunity on exposed individuals in order to obtain the effect of accounting for immunity in dose response. Epidemiological results and compartment studies for transmission dynamics of *Norovirus* have been incorporated to achieve a comprehensive incidence of illness on exposed individuals in risk assessment. The effects of the various temporary immunity induced DRI models are determined integrating the transmission dynamics scenarios. Simulated results on the transmission dynamics, illness inflation factor (the protective effect of induced immunity to compensate for future infection and disease transmission), partial and full immunity loss as well as effects of exposure duration are presented. The models are applied on vegetable consumers' exposure to *Norovirus* and the results are presented to estimate the effect of immunity based on the transmission dynamics scenarios for individual illness incidence.

5.2 The Model and its Analysis

In this section, the Fractional Poisson Model for estimating the risk of ingestion of Norovirus to account for the likelihood of individuals' heterogeneous response to infection is presented along with the modelling results. Temporary immunity models are also shown to take into account the loss of partial and full immunity and how this affects the estimation of the probable risk of illness.

5.2.1 The Fractional Poisson Model

A reference is made to equation 4.18 (Section 4.2.2.2) for estimating a response to dose inoculums. Accounts from Teunis et al., (2008) for quantification of probability of illness assume that individual complete virus genomic copies or particles ingested by each human subject share a common probability (r) of independently initiating infection in subjects. Moreover, under the beta-Poisson model r is a mixed distribution (equation 4.15) indicating, some subjects may have very small values of r, thus infection probabilities per individual virion near zero and vice versa. Hence, the aggregated Norovirus infection probability is a beta function with parameters described as[6]

$$\begin{aligned} P_{\inf}(dose,\alpha,\beta) &= 1-{}_2F_1\left(\alpha,,d(1-a)/a,\alpha+\beta;-a/(1-a)\right) \\ &= 1-{}_1F_1\left(\alpha,\alpha+\beta,-d\right) \end{aligned} \qquad 5.1$$

Where $\mu(a)$ is of the form

[6] The beta-poisson function in this case is the second order hyper geometric function which diverges at dose exceeding 342 genome copies of the virus particles replacing the first order hyper geometric function.

$$\mu(a) = -a/((1-a)\ln(1-a)) \tag{5.2}$$

Hence

$$P_{\inf}(d) = \begin{cases} e^{\frac{-d}{\mu(a)}}, d = 0 \\ \\ 1 - e^{\frac{-d}{\mu(a)}}, d \geq 1 \end{cases} \tag{5.3}$$

Assumption of probability of infection (Messner et al., 2014)across susceptible population $P_2(k|j) = \int_0^1 \left[\frac{j!}{k!(j-k)!}(1-r)^{j-k} r^k \right] f(r)dr$ (equation 4.14), leads to

$$P_{\inf}(d) = \begin{cases} \int_0^1 \left[\frac{j!}{k!(j-k)!}(1-r)^{j-k} r^k \right] f(r)dr \\ \\ r \sim Bernoulli \end{cases} \tag{5.4}$$

$$p(r,p) \sim Bernoulli = \begin{cases} 1-p, if \ \mathrm{r} = 0 \ (failure) \\ \\ p, \ if \ \mathrm{r} = 1 \ (\mathrm{Success}) \end{cases} \tag{5.5}$$

Where p is the proportion of susceptible individuals.

Thus individuals are either perfectly susceptible or perfectly protected against infection and cannot be anywhere in between in equation 5.5, hence, from 5.3 with perfectly susceptible leads to:

$$P_{\inf}(d,P) = 1 - e^{\frac{-d}{\mu(a)}} \tag{5.6}$$

Simulate Results for Probability of Infection with Fractional Poisson Dose Response

Simulated results for equation 5.6 shows a steady progression of probability of infection (Figure 5.1). The lower infection probability is less than 1 log of dose (genome copies) exhibiting a monotonic increment as dose ingestion increases. The absoluteness of infection was not reached at the highest dose of 8 logs of genomic copies. However, it should be noted that dose ingestion alone is not the only factor determining probability of infection. It depends also on exposure frequency. For instance, very frequent exposure may lead to a continuous ingestion of pathogens and hence to a higher probability of infection. Naturally, when exposure leads to infection (either symptomatic or asymptomatic) there is an associated acquired immunity for the protection of individuals who are exposed frequently to the pathogens. Such acquired immunity may explain the reasons behind the non-contract of illness/diseases of individuals frequently exposed to wastewater and henceforth the need to integrate such acquired immunity in risk assessment models.

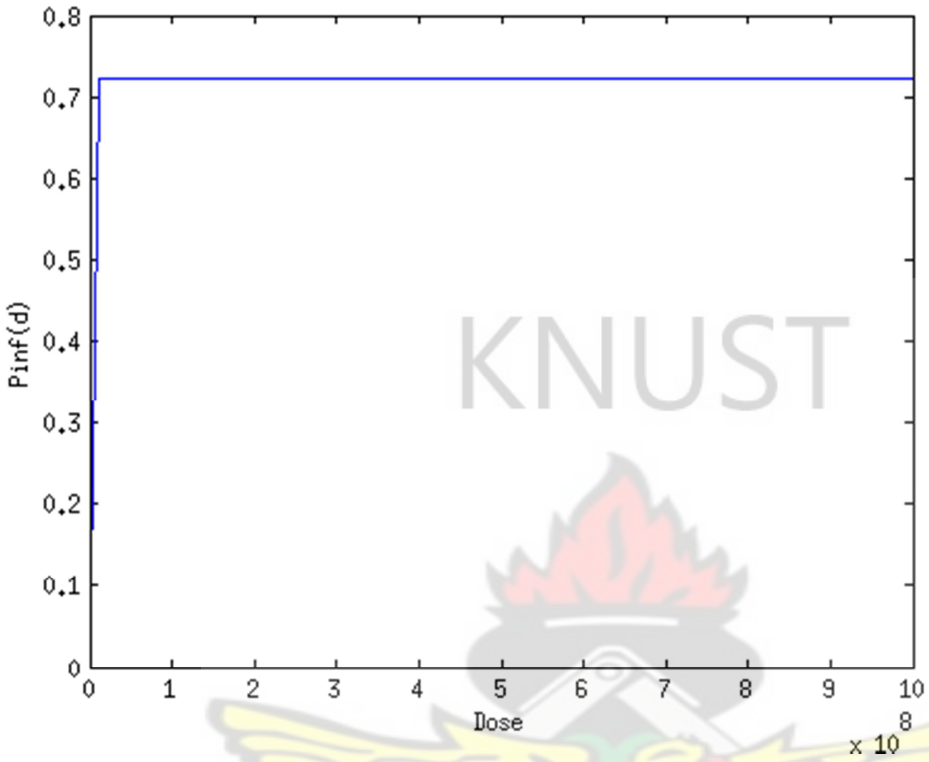


Figure 5.1: Simulated results for probability of infection with Fractional-Poisson model within a dose of genome copies.

5.2.2 Temporary Immunity Dose Response Incidence (DRI) Models

Dose Response models measure the response to ingestion of pathogens and are thereby crucial for estimating the risk associated with the ingestion of pathogens. In this section, four different DRI models are developed and their inclusion for assessing risk estimates as applied in a *Norovirus* risk assessment. The purpose is to estimate the impact as a result of exposures leading to temporary acquired immunity for an individual.

5.2.2.1 Naive DRI *Model*

Reference to equations 4.20 to 4.25, the conditional probability of illness from infected subjects is $P_{ill} = P_{ill|\inf} P_{\inf}$

Assumption

1. Illness outcome is independent on all previous exposure.
2. Probability of illness at exposure $P(d_j)$ is Bernoulli distributed.
3. Total exposure E for a specified period

Then the probability of illness is given as (Havelaar and Swart, 2014)

$$P_{ill} = \sum_{j=1}^{E} P_{ill|\inf}(d_j) P_{\inf}(d_j) \quad 5.7$$

For cases of independency, $P_{ill|\inf} = \varphi$, the conditional probability per exposure is $P_{ill} = \phi P_{\inf}(d)$. For a constant average dose, the probability of illness for individual exposure is

$$P_{ill} = E\varphi P_{\inf}(d) \quad 5.8$$

Where E is the total number of exposures within a year. Therefore, the probability of illness within the population is also given as $P_{ill} = NE\varphi P_{\inf}(d)$ (Havelaar and Swart, 2014), where N is the population size

5.2.2.2 Model for Multiple Exposures

Modelling risk of illness estimation with integrated immunity, the cases of illness for exposure is dependent on the waning of temporarily acquired immunity from previous transition from susceptible to partial protection and then back to be susceptible.

Effect of Acquired Immunity

In the case of a compartmental model to account for effect of acquired immunity, (Swart et al., 2012) presented the schematic process as shown below (Fig. 5.2)

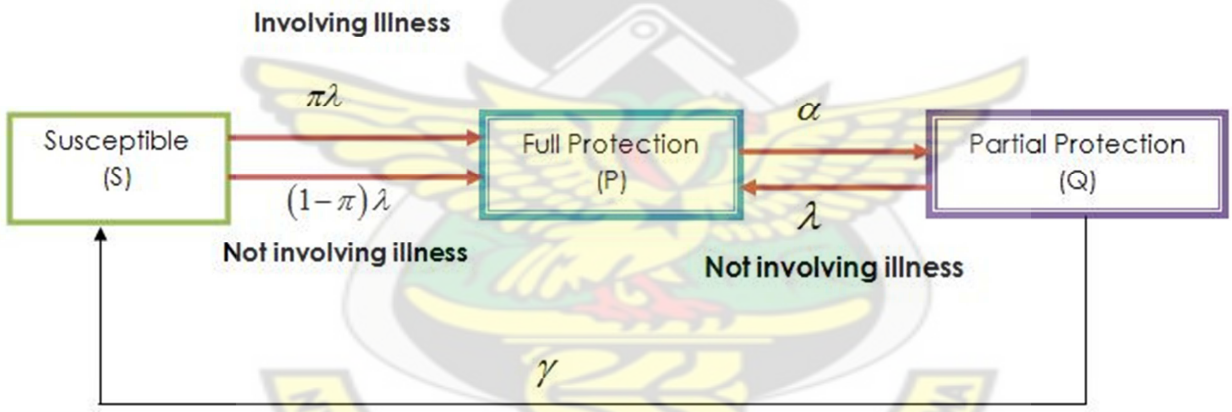


Figure 5.2: Schematic overview of infection and immunity

Table 5.1: Description of Parameters used in the Model

Symbols	Description
α	Loss of full immunity
γ	Loss of partial immunity
λ	Force of infection
π	Proportion of the susceptible involving of illness
S	Susceptible population
P	Fully protected after infection
Q	Partial protection

The deterministic first order differential equation for Figure 2.3 can be described by equation (5.9) as follows;

$$\left.\begin{aligned} f_1 &= \frac{dS}{da} = \gamma Q - \lambda S \\ f_2 &= \frac{dP}{da} = \lambda(Q+S) - \alpha P \\ f_3 &= \frac{dQ}{da} = \alpha P - (\gamma+\lambda)Q \end{aligned}\right\} \qquad 5.9$$

where, all parameters are positive, thus S(0),P(0) and Q(0)>0. The Jacobian of equation (5.9) is represented as;

$$J = \frac{df}{da_j} = \begin{bmatrix} \frac{df_1}{dS} & \frac{df_1}{dP} & \frac{df_1}{dQ} \\ \frac{df_2}{dS} & \frac{df_2}{dP} & \frac{df_2}{dQ} \\ \frac{df_3}{dS} & \frac{df_3}{dP} & \frac{df_3}{dQ} \end{bmatrix} = \begin{pmatrix} -\lambda & 0 & \gamma \\ \lambda & -\alpha & \lambda \\ 0 & \alpha & -(\lambda+\gamma) \end{pmatrix}$$

For $\det(J - XI) = 0$, the resulting characteristic equation is

$$J - XI = \begin{pmatrix} -\lambda - X & 0 & \gamma \\ \lambda & -\alpha - X & \lambda \\ 0 & \alpha & -(\lambda+\gamma) - X \end{pmatrix} \qquad 5.10$$

$$X^3 + \lambda X^2 + \gamma X^2 + \alpha X^2 + \alpha\lambda X + \alpha\gamma X + \lambda X^2 + \lambda^2 X + \lambda\gamma X = 0$$

$$X^3 + \lambda X^2 + \gamma X^2 + \alpha X^2 + \lambda X^2 + \alpha\lambda X + \alpha\gamma X + \lambda^2 X + \lambda\gamma X = 0$$

$$X^3 + (2\lambda + \gamma + \alpha) X^2 + (\alpha\lambda + \alpha\gamma + \lambda^2 + \lambda\gamma) X = 0$$

$$X\left[X^2 + (2\lambda + \gamma + \alpha) X + (\alpha\lambda + \alpha\gamma + \lambda^2 + \lambda\gamma) \right] = 0$$

Therefore, the eigenvalues of the characteristics equation are

$$X_1 = 0, X_2 = -(\alpha + \lambda), X_3 = -(\lambda + \gamma)$$

and the corresponding eigenvectors from equation 15.10 are: $(\alpha\gamma, \lambda(\lambda+\gamma), \alpha\lambda), (-\gamma, (\gamma-\alpha), \alpha), (1, 0, -1)$.

Let R be the expected number of transition of an individual from S to P before ultimately dying of natural cause, assuming that, illness does not leads to death. Again, let the expected number of transitions from S to P be defined as $R = \lambda \int_0^\infty s(a) F(a) da$, where s(a) is the probability of individual susceptibility at age 'a', and F(a) is the individual probability to survive until age 'a'. Therefore, the full and partial waning of immunity is given as;

$$\frac{d}{da}\begin{pmatrix} s \\ p \\ q \end{pmatrix} = \begin{pmatrix} -\lambda & 0 & \gamma \\ \lambda & -\alpha & \lambda \\ 0 & \alpha & -(\lambda+\gamma) \end{pmatrix}\begin{pmatrix} s \\ p \\ q \end{pmatrix} \qquad 5.11$$

$s(0)=1, p(0)=0, q(0)=0$.

The general solution for the equation 5.11 using the eigenvalues and eigenvectors of equation 5.10 and equation 5.11

$$\begin{pmatrix} s(a) \\ p(a) \\ q(a) \end{pmatrix} = C_1\begin{pmatrix} \alpha\gamma \\ \lambda(\gamma+\lambda) \\ \alpha\lambda \end{pmatrix} + C_2\begin{pmatrix} -\gamma \\ (\gamma-\alpha) \\ \alpha \end{pmatrix}\exp\left[-a(\alpha+\lambda)\right] + C_3\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\exp\left[-a(\gamma+\lambda)\right]$$

Using the initial conditions from equation 5.11 results

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = C_1\begin{pmatrix} \alpha\gamma \\ \lambda(\gamma+\lambda) \\ \alpha\lambda \end{pmatrix} + C_2\begin{pmatrix} -\gamma \\ (\gamma-\alpha) \\ \alpha \end{pmatrix} + C_3\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad 5.12$$

$$\begin{aligned} 1 &= \alpha\gamma C_1 - \gamma C_2 + C_3 \\ 0 &= \lambda(\gamma+\lambda)C_1 + (\gamma-\alpha)C_2 \\ 0 &= \alpha\lambda C_1 + \alpha C_2 - C_3 \end{aligned}$$

Hereafter, $C_1 = \frac{1}{(\alpha+\lambda)(\gamma+\lambda)}, C_2 = \frac{\lambda}{(\alpha-\gamma)(\alpha+\lambda)}, C_3 = \frac{\alpha\lambda}{(\alpha-\gamma)(\gamma+\lambda)}$. Therefore, the specific solution is given as;

$$\left.\begin{aligned}\begin{pmatrix} s(a) \\ p(a) \\ q(a) \end{pmatrix} &= \frac{1}{(\alpha+\lambda)(\gamma+\lambda)}\begin{pmatrix} \alpha\gamma \\ \lambda(\gamma+\lambda) \\ \alpha\lambda \end{pmatrix} \\ &+ \frac{1}{(\alpha-\gamma)(\alpha+\lambda)}\begin{pmatrix} -\gamma \\ (\gamma-\alpha) \\ \alpha \end{pmatrix}\exp\left[-a(\alpha+\lambda)\right] \\ &+ \frac{\alpha\lambda}{(\alpha-\gamma)(\gamma+\lambda)}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\exp\left[-a(\gamma+\lambda)\right]\end{aligned}\right\} \quad 5.13$$

The relevant solution for quantification of expected number of transition into the susceptible compartment after waning of immunity is given as;

$$s(a) = \frac{\alpha\lambda(\alpha+\lambda)\exp\left[-a(\gamma+\lambda)\right] + \alpha\gamma(\alpha-\gamma) - \gamma\lambda(\gamma+\lambda)\exp\left[-a(\alpha+\lambda)\right]}{(\alpha-\gamma)(\alpha+\lambda)(\gamma+\lambda)} \quad 5.14$$

Let the probability to survive until age 'a' be described by a survival function $F(a) = \begin{cases} 1, & 0 \le a \le A \\ 0, & a > A \end{cases}$, where A is the life expectancy of the population under study, consequently, the total transition from S to P of an individual is

$$\left.\begin{aligned} R &= \frac{\alpha\lambda^2}{(\alpha-\gamma)(\gamma+\lambda)^2}\left[1-\exp\left[-A(\gamma+\lambda)\right]\right] \\ &- \frac{\gamma\lambda^2}{(\alpha-\gamma)(\alpha+\lambda)^2}\left[1-\exp\left[-A(\alpha+\lambda)\right]\right] \\ &+ \frac{\alpha\lambda\gamma A}{(\alpha+\lambda)(\gamma+\lambda)}\left[1-\exp\left[-A(\gamma+\lambda)\right]\right]\end{aligned}\right\} \quad 5.15$$

$$\lim_{A\to\infty}\frac{R}{A}=\frac{\alpha\lambda\gamma}{(\alpha+\lambda)(\gamma+\lambda)} \qquad 5.16$$

Hence

$$R=\lambda A\frac{\alpha\gamma}{(\alpha+\lambda)(\gamma+\lambda)}\equiv\lambda A\tau \qquad 5.17$$

Where, α is loss of full immunity, γ is loss of partial immunity, λ is force of infection (the rate at which susceptible individual acquire an infection)and τ is illness inflation factor (is the protective effect of induced immunity to compensate for future infection and diseases transmission), hence $\tau=\frac{\alpha\gamma}{(\alpha+\gamma)(\gamma+\lambda)}$. Therefore, $\lambda=EP_{\inf}(d)$ and τ is obtained from binomial model of exposure (Section 3.5.3).

$$\tau=\frac{\alpha\gamma}{\left(\alpha+EP_{\inf}(d)\right)\left(\gamma+EP_{\inf}(d)\right)} \qquad 5.18$$

5.2.2.3 Immunity DRI Model

Characterising the impact of immunity by the inflation factor through scaling the naïve model, it leads to the *immunity model* given as

$$P_{ill}=\frac{\alpha\gamma\varphi EP_{\inf}(d)}{\left(\alpha+EP_{\inf}(d)\right)\left(\gamma+EP_{\inf}(d)\right)}$$

$$P_{ill}=\tau E\varphi P_{\inf}(d) \qquad 5.19$$

As $D\to\infty$ (for higher dose level), $P_{\inf}(d)\approx 1$, hence immunity model in such a scenario is represented as

$$P_{ill} = \frac{\alpha\gamma\varphi E}{\left(\alpha + EP_{\inf}(d)\right)\left(\gamma + EP_{\inf}(d)\right)}$$

$$P_{ill} = \tau E \varphi \qquad 5.20$$

Incorporating the Effect of Dose dependent conditional probability of illness

Modelling the hazard function of illness subject to exposure duration of infection (equation), the infected duration describes the period of which infection persists in an individual thus $0 \le t < \Lambda$, where Λ is the entire period of infection hence describing the hazard function $H(t)$ (Probability of illness given infection)

$$H(t) = P\left(ill \middle| \inf : u\right) = 1 - \exp\left[-\int_0^t h(t)\,dt\right], \quad 0 \le t < \Lambda \qquad 5.21$$

For a hazard function defined as

$$h(t) = -\frac{d}{dt}\ln\left[1 - F(t)\right] = -\frac{d}{dt}\ln S(t)$$

$$H(t) \overset{def}{=} \int_0^t h(u)\,du, \quad t > 0$$

$$= -\ln\left[1 - F(t)\right]$$
$$= -\ln S(t)$$

$$S(t) = \exp\left(-H(t)\right)$$
$$f(t) = h(t)\exp\left(-H(t)\right)$$

Scaling the infection period leads to an integral of the hazard function $H(t/\Lambda)$ over the period of infection and assuming an exponential model for the survival function (see equation 4.20 to 4.21). If the scale factor is η, then

$$P(ill|\inf;\Lambda) = 1-\exp(-\eta\Lambda)$$

Varying the distribution of the unknown duration Λ with Gamma distribution to account for individual heterogeneity in resistance and persistence of host to colonization of infection leads to

$$g(\Lambda;\omega,d) = \frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)$$

$$P(ill|\inf) = \int_{r=0}^{\infty}\left[1-\exp(-\eta\Lambda)\right]\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]d\Lambda$$

$$= \int_{r=0}^{\infty}\left[\frac{d^{-\omega}}{\Gamma(\omega)}\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]d\Lambda - \int_{r=0}^{\infty}\left[\Lambda^{\omega-1}\exp\left(\frac{-\Lambda}{d}\right)\right]\exp(-\eta\Lambda)\,d\Lambda$$

$$= 1-\frac{d^{-\omega}}{\Gamma(\omega)}\int_{r=0}^{\infty}\left[\Lambda^{\omega-1}\exp\left(\frac{-\Lambda(1+\eta)}{d}\right)\right]d\Lambda$$

$$P(ill|\inf) = 1-(1+\eta d)^{-\omega} \qquad 5.22$$

Where ω and ηd are the shape and scale parameters of an underlying Gamma distribution for duration of infection describing the heterogeneity in response of subjects.

5.2.2.4 Dose DRI Model

Replacing φ in this case for the immunity model leads to the ***dose model*** (Messner et al., 2014)as

$$I = E\left(1-\left(1+\eta d\right)^{-\omega}\right)P_{\inf}\left(d\right) \qquad 5.23$$

5.2.2.5 Combined Model with Immunity and Dose Dependence

The dose-immunity *DRI* model has the effects of acquired immunity and dose-dependent conditional probability of illness as

$$P_{ill} = \frac{\alpha\gamma E\left[1-\left(1+\eta d\right)^{-\omega}\right]P_{\inf}(d)}{\left(\alpha+EP_{\inf}\left(d\right)\right)\left(\gamma+EP_{\inf}\left(d\right)\right)}$$

$$P_{ill} = \tau E\left(1-\left(1+\eta d\right)^{-\omega}\right)P_{\inf}\left(d\right) \qquad 5.24$$

5.3 Model Implementation

In this section, the model is implemented with the use of simulation from epidemiological data sources for *Norovirus* for varying dose. An iterative approach of sampling with hypercube sampling procedure (Section 3.7) is employed, unknown parameters were estimated (Section 3.6) and model uncertainty quantified.

5.3.1 Simulated Model Implementation

Modelling the acquired temporary immune probability of illness requires multifaceted data input parameters to describe the various relations and probability distributions. Parameter description and generation are intrinsic and conservative, to evaluate the

incidence models with global values of *Norovirus*, a simulation plot was carried out for the illness inflation factor (Equation 5.18). Implementing the loss of full and partial immunity impact with the illness inflation factor, the plot shows a steeply decrement in illness inflation factor for frequent exposures (daily and weekly) as dose increases, a less steeply decrement was also recorded for monthly exposure with increasing dose as well. Nevertheless, the illness inflation factor does not respond sensitively to infrequent exposures (semi-yearly and yearly) as compared to the frequent exposures for increasing dose. Figure 5.3 confirmed that exposure frequency do have impact on the inflation factor of illness and contributes to its prediction of illness incidence with a loss of either partial or full immunity into the susceptible compartment.

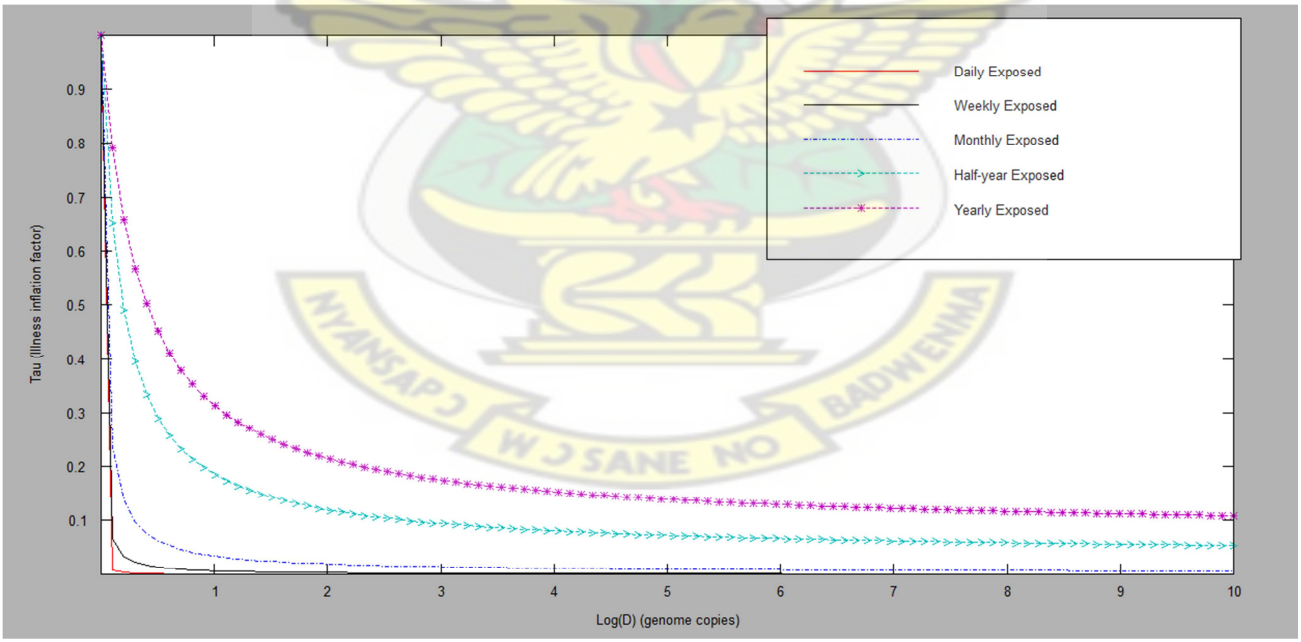


Figure 5.3: Illness inflation factor (tau) as a function of force of infection depending on dose and exposure intensity for Norovirus for parameters $\alpha \sim [0.1]$, $\gamma \sim [2.42 - 8.44] X 10^{-4}$

5.3.2 Dose Response (DRI) Models for Incidence of illness with Frequency of Exposure

In this thesis, the impact of temporary induced immunity and dose-dependent factors for incidence of illness DRI models are compared to the naïve and dose models currently in use. It's worth noting that, the estimated parameters describing the range of values for loss of full and partial immunity were used as a function of illness inflation factor to estimate the incidence of illness in *DRI* models.

From Figure 5.4 to Figure 5.7, the various DRI models have striking effect on the estimated incidence risk given exposure, as the illness incidence are strongly reliant on the dose levels, besides the impact of illness is dependent on the different frequency exposures as well as the characterisation of the DRI model with respect to whether there is an inclusion of immunity or not.

Results show (Equation 5.8), the naïve DR incidence model (Figure 5.4) increases sharply with increment in dose level and therefore highly sensitive to frequent exposure (daily and weekly exposures) whiles infrequent exposures also directly responded, however, with less impact. On dose DRI model (Figure 5.5) resulted from equation 5.23, there is a slight decrease in illness incidence for infrequent exposure (Monthly, Semi-yearly and yearly) whereas the decrease is sharp with increasing dose level for the frequent exposure (daily and weekly exposures).

For immunity (Equation 5.19) DRI model (Figure 5.6), the frequent exposure (daily and weekly) have a higher incidence level at low dose and decreases monotonically with increasing dose, nevertheless, for infrequency exposures (monthly, semi-yearly and yearly), illness incidence increases with increasing dose levels.

For dose-immunity (Equation 5.24) DRI model incidence (Figure 5.7) reaches maximum, the model has a higher incidence of illness at lower dose and at a frequent exposure level. Clearly, at a lower dose, the illness incidence is mostly dependent on the DRI model, thus dose-immunity DRI model exhibits a higher illness incidence reduction than the rest of the models, it is also influence by the frequency of exposure, the more frequent exposure, the higher the level of illness incidence reduction.

Illness incidence reduction is impacted significantly for DRI models with immunity inclusion and exhibits a less prominent of increasing effect of dose on incidence of illness at a higher dose levels. Generally, DRI models for risk estimates with immunity inclusion are approximately 2 logs lower than those without immunity included for *Norovirus*, and confirms a similar case for *C. jejuni* illness incidence reduction (Havelaar and Swart, 2014)on the prediction that, the current use of probable risk of illness (naïve and dose incidence DRI models approach) overestimate the true incidence of risk of illness.

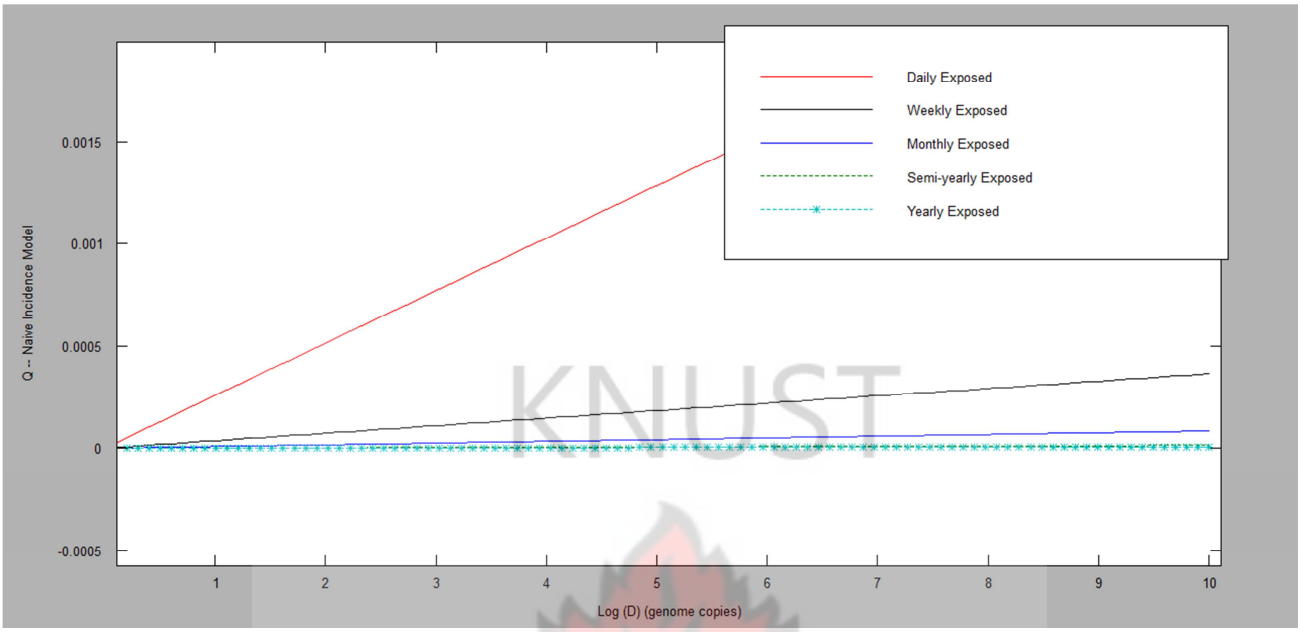


Figure 5.4: illness Incidence for Naïve Dose Response Incidence Model

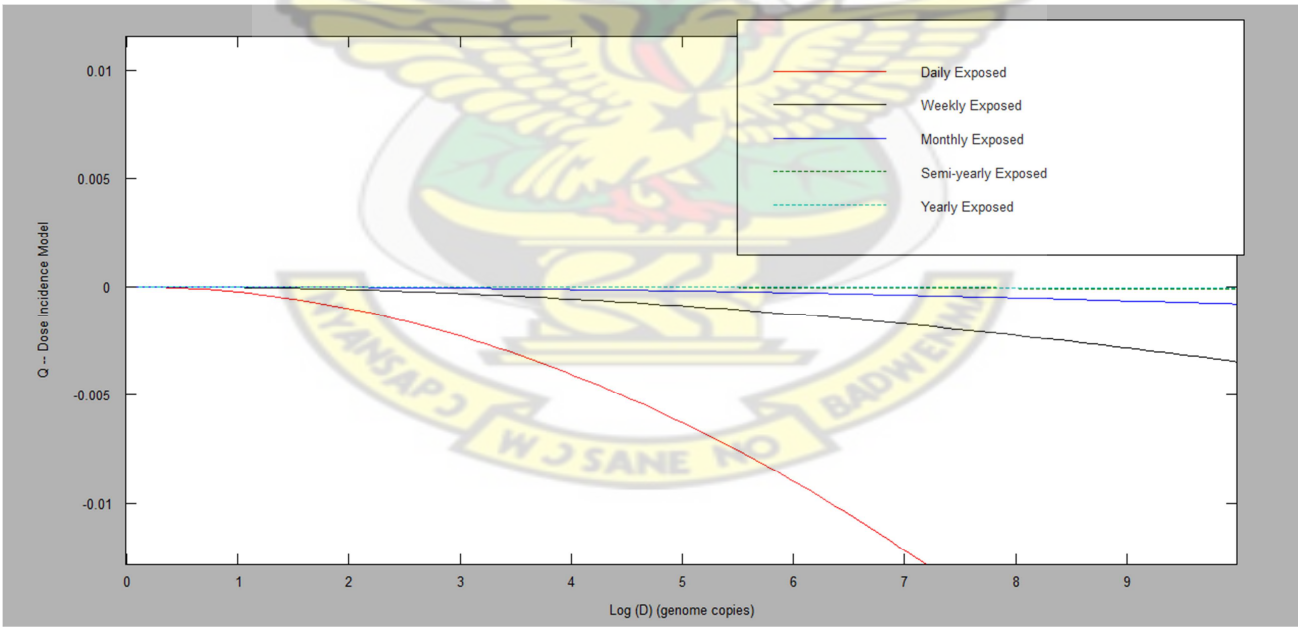


Figure 5.5: illness Incidence for Dose DRI Model

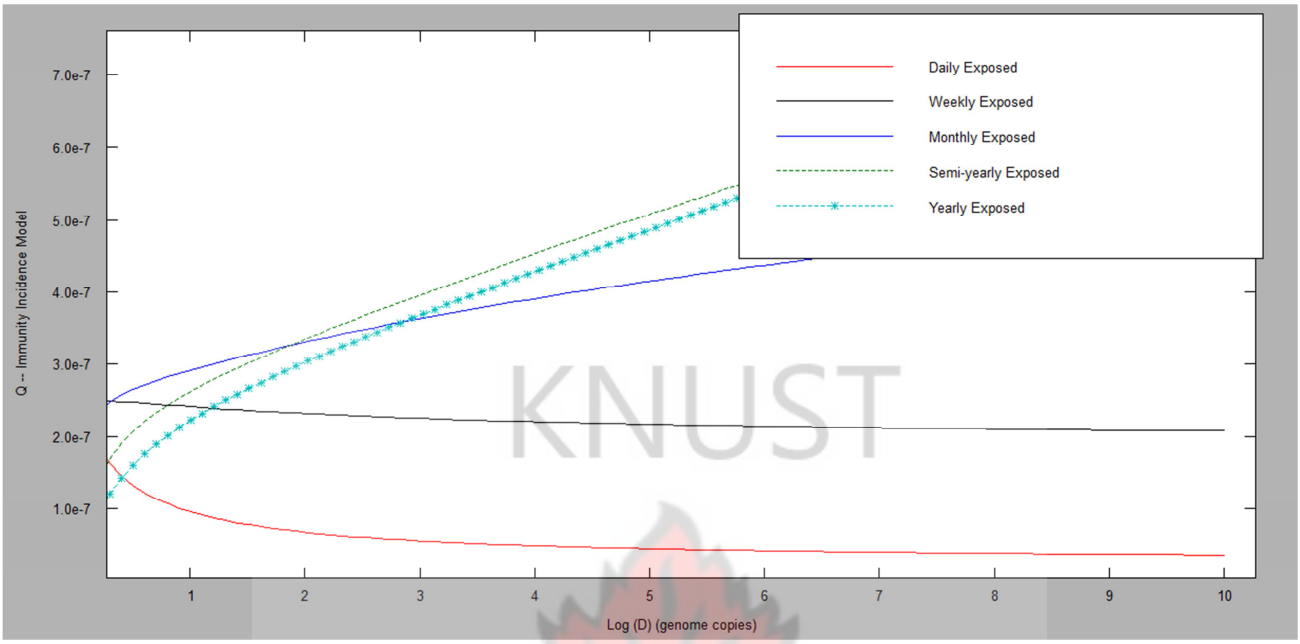


Figure 5.6: illness Incidence for Immunity Dose Response Incidence Model

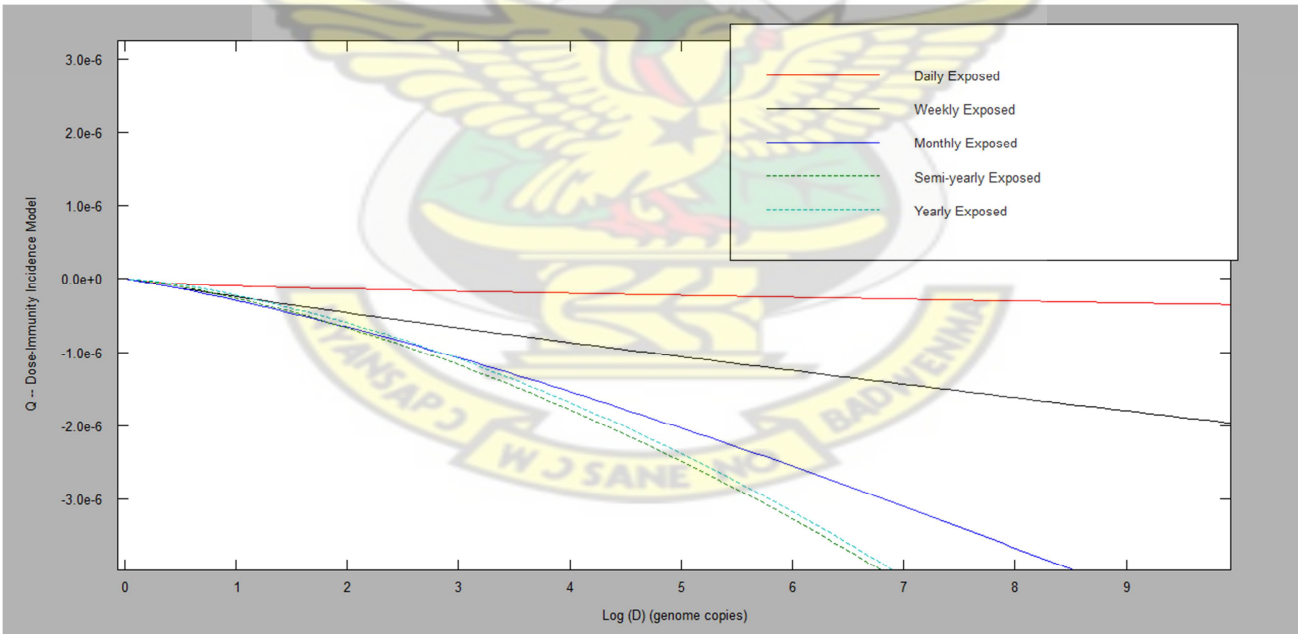


Figure 5.7: illness Incidence for Dose-immunity Dose Response Incidence Model

5.4 Illness Scenarios model Implementation with Norovirus Epidemiological Data

In this section, examination of simulated results on temporary immunity dose response models shall be carried out, population dynamics shall be integrated into the model to describe various scenarios, the scenarios model will be adopted from epidemiological studies on *Norovirus* across all different clinical transmissions, (Huynen et al., 2013; Simmons et al., 2013). The models will keep track of the following groups in the population: Symptomatic infectiousness, pre and post symptomatic infectiousness low and high, innate genetic resistance, geno-group type 4. Duration of induced immunity has been inconsistent from different studies(Simmons et al., 2013; Atmar, 2010; Frenck et al., 2012; Hamilton et al., 2006) especially in the case of *Norovirus* which has been believed previously to be from 6 months to 2 years, yet rare studies is seen to include both the influence of acquired temporary immunity and the transmission dynamics of *Nov* in risk assessment. The inclusion of the different transmission dynamics (Figure 5.8) will help to have an idea on how these dynamics within the population could influence risk assessment given its immunity influence on the different illness incidence reduction dose response models.

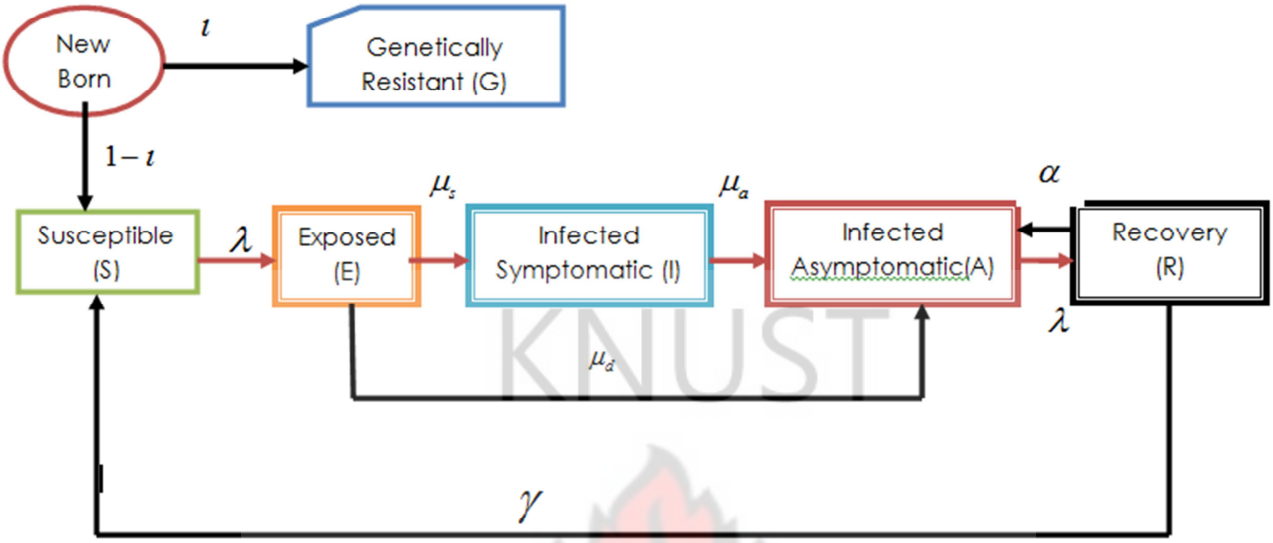


Figure 5.8: Modified Schematic overview of population dynamics of immunity states of Norovirus (Simmons et al., 2013)

Figure 5.8 describes an epidemiological model of modified Maternal Susceptible Exposed Infected Exposure (MSEIIR model) to describe *Norovirus* infectiousness in a population. Five different *Nov* transmission scenarios described below within the population adopted for estimating the illness inflation factor.

5.4.1 Scenario Description for Epidemiological Norovirus Transmission Dynamics in a Population

The various scenarios are described below;

- **Symptomatic Individual Infectious**

Only symptomatic individuals are infectious. This comprises individuals in the population under the assumption that all exposed individuals are susceptible to *Norovirus* infection and none is genetically resistance. It is worth noting that, naïve model dose not relates to symptomatic individuals infectiousness, the earlier refers to estimation of risk without inclusion of temporary acquired immunity.

- **Pre-symptomatic and Post-symptomatic infectiousness (Low)**

Pre-symptomatic persons in compartment (E) are individual Exposed but yet to be symptomatic of the infection (Ozawa et., 2007; Simmons et al., 2013; Sukhrie et al., 2012; Sukhrie et al., 2010; Teunis et al., 2014).

- **Pre-symptomatic and Post-symptomatic infectiousness (High)**

In this scenario, individuals exposed in the compartment (E) of the mathematical epidemiological model and asymptomatic compartment (Teunis et al., 2014).

- **Scenario D: Innate Genetic Resistance**

This is based on the assumption that part of the population is completely resistant to infection and diseases (G), thus they possess the non-secretor phenotype and plays no role in transmission process, however, they do make contact with persons included in empirical incidence estimate (Frenck et al., 2012). This is also different from immunity model or dose-immunity models; the innate genetic resistance is the

inclusion of individuals whose genetic make-up excludes them from infectiousness, yet forms part of the population.

Scenario E: Genogroup 2 Type 4 (GII.4)

Model scenarios A to D assume all *NoV* to be anti-genetically indistinguishable. In this scenario, it is assumed that only GII4 are infectious. The incidence of GII.4 is estimated based on values from (Vega et al., 2011; Huynen et al., 2013; Nordgren et al., 2010) ;Frenck et al., 2012) (Simmons et al., 2013)

- **Scenario F: No Immune Boosting by Asymptomatic Infection**

Persons do not travel from recovery (R) compartment to asymptomatic (A) compartment. The only pathway out of the recovery compartment is through waning of partial immunity to become susceptible (S) again.

Data input for modelling scenarios based on epidemiological studies are as shown in Table (5.2).

Table 5.2: Norovirus Epidemiological Data for Population Dynamics Immunity Modelling

Parameter	Scenario A	Scenario B	Scenario C	Scenario D	Scenario E	Scenario F
Loss of full immunity(α per year)	0.1	0.1	0.1	0.1	0.1	0.1
Loss of partial immunity (γ per year $\left(X10^{-4}\right)$	$4.22-7.02$	$4.12-6.85$	$2.42-4.02$	$5.39-8.44$	$3.42-4.91$	$4.17-7.02$
Duration of incubation μ_s (days)	1	1	1	1	1	1
Duration of asymptomatic infection ρ (days)	10	10	10	10	10	10
Duration of symptoms μ_a (days)	2	2	2	2	2	2
Relative infectiousness during asymptomatic infection period	0	0.05	0.25	0	0	0
Relative infectiousness during incubation period	0	0.05	0.25	0	0	0
Strains Included	All	All	All	All	GII.4	All
Boosting of immunity by asymptomatic infection	Yes	Yes	Yes	Yes	Yes	No
Total Exposure for Annual quantification (days)	1-365	1-365	1-365	1-365	1-365	1-365

The infection probability for subjects with disaggregated dose (P)	0.722	0.722	0.722	0.722	0.722	0.722
Parametric mean dose $\mu(a)$	1106	1106	1106	1106	1106	1106
Dose response parameters for illness given infection(r, η)	0.086, 2.60E-03	0.086, 2.60E-03	0.086, 2.60E-03	0.086, 2.60E-03	0.086, 2.60E-03	0.086, 2.60E-03
Life Expectancy A (years)	63	63	63	63	63	63

Parameter values (Simmons et al., 2013; Sukhrie et al., 2012, 2010; Tribble et al., 2010)

5.4.2 Applied Induced immunity model of Dose-Response Models for Consumers Exposure

In this section, consumers' exposure to *Norovirus* is modelled based on the dose-response immune induced models presented earlier, and characterized based on the Nov transmission dynamic scenarios presented in Section 5.4.1. Parameters describing the modelling process of the applied induced immunity model are given in Table 5.3

Table 5.3: Parametric Values for Model Implementation

Parameter	Description	Estimate	Estimate Value(s)	Reference
$P_{\inf}$	Probability of infection	Calculated[7]		Equation 5.61
d	Arithmetic Mean Dose per exposure per occasion	Variable		
$\mu(a)$	Parametric Mean dose		1106	(Teunis et al., 2008)
P	The infection probability for subjects with disaggregated dose		0.722	(Messner et al., 2014)
P_{ill}	Probability of illness	Calculated		Equation 5.8 to 5.24
$P_{ill\|\inf}$	Probability of illness given infection	Calculated		(Havelaar and Swart, 2014)
N	Population		2.50E+07	Population pyramids (2015)
E	Total Exposure	Calculated	[208.365]~[286.5]	(Seidu et al. 2008,

[7] Calculated values are based on the equations derived in the study and simulated to generated random numbers fitting onto a distribution.

				Mara et al. 2010)
ω, η	Dose response parameters for illness given infection		0.086, 2.60E-03	(Teunis et al., 2008)
α	Loss of full immunity	Calculated		(Simmons et al., 2013)
γ	Loss of partial immunity	Calculated	Values for different epidemiological scenarios refer (Table 5.1	Equation 5.11
λ	Force of infection	Calculated		Equation 5.11-
τ	Inflation factor	Calculated		Equation 5.12

Assigning a uniform distribution to the loss of full and partial immunity to characterise its influence on the immunity-DR models, estimation of illness incidence for the various transmission dynamics scenarios is therefore presented here with the *Norovirus* transmission dynamics using equations 5.8, 5.19, 5.23 and .5.24.

The transmission dynamics in all scenarios had illness incidence for dose-immunity DRI model within $1\times10^{-8} - 1\times10^{-7}$, immunity DRI model also falls within $1\times10^{-6} - 1\times10^{-3}$, dose DRI model also falls within $1\times10^{-5} - 1\times10^{-2}$ and Naïve DRI model falls within $1\times10^{-1} - 1\times10^{-0}$ The estimated difference for the dose-immunity DRI model and naïve DRI model is approximately close to 8 logs of magnitude, whiles, dose-immunity and dose DRI model also has a difference of approximately 6 logs of magnitude (Figure 5.9, 5.10, 5.11, 5.12, 5.13, 5.14). The individual illness incidence decreases from naïve, dose-model, immunity model and dose-immunity model.

The illness incidence risk estimate for various acquired immunity-incorporated dose models (dose-immunity, immunity) for all the transmission dynamic scenarios gives a much lesser estimation of risk as compare to the naïve and dose-model approach currently in use (Figure 5.9, 5.10, 5.11, 5.12, , 5.13, 5.14). Moreover, there is a significant change in log magnitude of illness incidence estimation among all dose response models and across all transmission dynamics with or without immunity inclusion.

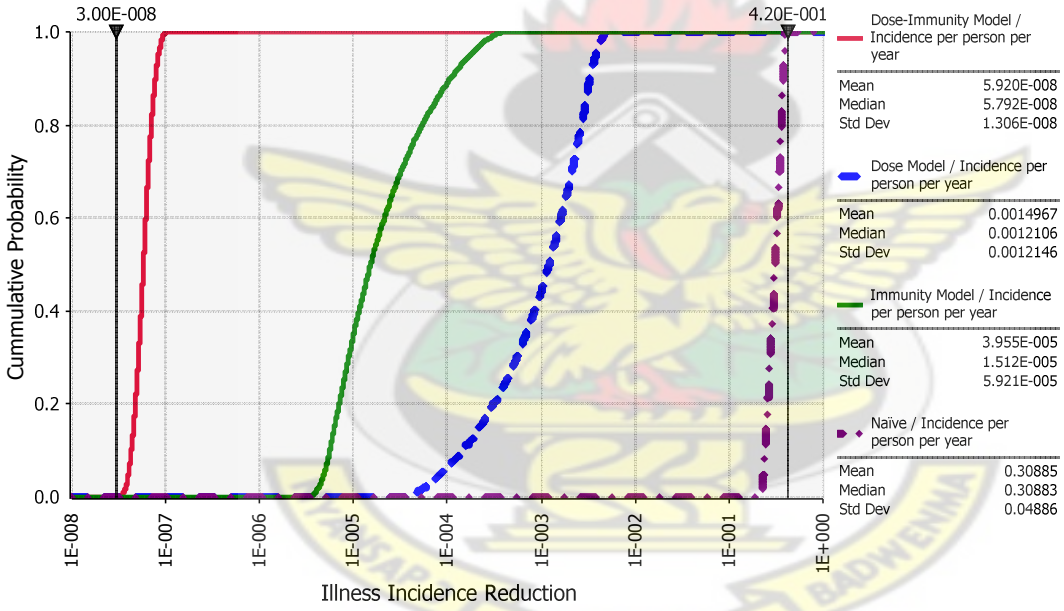


Figure 5.9: Illness Incident Reduction Models for 'Symptomatic infectiousness' of Norovirus per Person per Year

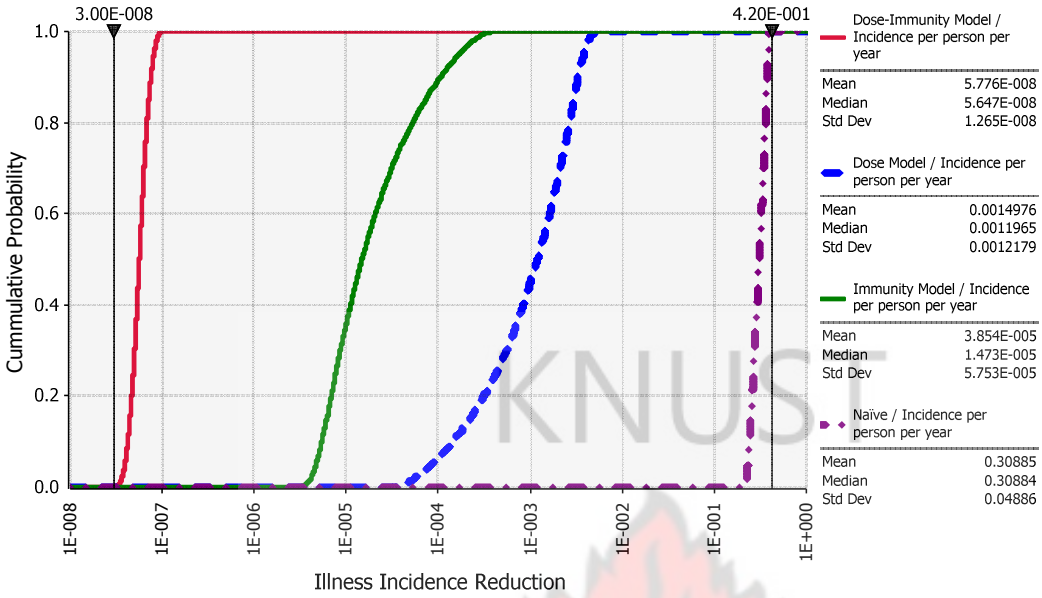


Figure 5.10: Illness Incident Reduction Models for 'Pre-Symptomatic and Post-Symptomatic infectiousness Low' of Norovirus per Person per Year

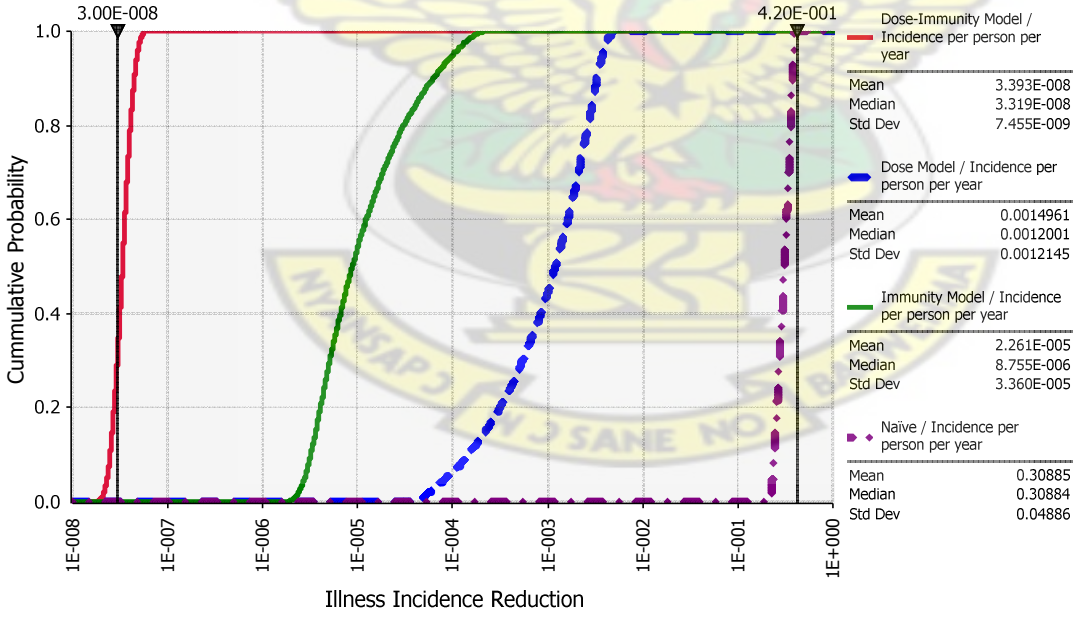


Figure 5.11: Illness Incident Reduction Models for 'Pre-Symptomatic and Post-Symptomatic infectiousness High' of Norovirus per Person per Year

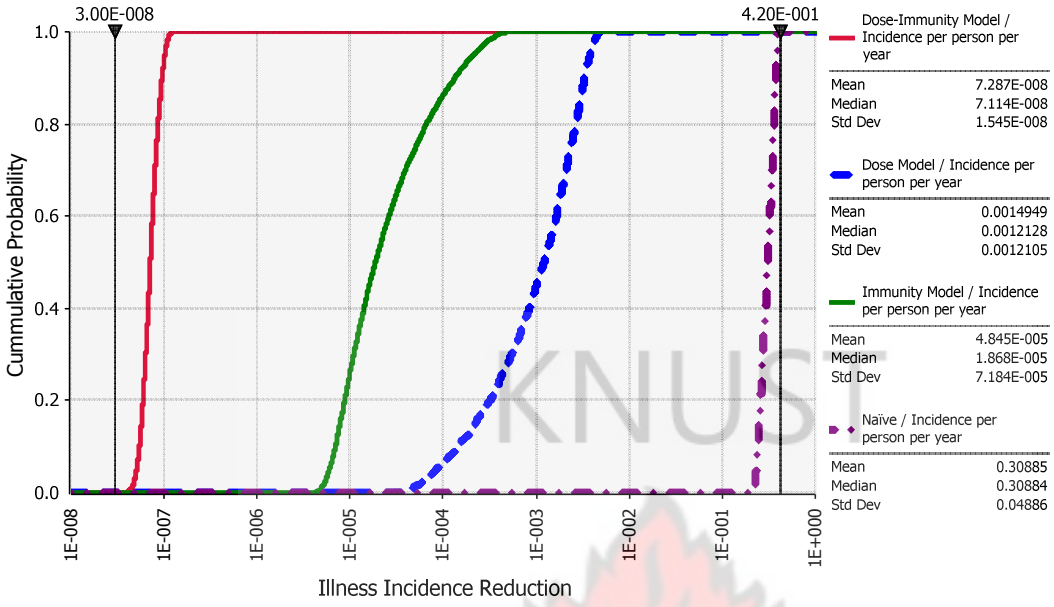


Figure 5.12: Illness Incident Reduction Models for 'Innate Genetic Resistance' of Norovirus per Person per Year

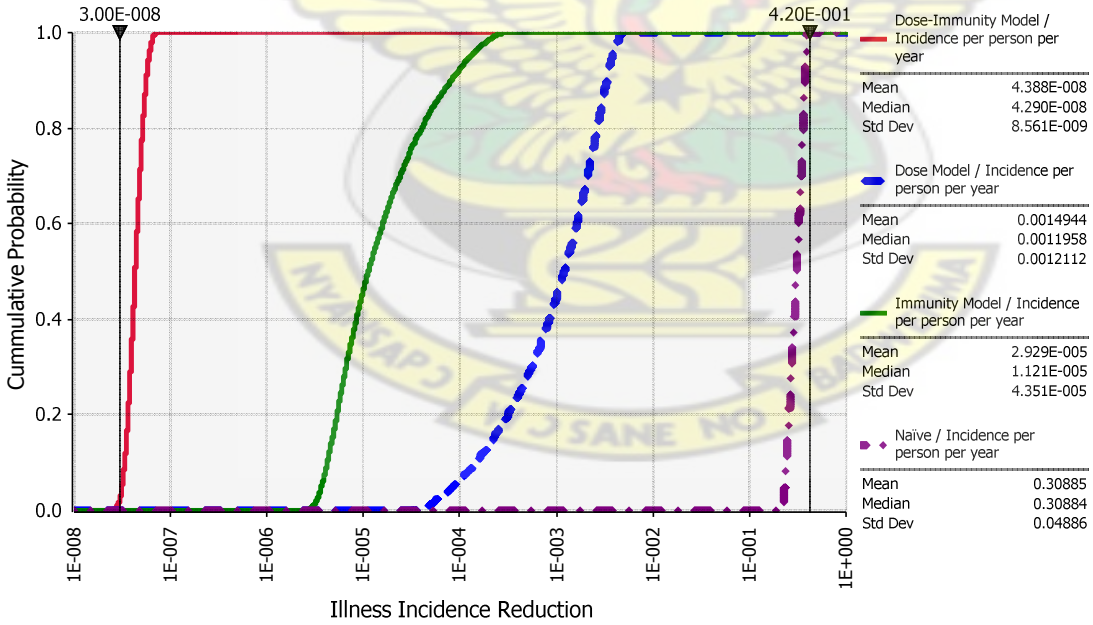


Figure 5.13: Illness Incident Reduction Models for 'Genogroup II Type 4' of Norovirus per Person per Year

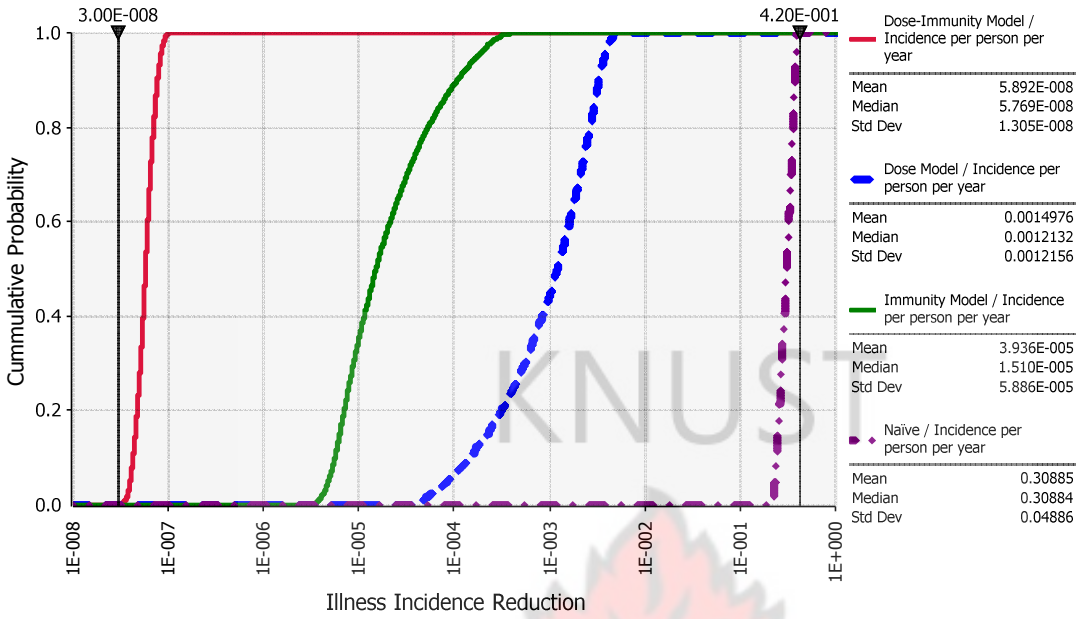


Figure 5.14: Illness Incident Reduction Models for 'No Immune Boosting after Asymptomatic Infection' of Norovirus per Person per Year

Comparatively, for the median risk of illness estimates, the dose DRI model has an approximately 2 logs of magnitude less than the naïve DRI model, the immunity DRI model recorded a 4 logs of magnitude less to the naïve DRI model and a 2 logs of magnitude less to the dose model. Furthermore, by incorporating the temporary immunity protection for the probability of illness given infection, the dose-immunity DRI model has 7 logs, 5 logs and 3 log (in most cases) of magnitude less to the naïve DRI model; dose DRI model and immunity DRI model respectively (Table 5.4).

Across the different transmission dynamics scenarios with respect to their loss of partial and full induced immunity protection levels, a comparison of the various DRI models with each of the epidemiological scenarios for individual illness incidence did not

show a difference. The difference in the illness incidence DRI models is not sensitive to the epidemiological scenarios, thus, *Norovirus* transmissions dynamics does not affect the various DRI models predictions. In all cases, the immunity DRI model and the dose-immunity DRI models resulted in significant lower levels of illness as compared to naïve and dose-DRI models (Table 5.4). A comparison difference of 7 logs of magnitude exists for dose-immunity and naïve DRI models for all epidemiological transmission dynamics, 5 logs, and 3 logs differences for dose-immunity and dose-DRI model, immunity DRI model respectively. Though the DRI models exhibit no differences in terms of logs magnitude of incidence estimates of illness among the transmission dynamics (with the exception of immunity dose model which had 1 log less for pre-symptomatic and post-symptomatic low), there exist some differences in terms of values which can translate into logs magnitude difference when the estimate is extended for a large population. Furthermore, there is no difference for all DRI models of the transmission dynamics of symptomatic infectiousness' and the 'no immune boosting after asymptomatic infectiousness' this confirms Teunis et al., (2014) study indicating, shedding of virus is similar for both symptomatic and asymptomatic infectiousness, however, it is also worth noting, some differences exist between studies for shedding of virus of infected subjects (Atmar et al., 2008), this differences is attributed to genotype studied, nevertheless, the difference in numbers shed could not have clinical significance, hence such indifference in risk estimate of illness incidence as seen is not unusual (Teunis et al., 2014).

Table 5.4: Annual Individual risk of Illness for Dose-Response Models with Epidemiological Scenarios of Norovirus

Scenarios/Models	**Naïve Model**	**Immunity Model**	**Dose Model**	**Dose-Immunity Model**
Symptomatic Individual Infectious	3.09×10^{-1}	1.51×10^{-5}	1.21×10^{-3}	5.7×10^{-8}
Pre-symptomatic and Post-symptomatic infectiousness (Low)	3.09×10^{-1}	1.47×10^{-5}	1.19×10^{-3}	5.65×10^{-8}
Pre-symptomatic and Post-symptomatic infectiousness (High)	3.09×10^{-1}	8.76×10^{-6}	1.20×10^{-3}	3.32×10^{-8}
Innate Genetic Resistance	3.09×10^{-1}	1.87×10^{-5}	1.21×10^{-3}	7.11×10^{-8}
Genogroup 2 Type 4 (GII.4)	3.09×10^{-1}	1.12×10^{-5}	1.19×10^{-3}	4.29×10^{-8}
No Immune Boosting by Asymptomatic Infection	3.09×10^{-1}	1.51×10^{-5}	1.21×10^{-3}	5.77×10^{-8}

Population Risk Estimate

In order to estimate for an approximate population of 25million total population for Ghana, range of values characterising the transmission dynamics to estimate the illness incidence in population with varying loss of partial and full immunity, shows a significant decrease in illness incidence considering at different percentage level. According to results (Appendix B, See basic statistics results, and output distributions), the illness incidence level saw a decrease when dose-immunity or immunity DRI model is used instead of dose model or naïve model, and hence by incorporating the effect of the immunity and dose-dependent lead to a further approximately logs magnitude of less prediction of illness incidence in the population.

5.5 Daily Adjusted Life Years (DALYs) with immunity incorporated DR models for different transmission dynamics for *Norovirus*

Estimating the annual risk of infection or illness with temporary immunity dose response models can be translated into estimating the overall DALY across transmission dynamics as follows:

$$P_{ill\_k} = 1 - \left[1 - \prod_{i=1}^{E}(1 - Q_i)\right] \quad 5.25$$

Where E is the total exposure and Q is illness incidence per exposure, hence

For Naïve DRI model

$$P_{ill\_naive} = 1 - \left[1 - \prod_{i=1}^{E}\left(1 - \left(\varphi P_{\inf}(d)\right)_i\right)\right] \quad 5.26$$

Immunity DRI Model

$$P_{ill\_immunity} = 1 - \left[1 - \prod_{i=1}^{E}\left(1 - \left(\tau\varphi P_{\inf}(d)\right)_i\right)\right] \quad 5.27$$

Dose DRI Model

$$P_{ill\_immunity} = 1 - \left[1 - \prod_{i=1}^{E}\left(1 - \left(\left(1 - (1+\eta d)^{-\omega}\right)P_{\inf}(d)\right)_i\right)\right] \quad 5.28$$

Dose-Immunity DRI Model

$$P_{ill\_dose-immunity} = 1 - \left[1 - \prod_{i=1}^{E}\left(1 - \left(\tau\left(1 - (1+\eta d)^{-\omega}\right)P_{\inf}(d)\right)_i\right)\right] \quad 5.29$$

Therefore, the Daily Adjusted life years is estimated as

$$P_{ill\_k} = 1 - \left[1 - \prod_{i=1}^{E}(1 - Q_i)\right]B \quad 5.30$$

Where k, is the DRI model used for the estimation, B is the diseases burden within the population

Daily Adjusted Life Years, which measures the diseases or health conditions of people as the sum of the years of life lost, due to premature mortality and disability (WHO, 2015). The DALY for the various DRI models across the transmission dynamics follows similar patters as the risk estimate of individual illness incidence. In all scenarios, the dose-immunity DRI model falls within $1.0\times10^{-11}-1.0\times10^{-9}$ DALY pppy, Immunity DRI model also falls within $1.0\times10^{-9}-1.0\times10^{-6}$ DALY pppy, the dose DRI model falls within $1.0\times10^{-8}-1.0\times10^{-4}$ DALY pppy and the naïve DRI model falls within $1.0\times10^{-5}-1.0\times10^{-2}$ DALY pppy (Figure 5.15, 5.16, 5.17, 5.18, 5.19, and 5.20).

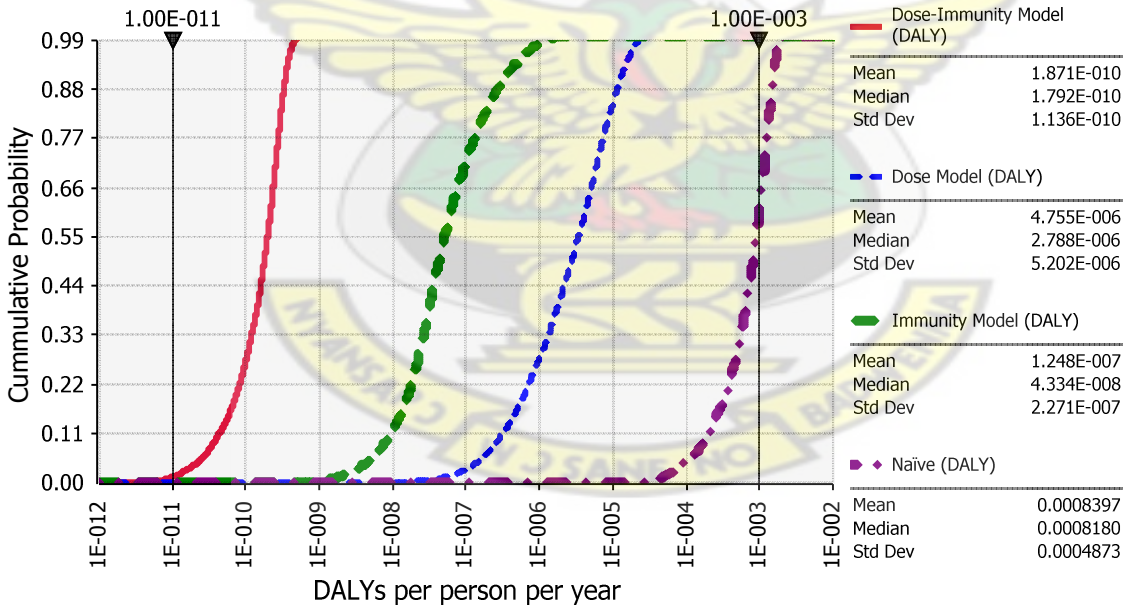


Figure 5.15: DALY for 'Symptomatic Infectiousness' transmission dynamics

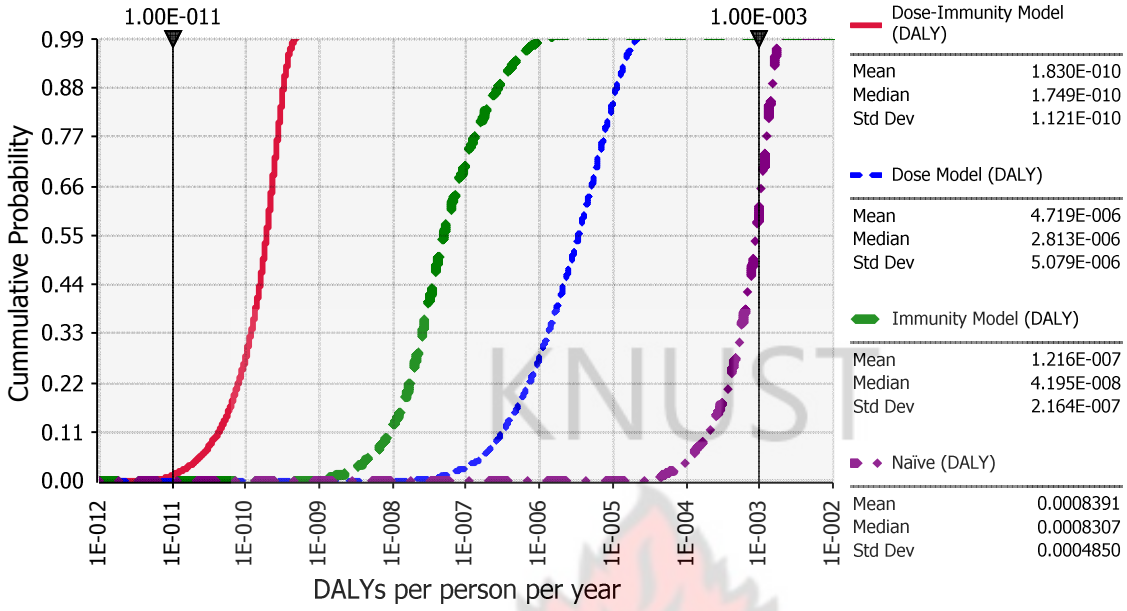


Figure 5.16: DALY for 'Pre-Symptomatic Post Symptomatic Infectiousness Low' for transmission dynamics

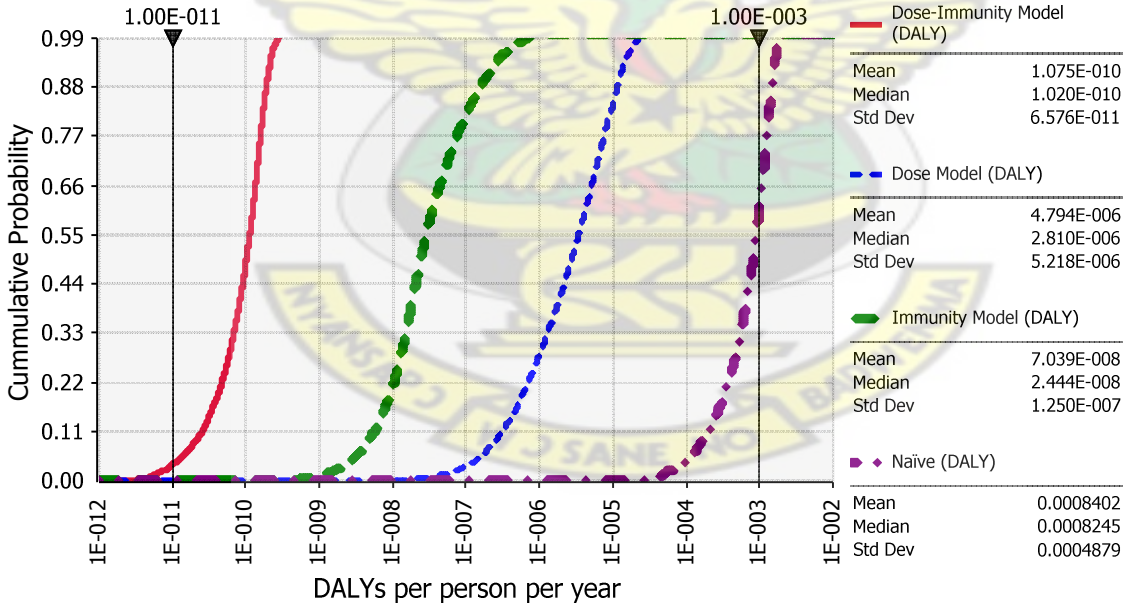


Figure 5.17: DALY for 'Pre-Symptomatic Post Symptomatic Infectiousness High' for transmission dynamics

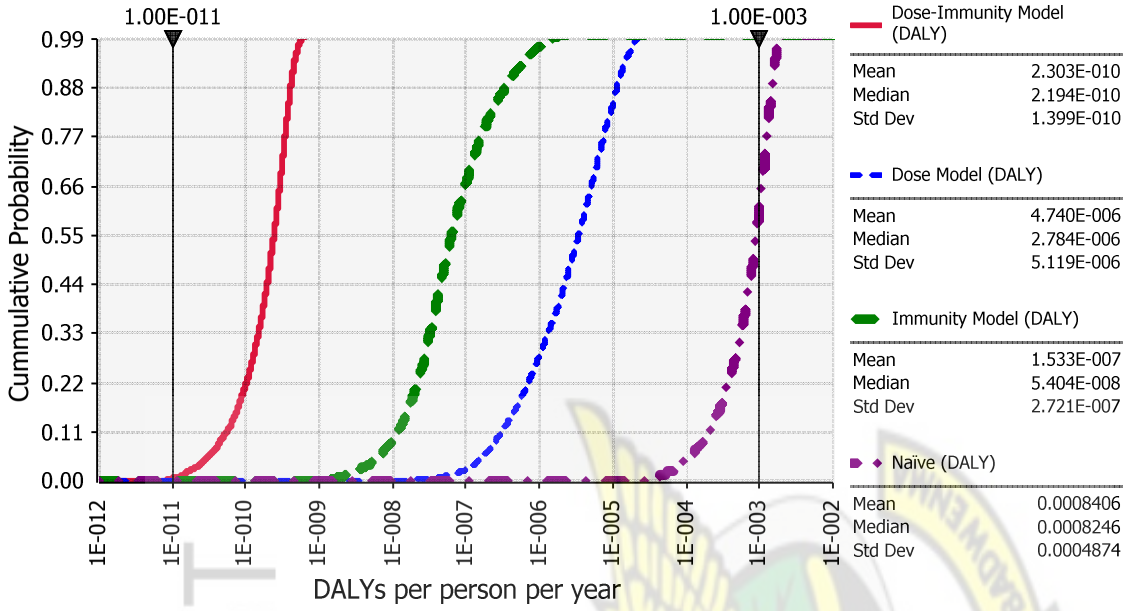


Figure 5.18: DALY for 'Innate Genetic Resistance' for transmission dynamics

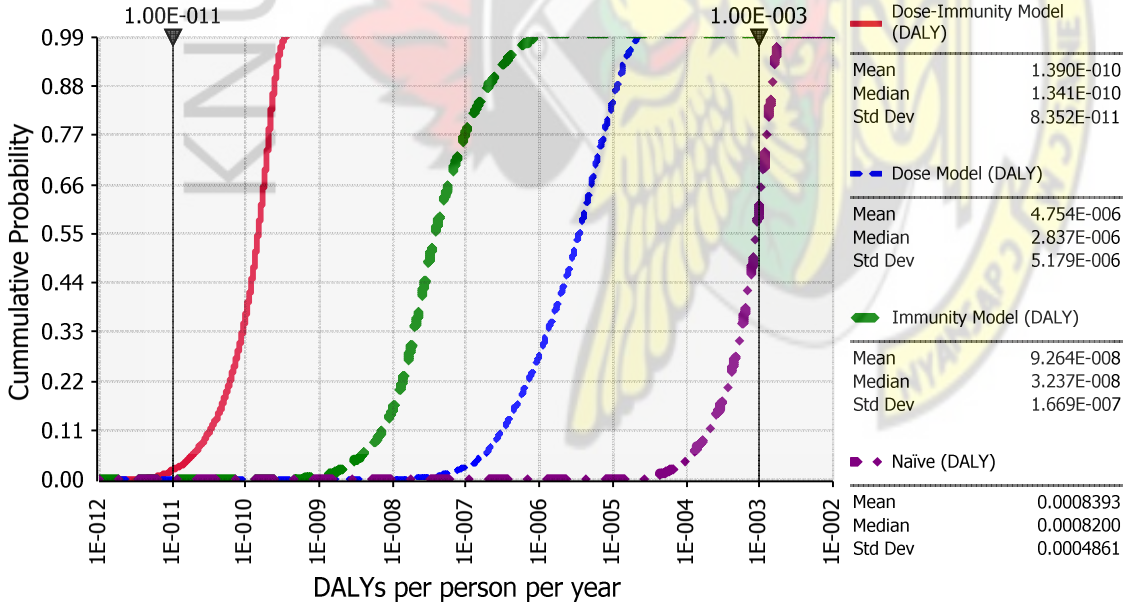


Figure 5.19: DALY for 'Genogroup 2 Type 4' transmission dynamics

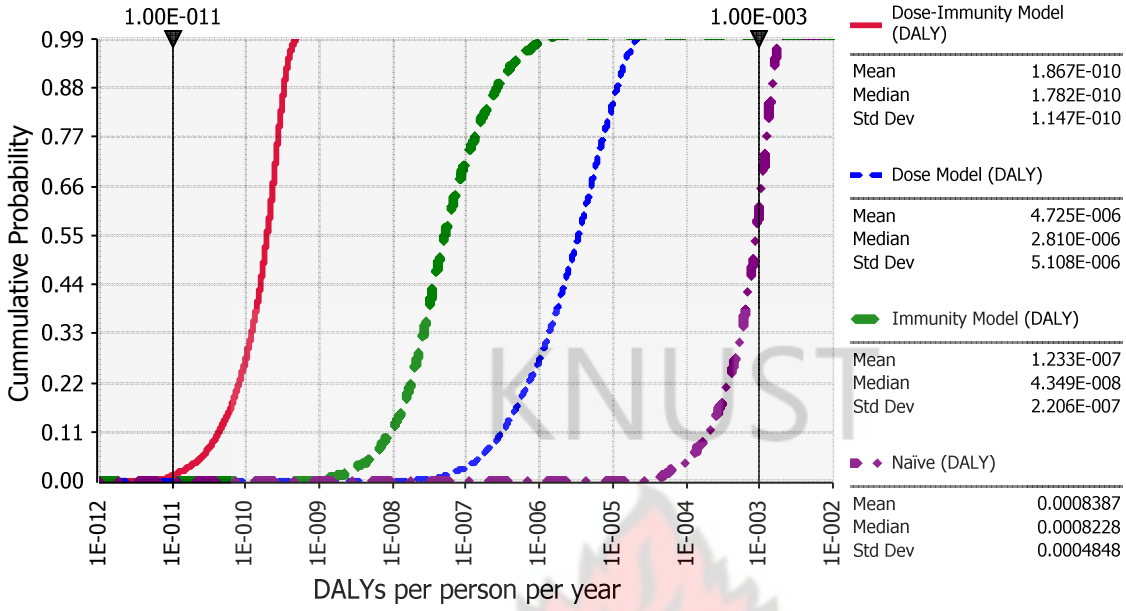


Figure 5.20: DALY for 'No Immune Boosting by Asymptomatic Infection' for transmission dynamics

The median DALY estimate, puts all scenarios describing the transmission dynamics acceptable for both immunity DRI model and dose-immunity DRI model by using the WHO standard of 1.0×10^{-4} DALY pppy (Mara and Sleigh, 2009) whiles, the dose DRI model and naïve DRI model do not meet the WHO standard. The immunity included dose response models had ≤1logs order of magnitude less than the WHO standard, on the other hand, the dose-dependent DRI models also had ≥1log order of magnitude higher than the WHO standard (Table 5.5). An inclusion of immunity into DRI models predicts a lower risk of illness incidence due to temporary induced protection, this confirms (Linnemann et al., 1984) that, there is limited risk of infection/illness among farmers due to acquired immunity from continuous exposure

Table 5.5: DALY(s) for Dose-Response Models with Epidemiological Transmission Dynamics of Norovirus

Scenarios/Models	Naïve Model	Immunity Model	Dose Model	Dose-Immunity Model
Symptomatic Individual Infectious	3.09×10^{-1}	1.51×10^{-5}	1.21×10^{-3}	5.7×10^{-8}
Pre-symptomatic and Post-symptomatic infectiousness (Low)	3.09×10^{-1}	1.47×10^{-5}	1.19×10^{-3}	5.65×10^{-8}
Pre-symptomatic and Post-symptomatic infectiousness (High)	3.09×10^{-1}	8.76×10^{-6}	1.20×10^{-3}	3.32×10^{-8}
Innate Genetic Resistance	3.09×10^{-1}	1.87×10^{-5}	1.21×10^{-3}	7.11×10^{-8}
Genogroup 2 Type 4 (GII.4)	3.09×10^{-1}	1.12×10^{-5}	1.19×10^{-3}	4.29×10^{-8}
No Immune Boosting by Asymptomatic Infection	3.09×10^{-1}	1.51×10^{-5}	1.21×10^{-3}	5.77×10^{-8}

5.6 Summary

This chapter presented various dose-response models for *Norovirus*, where the impact of the illness incidence was determined based on whether the dose-response models were based on immunity or dose-dependent probability. For immunity included dose-response models, illness incidence was low, whereas for dose-dependency models, the illness incidence shifted and a steep rise is observed. The naïve DRI model increases monotonically with increasing exposure frequency. The influence of pathogen dose was evident. At a low pathogen dose level, the illness incidence is affected by the choice of the DRI model (thus whether the model has induced immunity included or not), while at high pathogen dose, the impact of immunity protection dominates. Applying the models to the Norovirus data, all epidemiological

scenarios had the same trend of movement of the various dose-response models, and individual and population level of illness incidence reduction was much better measured by the dose-immunity DRI model, followed by the immunity DRI models. The study recorded a difference of 7 logs of magnitude less when the dose-immunity DRI model is used compared to the naïve model, whiles a 4 log of magnitude less is recorded if immunity alone is integrated to get the immunity DRI model as compared to the naïve model across all the transmission dynamics. Applying the DALY showed a similar trend of DRI models across all transmission dynamics of *NoV*. The immunity incorporated models tend to predict a lower incidence all year round, while the non-immunity incorporated models do not. It was also found that, the immunity dependent models (immunity and dose-immunity models) meet the WHO standard of $1.0X10^{-4}$. Besides the dose-immunity DRI model meets the more stringent WHO standard target of $1.0X10^{-6}$ in all *NoV* transmission scenarios. It is worth noting that, the transmission dynamics of *NoV* influence on predicting risk estimate is similar in all scenarios and tends to have a minimal difference in terms of values for each scenario.

CHAPTER 6

STATISTICAL MEASUREMENT MODELING FOR UNCERTAINTY QUANTIFICATION OF LOW QUALITY WATER EFFLUENT DISCHARGE

6.1 Introduction

This section presents measurement of uncertainty quantification for Waste Stabilization Pond (WSP) effluent discharge as against policy standard values for such discharge. In risk assessment for microbial pathogens, there is a strong correlation between risk estimates and the pathogen concentration in wastewater. Moreover, strong correlation also exists between physico-chemical parameters and pathogen concentrations as well. Total coliform count reveals an existence of strong positive association with temperature, turbidity, pH and alkalinity. Chloride, fluoride and Dissolve Oxygen (DO) are negatively correlated with total coliform count (Maheepal and Singh, 2014). In most cases significant positive correlation is observed between pollution indicator bacteria and pathogenic bacteria which may imply their co-presence.

The study integrates the use of policy standards and acceptable compliance level based on design model of WSP to establish refer charts for some physico-chemical and biological parameters measure for discharge effluents from WSP and treatment plants. Also, the study established the need for such a chart as a guide to monitor effluent discharge parameter values and help in controlling pathogenic

concentration reduction as a result of discharging effluents either onto streams or for irrigation purposes based on accepted Environmental Protection Agency (EPA) standards in Ghana.

6.2 Modelling Statistical Framework for Reliability

Several studies have defined reliability as the ability to perform the specified requirements free from failure (Niku et al.,1979) e.g. the percentage of times a wastewater treatment plant complies to discharge standards (Mcbride and Ellis, 2001; McBride, 2003; Smith et al., 2001). The WSPs will be completely reliable if the process performance does not violate the target standards of the regulatory bodies specifications (Oliveira and Von Sperling, 2008). Mathematically,

$$Failure = effluent\ concentration > effluent\ requirements \tag{6.1}$$

A risk of failure is always unavoidable, hence

$$Reliability = 1 - P(\ failure) \tag{6.2}$$

From equation 6.1, equation 6.2 becomes

$$Reliability = 1 - P(effluent\ concentration > effluent\ requirements) \tag{6.3}$$

In measuring the effluent discharge, a suitable distribution function to describe such effluent discharge is the use of lognormal distribution (Niku et al., 1979). The lognormal distribution owning to its deviation in symmetry measured by the skewness coefficient,

has positive skewness since there is usually a lower bound for effluent concentration, but there are no upper bounds.

Let $X=(X_1, X_2, ..., X_n)$ be a random variable effluent quality values of physic-chemical/biological property of low quality water having a lognormal distribution and μ and σ^2 respectively denoted the mean and the variance of Y where $Y=\ln(X)\sim N(\mu,\sigma^2)$. The probability density function of the lognormal distribution is

$$f(x,\mu,\sigma^2)=\begin{cases}\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left(-\dfrac{(\ln x-\mu)^2}{2\sigma^2}\right); \text{ for } x>0\\ 0 \qquad\qquad\qquad\qquad ; \text{ for } \le 0\end{cases} \qquad 6.4$$

The mean for the lognormal population is $E(X)=\exp\left(\mu+\sigma^2/2\right)$ where $E(X)$ denotes the expectation of X. For a known arithmetic mean and standard deviation of the effluent discharge values, then the location parameters can be determined, thus

$$\sigma^2=\sigma_{\ln x}^2=\ln\left(1+\frac{Var[X]}{\left(E[X]^2\right)}\right) \quad \text{and} \quad \mu=m_{\ln x}=\ln\left(E[X]\right)-\frac{1}{2}\ln\left(1+\frac{Var[X]}{\left(E[X]^2\right)}\right)=\ln\left(E[X]\right)-\frac{1}{2}\sigma^2,$$

Hence at different location parameter values, the probability density function and the cumulative density function are as shown.

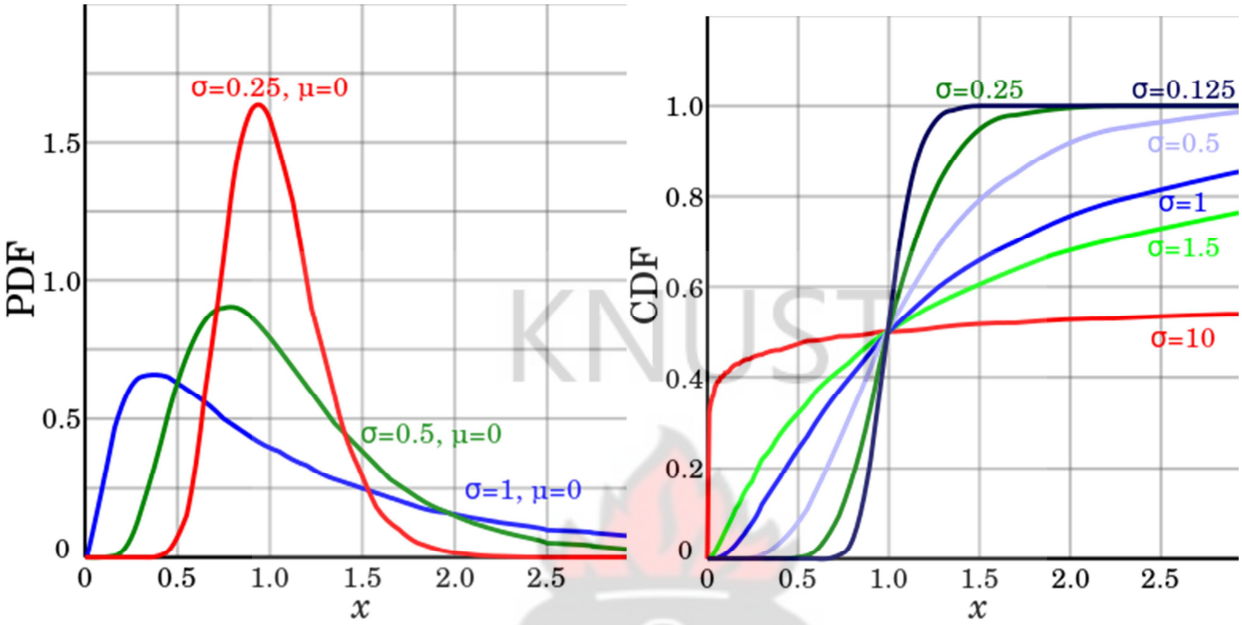


Figure 6.1: The probability density function and the cumulative density function of the log-normal distribution. Hence the density function for effluent quality is given as

$$f_x(x) = \frac{1}{x\sigma_{\ln x}\sqrt{2\pi}} \exp\left\{\frac{-1}{2}\left[\frac{1}{\sigma_{\ln x}} \ln\left(\frac{x}{\widetilde{m}_x}\right)\right]^2\right\} \quad x \geq 0 \qquad 6.5$$

Where x represents effluent variable concentration, $\sigma_{\ln x}$ represents standard deviation of the natural logarithm of X and $\widetilde{m}_x$ represents mean of x. For the r^{th} moment about the origin in the moment generation function.

$$\left.\begin{aligned} E\left(x^r\right) &= \left(\widetilde{m}_x\right)^r \exp\left[\frac{1}{2}r^2\sigma^2_{\ln x}\right] \\ m_x &= \widetilde{m}_x \exp\left(\frac{1}{2}\sigma^2_{\ln x}\right) \\ \sigma^2_x &= m^2_x\left[\exp\left(\sigma^2_{\ln x}\right)-1\right] \end{aligned}\right\} \quad 6.6$$

Where $E[X]=m_x$ and $Var[X]=\sigma_x$, thus equation (6.6) is the mean and variance of the original data of effluent discharge. Re-arranging equations 6.6 accounting for the relationship of parameters of probability density function of lognormal distribution in terms of moment of variable X leads to;

$$\left.\begin{aligned} \sigma^2_{\ln x} &= \ln\left(\frac{\sigma^2_x}{m^2_x}+1\right) \\ m_{\ln x} &= \ln m_x - \frac{1}{2}\sigma^2_{\ln x} \end{aligned}\right\} \quad 6.7$$

Where $m_{\ln x}$ is the average natural logarithm of X

The maximum likelihood estimation of parameters for the log-normal distribution parameters is represented by

$$f_L = (x, m_x, \sigma_{\ln x}) = \prod_{i=1}^{n}\left(\frac{1}{x_i}\right) f_N\left(\ln x; m_x, \sigma_{\ln x}\right),\ f_L$$

denotes the probability density function of the distribution and f_N that of the normal distribution, hence the log-likelihood function

$$l_L\left(m_x,\sigma_{\ln x}\middle|x_1,...,x_n\right)$$

$$=-\sum_i \ln x_i + l_N\left(m_x,\sigma_{\ln x}\middle|\ln x_1,...,\ln x_n\right)$$

$$=\varpi + l_L\left(m_x,\sigma_{\ln x}\middle|\ln x_1,...,\ln x_n\right)$$

hence it holds that, the logarithmic likelihood function reaches their maximum for mean and variance as;

$$\widehat{m_x}=\frac{\sum_i \ln x_i}{n}$$

$$\widehat{\sigma^2}=\frac{\sum_i\left(\ln x_i - m_x\right)^2}{n}$$

For some probability of failure at α the lognormal distribution will have a property of X, thus

$$P\left(X\le X_s\right)=1-\alpha \qquad 6.8$$

Where X_s is the effluent concentration standard fixed for policy assessment. Hence choosing the parameters of the lognormal distribution, equation 6.7 becomes

$$P\left(Z\le\frac{\ln X_s - m_{\ln x}}{\sigma_{\ln x}}\right)=1-\alpha \qquad 6.9$$

The standard Z normal distribution can also be defined from equation 6.10 as

$$P\left(Z\le Z_{1-\alpha}\right)=1-\alpha \qquad 6.10$$

Hence at reliability level of $1-\alpha$ of a failure level of α, a known standard of effluent concentration level could be calculated given a coefficient of variation, the $Z_{1-\alpha}$ values are as shown in Table 6.1 for the cumulative probability at $(1-\alpha)$ and its percentiles.

Table 6.1: Values of Standard Normal Distribution

Cumulative Probability $1-\alpha$	**Percentiles** $Z_{1-\alpha}$
50	0.000
60	0.253
70	0.525
80	0.842
90	1.282
92	1.405
95	1.645
98	2.054
99	2.326
99.9	3.090

It should be noted that, the higher the normal variate value the higher the corresponding compliance level (cumulative probability). Hence substituting equation 6.7 (mean and variance) into equation 6.9 leads to;

$$\frac{\ln X_s - \left[\ln m_x - \frac{1}{2}\ln\left(V_x^2+1\right)\right]}{\left[\ln\left(V_x^2+1\right)\right]^{\frac{1}{2}}} = Z_{1-\alpha} \qquad 6.11$$

Making the mean value the subject of equation 6.11 becomes

$$m_x = \left[\left(V_x^2+1\right)\right]^{\frac{1}{2}} \exp\left\{-Z_{1-\alpha}\left[\ln\left(V_x^2+1\right)\right]^{\frac{1}{2}}\right\} X_s \qquad 6.12$$

By simplification, equation 6.12 results

$$Z_{1-\alpha} = -\frac{\ln\left[\frac{m_x}{X_s}\left(V_x^2+1\right)\right]^{-\frac{1}{2}}}{\left[\ln\left(V_x^2+1\right)\right]^{\frac{1}{2}}} \qquad 6.13$$

The statistical parameters used in the reliability to relate the mean constituent value m_x to standard X_s defines the coefficient of variation (CV) as V_x

$$V_x = CV = \frac{\sigma_x}{m_x} \qquad 6.14$$

From equation 6.12, the Coefficient of Reliability (COR) is given as

$$COR = \left[\left(V_x^2+1\right)\right]^{\frac{1}{2}} \exp\left\{-Z_{1-\alpha}\left[\ln\left(V_x^2+1\right)\right]^{\frac{1}{2}}\right\} \qquad 6.15$$

Putting equation 6.14 into equation 6.15

$$COR = \left[\left(CV^2+1\right)\right]^{\frac{1}{2}} \exp\left\{-Z_{1-\alpha}\left[\ln\left(CV^2+1\right)\right]^{\frac{1}{2}}\right\} \qquad 6.16$$

The COR values are obtained as a function of coefficient of variation and reliability level, the different values of the coefficient of variation depends on the different mean and standard deviation parameters. Hence equation 6.16 becomes

$$m_x = COR\, X_S \qquad 6.17$$

Where X_S is the effluent quality standard;

COR is the coefficient of reliability, and m_x is the mean effluent concentration needed to achieve a certain compliance level of effluent quality standard.

6.3 Results on Statistical Measurement Modelling for Uncertainty Quantification

6.3.1 Simulated Results for Coefficient of Reliability

Simulated results indicate influences of coefficient of reliability as a function of coefficient of variation and normal-variate values show that, a lower coefficient of reliability is dependent on the high normal-variate value which signifies a lower failure rate and a high coefficient of variation. As shown in Figure 6.2, the CV values are inversely related to the COR values, nevertheless, the CV values are directly related to the normal variate values that measure the reliability.

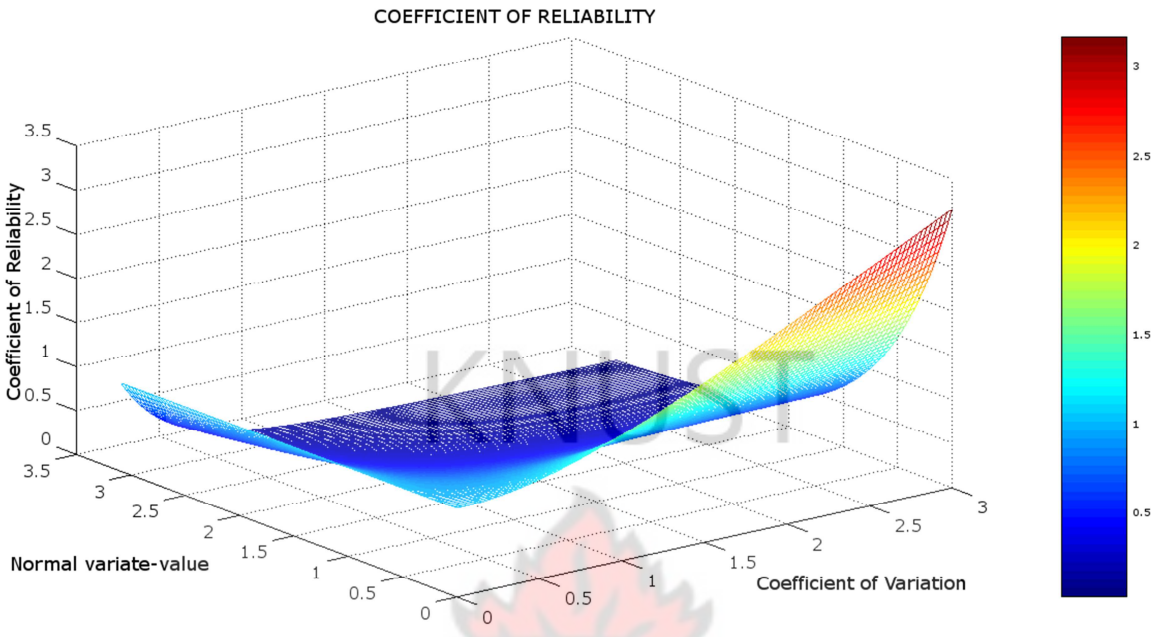


Figure 6.2: Simulated Coefficient of Reliability values

6.3.2 Results on Application of Development Reference Charts

Result of effluent discharge compliance and performance are much to be desired, when the effluent discharge is compared to the fixed standard value only and where it is assume that discharge concentration is less than the standard value. The performance is good and indicates a better compliance level of the WSPs. However with inclusion of reliability and compliance levels, the reference chart developed (Table 6.2 (See Appendix C also)), such an assumption is not always true.

From the developed chart, the required standard of effluent discharge concentration of BOD_5 or TSS is 50mg/L (EPA, Ghana Standard). If a sample taken from a specified WSP gives a mean effluent quality of 48.00 and with standard deviation of 14.4, will

have a CV value of 0.3, as indicated in Table 6.2 corresponds to a compliance level (COR) of 0.6 (60%) i.e. less than the required less stringent compliance level of 0.8 (80%) for WSPs (Oliveira and Von Sperling, 2008), This is despite the fact that the mean effluent concentration from the WSP as compared to the EPA standard falls below the standard value of 50mg/L and can be classified as a good discharged value. Nevertheless, the WSP is underperforming, its effluent discharge value is just less than that of a compliance level of 0.60, hence its compliance level is below what's generally accepted (even in a less stringent level of 0.80). Such information, if available can trigger a further check to be done to identify the segment of the wastewater stabilization pond (Anaerobic, facultative and maturation) that is underperforming, which could support the routine maintenance of the ponds.

The same procedure could be used by comparing the expected mean effluent concentration of the segments of the pond to its samples using its design compliance. Hence finding the compliance level of effluent to check for malfunctioning of pond segments, this is necessary due to the different expected work to be done by each segment to enhance the maintenance of the ponds regularly. Nonetheless, a critical look should be taken because CV values directly relate to reliability and inversely to COR values, the CV value with high standard deviation and lower mean of an effluent can have the same value as a CV of high mean and low standard deviation value of an effluent, the later shows a more

consistent discharge as shown in Table 6.2. This shows that, a lower value of CV does not necessarily indicate better results.

Moreover, with the use of the developed charts (Appendix C)for various parameters of WSPs in Ghana, once an effluent concentration average is known and its compliance level at design is also known, a quick reference point can be made to find what was expected to be discharging and compare to its current discharge to be assure of its compliance without necessary comparing it to fixed standard values. These reference charts were developed to serve as reference points in assessing the various characteristics of compliance and performance of WSPs in Ghana. Table 6.2 to 6.3 (See Appendix C for the rest) are intended to make it easier to assess the performance of WSPs and its corresponding reliability and compliance level without going through the task of using the log-normal procedure as shown above.

Table 6.2: Reference Chart of Compliance of Mean Effluent Discharge of BOD_5,TN and TSS for 50mg/L and Trichloroethylene, Benzene for 50 µg/l

COR	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.2	1.4	1.6	1.8	2	2.5	3
50%	50.00	50.25	50.99	52.20	53.85	55.90	58.31	61.03	64.03	67.27	70.71	78.10	86.02	94.34	102.96	111.80	134.63	158.11
60%	50.00	49.00	48.50	48.46	48.85	49.60	50.68	52.02	53.59	55.36	57.28	61.50	66.09	70.94	75.96	81.11	94.30	107.71
70%	50.00	47.69	45.95	44.75	43.99	43.62	43.58	43.81	44.26	44.89	45.67	47.57	49.78	52.21	54.78	57.44	64.30	71.28
80%	50.00	46.20	43.16	40.77	38.93	37.56	36.56	35.86	35.42	35.17	35.08	35.26	35.78	36.53	37.42	38.42	41.16	44.06
90%	50.00	44.22	39.56	35.83	32.86	30.51	28.64	27.16	25.99	25.06	24.32	23.27	22.63	22.25	22.05	21.98	22.16	22.60
92%	50.00	43.68	38.61	34.56	31.34	28.79	26.75	25.13	23.84	22.79	21.95	20.72	19.91	19.37	19.02	18.81	18.63	18.75
95%	50.00	42.64	36.81	32.21	28.57	25.70	23.42	21.60	20.13	18.95	17.98	16.52	15.50	14.78	14.26	13.87	13.29	13.03
98%	50.00	40.94	33.95	28.56	24.41	21.19	18.67	16.68	15.10	13.83	12.79	11.22	10.12	9.32	8.72	8.26	7.47	7.00
99%	50.00	39.84	32.17	26.37	21.98	18.63	16.05	14.05	12.47	11.21	10.20	8.68	7.63	6.86	6.29	5.85	5.10	4.64
999%	50.00	36.92	27.65	21.07	16.38	12.99	10.51	8.67	7.29	6.22	5.40	4.22	3.44	2.90	2.51	2.22	1.74	1.45

Table 6.3: Reference Chart of Compliance of Mean Effluent Discharge of TP for 2.0mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	2.00	2.01	2.04	2.09	2.15	2.24	2.33	2.44	2.56	2.69	2.83	3.12	3.44	3.77	4.12	4.47	5.39	6.32	7.28	8.25
60%	2.00	1.96	1.94	1.94	1.95	1.98	2.03	2.08	2.14	2.21	2.29	2.46	2.64	2.84	3.04	3.24	3.77	4.31	4.85	5.39
70%	2.00	1.91	1.84	1.79	1.76	1.74	1.74	1.75	1.77	1.80	1.83	1.90	1.99	2.09	2.19	2.30	2.57	2.85	3.13	3.41
80%	2.00	1.85	1.73	1.63	1.56	1.50	1.46	1.43	1.42	1.41	1.40	1.41	1.43	1.46	1.50	1.54	1.65	1.76	1.88	2.00
90%	2.00	1.77	1.58	1.43	1.31	1.22	1.15	1.09	1.04	1.00	0.97	0.93	0.91	0.89	0.88	0.88	0.89	0.90	0.93	0.95
92%	2.00	1.75	1.54	1.38	1.25	1.15	1.07	1.01	0.95	0.91	0.88	0.83	0.80	0.77	0.76	0.75	0.75	0.75	0.76	0.77
95%	2.00	1.71	1.47	1.29	1.14	1.03	0.94	0.86	0.81	0.76	0.72	0.66	0.62	0.59	0.57	0.55	0.53	0.52	0.52	0.52
98%	2.00	1.64	1.36	1.14	0.98	0.85	0.75	0.67	0.60	0.55	0.51	0.45	0.40	0.37	0.35	0.33	0.30	0.28	0.27	0.26
99%	2.00	1.59	1.29	1.05	0.88	0.75	0.64	0.56	0.50	0.45	0.41	0.35	0.31	0.27	0.25	0.23	0.20	0.19	0.17	0.16
99.9%	2.00	1.48	1.11	0.84	0.66	0.52	0.42	0.35	0.29	0.25	0.22	0.17	0.14	0.12	0.10	0.09	0.07	0.06	0.05	0.05

6.3.3 Statistical Measurement of Performance Based on Reference Chart

This study used two plants (KNUST waste stabilization pond, Ahinsan Waste Stabilization pond data from 2009 student projects). From the comparison of the various discharge qualities to the standards of EPA, it is very evident that, some of the parameters for the KNUST plant do not conform to EPA standards (Table 6.4). Though some exceptions like the temperature (24.66), TN (10.83), TC(79.69), pH (6.78) and turbidity (59.69) level, which recorded a lower discharge values than the EPA standard, all other parameters such as TSS (51.53), TP (12.2), BOD_5 (81.75) and E.coli (26.50) were higher than the standard. In contrast, the Ahinsan WSP was performing better in terms of discharge values than the KNUST plant. This WSP had most of its effluent discharge values lower than the EPA, which included; temperature (26.6), pH (7.3), TN (0.01), Ammonia (0.36), BOD_5 (38), COD (99), Conductivity (484), TDS (242). The performing discharge values of the Ahinsan WSP is attributed to some form of maintenance during the trial work of aqua-culture in the ponds, whereas, the KNUST plant did not receive any form of maintenance over quite a number of years, which can explain its under-performance.

Table 6.4: Effluent Discharge Values and the EPA standard

Parameter	**EPA standard**	**KNUST Plant**	**Ahinsan WSP**
Temperature (^{0}C)	< 30	24.66	26.60
pH	6 – 9	6.78	7.30
TSS (mg/L)	50	51.63	52.00
TP (mg/L)	2	12.20	6.10
Turbidity (NTU)	75	59.69	-
TN (mg/L)	50	10.83	0.01
Ammonia/Ammonium (mg/L)	1	-	0.36
BOD_5 (mg/L)	50	81.75	38.00
COD (mg/L)	250	-	99.00
Conductivity ($\mu S/cm$)	750	-	484.00
TDS (mg/L)	1500	-	242.00
DO (mg/L)	1	-	0.80
TC (MPN/100ml)	400		$1.7x10^8$
E.coli (MPN/100ml)	10	26.50	$7.1x10^5$

EPA's Discharge Standards to be achieved in Operation Concentration

The measured reliability of the KNUST treatment plant and Ahinsan WSP (Table 6.5) shows different compliance level of the discharge values to the standard values used for the design. Only two discharge values (TN; 98.65) and pH; 48.80 - 80.00) met the

less stringent design specification of 80.00 compliance on the KNUST plant and as well recorded an observed value less than its mean design concentration value (Table 6.6). The different compliance levels of the actual effluent discharge were: Temperature (73.60), TSS (63.70), TP (1.70), Turbidity (75.20), BOD_5 (63.30) and E.coli (31.90), and discharge values for temperature, pH, turbidity and TC were all lower than the EPA standards, but fall short of meeting the design compliance of 95 percent.

The Ahinsan WSP had five of its discharge values (TN, 99.90, Conductivity 98.30 Ammonia, 96.70; COD, 95.20 and TDS, 98.20) conforming to the standard compliance of 95 percent compliance and achieving its observed effluent discharge being less than the mean design concentration with the required compliance level (Table 6.6), whereas the rest were not complying with the design compliance level. These included; Temperature (69.50), pH (48.00 to 74.20), TSS (64.10), TP (8.40), BOD_5 (77.30), DO (74.90), TC (0.00) and E.coli (0.00). Again, temperature, pH, BOD_5, conductivity and DO effluent discharge meet the EPA standard but its compliance level does not meet the design specification.

Table 6.5: Actual Mean Effluent Discharge, Reliability and its Compliance

Parameters	**KNUST Treatment Plant**			**AHINSAN WSP**		
	Mean Effluent Discharge	Reliability	Compliance	Mean Effluent Discharge	Reliability	Compliance
Temperature (^{0}C)	24.66	0.63	73.60	26.60	0.51	69.50
pH	6.78	-0.03-0.84	48.8-80.0	7.30	-0.05-0.65	48.0-74.2
TSS (mg/L)	51.63	0.35	63.70	52.00	0.36	64.10
TP (mg/L)	12.20	-2.11	1.70	6.10	-1.38	8.40
Turbidity (NTU)	59.69	0.68	75.20	-	-	-
TN (mg/L)	10.83	2.21	98.65	0.01	17,09	99.90
Ammonia/Ammonium (mg/L)	-	-	-	0.36	1.84	96.70
BOD_5 (mg/L)	81.75	0.34	63.30	38.00	0.75	77.30
COD (mg/L)	-	-	-	99.00	1.66	95.20
Conductivity ($\mu S/cm$)	-	-	-	484.00	2.12	98.30
TDS (mg/L)				242.00	2.09	98.20
DO (mg/L)	-	-	-	0.80	0.67	74.90
TC (MPN/100ml)				1.7x10^8	-13.17	00.00
E.coli (MPN/100ml)	26.50	-0.47	31.90	7.1x10^5	-14.92	00.00

Table 6.6: Mean design effluent concentration to achieve 95% compliance with the standard and observed actual effluent concentrations

Parameters	**KNUST**				**AHINSAN**			
	CV	COR	Mean Design Conc.	Observed Actual Mean Conc.	CV	COR	Mean Design Conc.	Observed Actual Mean Conc.
Temperature (^{0}C)	0.57	0.48	14.40	24.66	0.72	0.43	12.90	26.60
pH	0.49	0.52	3.12	2.78	0.63	0.46	2.76	7.30
TSS (mg/L)	0.92	0.37	18.50	51.63	0.98	0.36	18.00	52.00
TP (mg/L)	0.84	0.39	0.78	12.20	0.73	0.42	0.84	6.10
Turbidity (NTU)	0.73	0.42	31.50	59.69	-	-	-	-
TN (mg/L)	1.04	0.35	17.54	10.83	0.54	0.49	24.50	0.01
Ammonia/Ammonium (mg/L)	-	-	-	-	0.77	0.41	0.41	0.36
BOD_5 (mg/L)	2.41	0.27	13.50	81.75	0.69	0.44	22.00	38.00
COD (mg/L)	-	-	-	-	0.81	0.40	100.00	99.00
Conductivity ($\mu S/cm$)	-	-	-	-	0.69	0.44	660.00	484.00
TDS (mg/L)	-	-	-		1.03	0.35	350.00	242
DO (mg/L)				-	0.85	0.39	0.39	0.80
TC (MPN/100ml)					1.21	0.33	132.00	$1.7x10^8$
E.coli (MPN/100ml)	1.32	0.32	3.20	26.50	0.84	0.39	3.90	$7.1x10^5$

6.4 Summary

In this chapter, uncertainty measurements were measured with a log-normal distribution function. A measurement table incorporating compliance level, reliability and standard policy value for effluent discharge were used to estimate various coefficient of variation. Coeffecient of reliability were obtained based on the discharge values and the designed compliance level of the waste stabilization ponds. Compliance measurements were found that, by comparison of discharge values to policy standard values for acceptance. In some cases, effluent values might meet standard value by comparison by fails to meet the expected value when compliance is to be measured as part of acceptable effluent discharge value.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the conclusions and recommendations of the study.

Probabilistic quantitative risk assessment models were presented, illustrating the impact of two contrasting ideas, namely the use of fecal indicators as a ratio conversion method due to lack of specific pathogen data, and the use of the scarce pathogen data available as a best alternative to the ratio conversion method. The quantitative model was analysed and the simulation results were discussed. Extension of the model to include illness incidence DR models were presented for the case of Norovirus and different population epidemiological dynamics were integrated into the model to find its effects on predicting illness given infection. An all-inclusive modelling of measuring effluent discharge was presented by integrating both reliability and compliance into such measurements. The results obtained were presented in Chapter 4, 5 and 6 respectively and shall be summarized in this chapter.

7.2 Conclusion Findings

The study revealed the following;

7.2.1 Risk Assessment using Dose Estimates based onGenome Copies and Conversion Ratio

We estimated the risk of illness and disease burden with the use of fecal indicator ratio conversion or genome copies *Norovirus* for consumption of vegetables in

Ghana. A QMRA model was developed to estimate the differences in disease burden, and the results showed that:

1. All model scenarios for consuming vegetables irrigated with wastewater (stream or drain) met the 10^{-4} DALY pppy threshold for Norovirus. However, models that use genomic copies *Norovirus* are considered highly conservative estimates.
2. In all cases, stream water recorded a higher probability of illness and disease burden than the drain water sources and again represents conservative estimates due to insufficient data availability.
3. In the model of the same scenarios, the use of fecal indicator ratio conversion tends to underestimate the risk of disease burden DALY pppy as compared to the use of genome copies of Norovirus. This indicates that a shift from using fecal indicator to data on the actual pathogen (virus) of interest might give a more realistic output of the risk estimates.
4. A 2 order of magnitude was recorded in terms of differences in DALY for fecal indicator ratio conversion and genome copies Norovirus for stream and drain water. However, when pooled data were used, more than 1 order magnitude difference was recorded. Cryptosporidium spp. also showed a similar difference of close to a 2 order of magnitude difference in DALY for fecal ratio conversion and oocyst data.

7.2.2 Results on Illness Incidence Model with Induced Immunity

The adoption of the best DR model to estimate the illness dependency on infection/illness as well as DALY shows, inclusion of dose-dependent and immunity, substantially reduces the uncertainty surrounding the estimation of illness due to infection to a tune of ≥ 4 logs order of magnitude less than the naïve DR model, as immunity plays a substantial role in estimating the illness. The study revealed that transmission dynamics on response to disease infectiousness in epidemiological modelling do not have a significant impact on the extent of illness reduction given infection in terms of magnitude orders but a slight difference in values under the same order of magnitude. On the part of the DALY calculation, the inclusion of immunity in the dose-immunity DR model was found to result in acceptable pathogen levels under WHO stringent condition of $1.0X10^{-6}$ In all scenarios for the epidemiological analysis. The dose-immunity DRI model was found to be performing better and hence gives the best estimate as copare to Naïve, Dose and Immunity DRI models. Nevertheless, the immunity DRI model without the dose-dependent inclusion also had a significant difference of ≥ 4 logs order of magnitude less as compared to the naïve DR model and ≥ 2 logs order of magnitude as compared to dose DR model for DALYs, and can serve as an alternative in other scenarios where the dose-dependent DR model estimation is not available for use.

7.2.3 Statistical Measurement for Quantifying Uncertainty in Low Quality Water Effluent

From this study, it was observed that, measuring performance of WSP and treatment plants using effluent discharge values in comparison to standards alone is sufficient only for knowing effluent quality but cannot be used to evaluate compliance of the WSP or treatment plant. It is evident that compliance that considers both effluent quality discharges and design capability and its performance measure, is more appropriate than the use of removal efficiency and fixed standard values alone. In this study, we developed reference charts (Appendix C: Supplementary Results), which can be used for assessing effluent discharge qualities. These were carried out for different compliance levels from the Ghana EPA standard discharge values.

The importance of a stable operation and thus low CV should be remembered at all times, so that the WSP or treatment plant should not need to be designed to achieve very low mean effluent concentration. The effluent discharge values of the sites used for the study were not complying fully with the design specification (for the less stringent specifications of WSP and treatment plant). However, the Ahinsan WSP had some of its water quality parameters (TN, Ammonia, TDS and COD) meeting both the compliance level at 95% and the EPA discharge standards. Irrespective of the presence of two maturation ponds in series for the Ahinsan WSP, it could not meet the pathogen reduction standard values expected.

7.3 Limitations and Recommendations

In this section, the limitation and recommendations of the study are presented. The limitations are basically based on the various models built in this study. The recommendations are based on the outcome of the study.

7.3.1 Limitations

The study has some limitations with regards to the modelling process and the data validation, just like all other quantitative modelling approach, using probability distributions, this add up to the model uncertainty as some of the assumptions made and data applied are based on various studies differing in geographical location, agricultural practices and human race. Response to disease is also sensitive to geographical locations, genetic make-up and environmental conditions. In this study, not all model parameters used primary data that could be fitted and the predictions validated against some observed outputs. Hence, we relied on assumptions of probability distributions and theoretical theorems for such estimations. In some cases, scarce data were used and some were combined with experts' opinion. Though such opinions are accepted in quantitative risk modelling, their subjectivity in nature makes them less desirable for uncertainty quantification.

7.3.2 Recommendations

The study revealed that, modelling risk of infection and illness with ratio conversion method underestimated the risk of infection and hence can lead to an insufficient mitigation process. The study also revealed that in the face of scarce data for Norovirus, using the pathogen of interest estimation was far better than the use of ratio conversion, which was also confirmed forCryptosporidium spp, for which sufficient data for quantifying its distribution were available.

The study further established that, estimating illness incidence is better measured when dose-dependency and induced immunity are included in the dose-response modelling step. It is recommended that such models are used in estimating risk of illness given infection. However, in cases, where dose-dependent models are unavailable, immunity DR models are recommended to be used instead of the naïve DR model.

The study also revealed that measures of effluent discharge that are compared to policy standard should be used in conjunction with expected compliance and reliability level. For measurements, where the effluent doesn't record negative values and follows a log-normal distribution, it is recommended that the chart produced by this study are used as a guide to measure the effluent outputs with a known compliance level. Finally, this study can guide towards modelling risk assessment from

the water treatment to estimating illness given infection for practitioners in the field of risk assessment modelling.

7.4 Areas of Possible Further Research

The avenue for further research based on the resuts of this study is outlined below:

- Application of different methods for characterising mixed distributions to account for variability in cases where such variability description is ambiguous in nature.
- Estimation of shape parameter using different pathogen concentrations to pave a way for more dose-dependent DR models to be constructed, which will make the dose-immunity DR models available for most pathogens.
- Estimation and modelling of dose-dependent models for other pathogens of interest with the use of parsimonious DR models like fractional poison, exponential model and other empirical models.
- Studies to involve different mathematical epidemiological models to construct appropriate illness inflation factor for various pathogen transmission dynamics

REFERENCES

Ackerson, N., Awuah, E., (2012). Microbial risk assessment of urban agricultural farming: A case study on Kwame Nkrumah University of Science and Technology campus, Kumasi, Ghana. … Technol. 2, 356–363.

Albert, I., Grenier, E., Denis, J., Rousseau, J., (2008). Quantitative Risk Assessment from Farm to Fork and Beyond: A Global Bayesian Approach Concerning Food-Borne Diseases. Risk Anal. 28. doi:10.1111/j.1539-6924.2008.01000.x

Allwood, P., Malik, Y., (2004). Effect of temperature and sanitizers on the survival of feline calicivirus, Escherichia coli, and F-specific coliphage MS2 on leafy salad vegetables. J. Food … 67, 1451–1456.

Amoah, P., Drechsel, P., (2007). Effectiveness of common and improved sanitary washing methods in selected cities of West Africa for the reduction of coliform bacteria and helminth eggs on. Trop. Med. … 12, 40–50. doi:10.1111/j.1365-3156.2007.01940.x

Amoah, P., Drechsel, P., Abaidoo, R., (2005). Irrigated urban vegetable production in Ghana: Sources of pathogen contamination and health risk elimination. Irrig. Drain. 61, 49–61.

Amoah, P., Drechsel, P., Abaidoo, R.C., Ntow, W.J., (2006). Pesticide and pathogen contamination of vegetables in Ghana's urban markets. Arch. Environ. Contam. Toxicol. 50, 1–6. doi:10.1007/s00244-004-0054-8

Amoah, P., Drechsel, P., Henseler, M., Abaidoo, R.C., (2007). Irrigated urban vegetable production in Ghana: microbiological contamination in farms and markets and associated consumer risk groups. J. Water Health 5, 455–66. doi:10.2166/wh.2007.041

An, Y.-J., Yoon, C.G., Jung, K.-W., Ham, J.-H., (2007). Estimating the microbial risk of E. coli in reclaimed wastewater irrigation on paddy field. Environ. Monit. Assess. 129, 53–60. doi:10.1007/s10661-006-9425-0

Andersson, H., Britton, T., (2000). Stochastic epidemic models and their statistical analysis. Springer Lect. Notes Stat.

Anonymous, (2014). Water for agriculture [WWW Document]. UNESCO. URL http://webworld.unesco.org/water/ihp/publications/waterway/webpc/pag19.ht

ml (accessed 6.7.15).

Armstrong, T.W., (2005). by Armstrong, Thomas W. Drexel University.

Atmar, R.L., (2010). Noroviruses - State of the Art. Food Environ. Virol. 2, 117–126. doi:10.1007/s12560-010-9038-1

Atmar, R.L., Opekun, A.R., Gilger, M. a., Estes, M.K., Crawford, S.E., Neill, F.H., Graham, D.Y., (2008). Norwalk virus shedding after experimental human infection. Emerg. Infect. Dis. 14, 1553–1557. doi:10.3201/eid1410.080117

Ayuso-Gabella, N., Page, D., Masciopinto, C., Aharoni, A., Salgot, M., Wintgens, T., (2011). Quantifying the effect of Managed Aquifer Recharge on the microbiological human health risks of irrigating crops with recycled water. Agric. Water Manag. 99, 93–102. doi:10.1016/j.agwat.2011.07.014

Ayyub, B.M., Klir, G.J., (2006). Uncertainty Modeling and Analysis in Engineering and the Sciences.

Bae, H.R., Grandhi, R. V., Canfield, R. a., (2004). Epistemic uncertainty quantification techniques including evidence theory for large-scale structures. Comput. Struct. 82, 1101–1112. doi:10.1016/j.compstruc.2004.03.014

Bae, H.-R., Grandhi, R. V., Canfield, R. a., (2004). An approximation approach for uncertainty quantification using evidence theory. Reliab. Eng. Syst. Saf. 86, 215–225. doi:10.1016/j.ress.2004.01.011

Baert, L., Uyttendaele, M., (2008). Survival and transfer of murine norovirus 1, a surrogate for human noroviruses, during the production process of deep-frozen onions and spinach. J. Food … 71, 1590–1597.

Baert, L., Vandekinderen, I., Devlieghere, F., Coillie, E.L.S.V.A.N., Debevere, J., (2009). sodium hypochlorite and peroxyacetic acid to reduce murine norovirus 1, B40-8, Listeria monocytogenes, and Escherichia coli O157: H7 on shredded iceberg lettuce. J. Food … 72, 1047–1054.

Barker, S.F., (2014). Risk of norovirus gastroenteritis from consumption of vegetables irrigated with highly treated municipal wastewater-evaluation of methods to estimate sewage quality. Risk Anal. 34, 803–17. doi:10.1111/risa.12138

Barker, S.F., Amoah, P., Drechsel, P., (2014). A probabilistic model of gastroenteritis risks associated with consumption of street food salads in Kumasi, Ghana: Evaluation

of methods to estimate pathogen dose from water, produce or food quality. Sci. Total Environ. 487, 130–42. doi:10.1016/j.scitotenv.2014.03.108

Benavoli, A., Ristic, B., Chisci, L., Farina, A., (2009). Modelling Uncertain Implication Theory in Evidence.

Boomsma, W., Mardia, K. V, Taylor, C.C., Ferkinghoff-Borg, J., Krogh, A., Hamelryck, T., (2008). A generative, probabilistic model of local protein structure. Proc. Natl. Acad. Sci. U. S. A. 105, 8932–7. doi:10.1073/pnas.0801715105

Casella, G., George I, E., (1992). Explaining the Gibbs sampler. Am. Stat. 46, 167–174.

Charles, N. Haas; Joan, B. Rose; Charles, P.G., (1999). Quantitative Microbial Risk Assessment -. Wiley.

Chiyaka, C., Garira, W., Dube, S., (2010). Using Mathematics to Understand Malaria Infection During Erythrocytic Stages. Infection 5, 1–11.

Coughlan, J., (2010). Maximum Entropy Distributions and Their Re- lationship to Maximum Likelihood.

Crabtree, K., Gerba, C., Rose, J., Haas, C., (1997). Waterborne adenovirus: a risk assessment. Water Sci. … 35, 1–6.

Croci, L., De Medici, D., Scalfaro, C., Fiore, A., Toti, L., (2002). The survival of hepatitis A virus in fresh produce. Int. J. Food Microbiol. 73, 29–34. doi:10.1016/S0168-1605(01)00689-4

Dawber, J., Horn, B., Brown, J., (2009). Performing a Quantitative Microbial Risk Analysis Using Second-Order Monte Carlo Simulation, in: Third Annual ASEARC Conference.

Drechsel, P., Keraita, B., Seidu, R., Abaidoo, R.C., (2007). Human Health Risks from Wastewater-irrigated Vegetable Farming 104–115.

Drechsel, P., Scott, C., Raschid-sally, L., Redwood, M., Bahri, A., (2009). Wastewater irrigation and health, International Water Management Institute and International Development Research Centre. Earthscan, London.

Duarte, S.R., Nauta, M.J., (2015). Impact of microbial count distributions on human health risk estimates. Int. J. Food Microbiol. 195, 48–57. doi:10.1016/j.ijfoodmicro.2014.11.024

Dudley, R.H., Hekimian, K.K., Mechalas, B.J., Dudley, H., Hekimian, K., M, J., (1976). A scientific basis for determining recreational water quality. Water Environ. Fed. 48, 2761–2777.

Elod, B., (2013). Maximum Entropy Models.

Ferson, S., Ginzburg, L.R., (1996). Different methods are needed to propagate ignorance and variability. Reliab. Eng. Syst. Saf. 54, 133–144. doi:10.1016/S0951-8320(96)00071-3

Fiona Barker, S., O'Toole, J., Sinclair, M.I., Leder, K., Malawaraarachchi, M., Hamilton, A.J., (2013). A probabilistic model of norovirus disease burden associated with greywater irrigation of home-produced lettuce in Melbourne, Australia. Water Res. 47, 1421–32. doi:10.1016/j.watres.2012.12.012

Flannery, J., Keaveney, S., Rajko-nenow, P., Flaherty, V.O., Doré, W., (2012). Concentration of Norovirus during Wastewater Treatment and Its. Appl. Environ. Microbiol. 78. doi:10.1128/AEM.07569-11

Food and Agricultural Organization, W.H., (2003). Hazard characterization for pathogens in food and water: guidelines.

Frenck, R., Bernstein, D.I., Xia, M., Huang, P., Zhong, W., Parker, S., Dickey, M., McNeal, M., Jiang, X., (2012). Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans. J. Infect. Dis. 206, 1386–1393. doi:10.1093/infdis/jis514

Fuhs, W.., (1975). A PROBABILISTIC MODEL OF BATHING BEACH SAFETY An improved mathematical model for bathing beach safety is proposed . It is derived by joining the probability of infection from a given dose (Poisson distribution and the probability of acquiring such a dose. Sci. Total Environ. 4, 165–175.

Fung, J., (2011). Microbiological quality of urban-vended salad and its association with gastrointestinal diseases in Kumasi, Ghana. Int. J. Food Safety, Nutr. Public Heal. 4, 152 – 166.

Furman, M., Pivi, M., (2002). Probabilistic model for the simulation of secondary electron emission. Phys. Rev. Spec. Top. - Accel. Beams 5, 124404. doi:10.1103/PhysRevSTAB.5.124404

Furumoto, W. a, Mickey, R., (1970). Mathematical analyses of the interference

phenomenon of tobacco mosaic virus: experimental tests. Virology 40, 322–328. doi:10.1016/0042-6822(70)90408-3

Furumoto, W.A., Mickey, R., (1967a). A mathematical model for the infectivity-dilution curve of tobacco mosaic virus: Theoretical considerations. Virology 32, 216–223. doi:http://dx.doi.org/10.1016/0042-6822(67)90271-1

Furumoto, W.A., Ray, M., (1967b). A Mathematical Model for the Infectivity-Dilution Curve of Tobacco Mosaic Virsu: Experimental Tests. Virology 32, 224–233.

Glass, R.I., Parashar, U.D., Estes, M.K., (2009). Norovirus Gastroenteritis. N. Engl. J. Med. 361, 1776–1785. doi:10.1056/NEJMra0804575

Greenwood, Major; Yule, U.., (1920). An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents on JSTOR. R. Stat. Soc. 83, 225–279.

Greiner, M., Smid, J., Havelaar, A.H., Müller-Graf, C., (2013). Graphical models and Bayesian domains in risk modelling: application in microbiological risk assessment. Prev. Vet. Med. 110, 4–11. doi:10.1016/j.prevetmed.2013.02.008

Hall, J.N., Moore, S., Harper, S.B., Lynch, J.W., (2009). Global Variability in Fruit and Vegetable Consumption. Am. J. Prev. Med. 36, 402–409.e5. doi:10.1016/j.amepre.2009.01.029

Hamby, D., (1995). A comparison of sensitivity analysis techniques. Health Phys. 1–20.

Hamby, D.M., (1994). A review of techniques for parameter sensitivity analysis o f e n v i r o n m e n t a l m o d e l s d. m. hamby* 135–154.

Hamilton, A.J., Stagnitti, F., Premier, R., Boland, A., Hale, G., (2006). Quantitative Microbial Risk Assessment Models for Consumption of Raw Vegetables Irrigated with Reclaimed Water Quantitative Microbial Risk Assessment Models for Consumption of Raw Vegetables Irrigated with Reclaimed Water. Appl. Environ. Microbiol. 72. doi:10.1128/AEM.72.5.3284

Haramoto, E., Katayama, H., Oguma, K., Yamashita, H., Tajima, A., Nakajima, H., Ohgaki, S., (2006). Seasonal profiles of human noroviruses and indicator bacteria in a wastewater treatment plant in Tokyo, Japan. Water Sci. Technol. 54, 301. doi:10.2166/wst.2006.888

Harremoës, P., Topsøe, F., (2001). Maximum entropy fundamentals. Entropy 3, 191–226. doi:10.3390/e3030191

Hassine-Zaafrane, M., Sdiri-Loulizi, K., Kaplon, J., Salem, I. Ben, Pothier, P., Aouni, M., Ambert-Balay, K., (2014). Molecular Detection of human Noroviruses in Influent and Effluent Samples From Two Biological Sewage Treatment Plants in the Region of Monastir, Tunisia. Food Environ. Virol. 6, 125 – 131. doi:10.1007/s12560-014-9147-3

Havelaar, A.H., Swart, A.N., (2014). Impact of Acquired Immunity and Dose-Dependent Probability of Illness on Quantitative Microbial Risk Assessment. Risk Anal. 34, 1807–1819. doi:10.1111/risa.12214

Heath, J., Kwiatkowska, M., Norman, G., Parker, D., Tymchyshyn, O., (2008). Probabilistic model checking of complex biological pathways. Theor. Comput. Sci. 391, 239–257. doi:10.1016/j.tcs.2007.11.013

Hegazy, B.E.E., (2013). Evaluation Of Waste Stabilization Ponds Performance In Egypt. J. Appl. Sci. Res. 9, 638–642.

Heidinger, J.C., (2009). Open Access Dissertation - ProQuest. University of California.

Heylings, J.R., (2011). Transdermal and Topical Drug Delivery, Transdermal and Topical Drug Delivery: Principles and Practice. John Wiley & Sons, Inc., Hoboken, NJ, USA. doi:10.1002/9781118140505

Hogg, R., McKean, J., Craig, A., (2005). Introduction to Mathematical Statistics (6th Edition). All Books Monogr. by WMU Authors.

Huynen, P., Mauroy, A., Martin, C., Savadogo, L.G.B., Boreux, R., Thiry, E., Melin, P., De Mol, P., (2013). Molecular epidemiology of norovirus infections in symptomatic and asymptomatic children from Bobo Dioulasso, Burkina Faso. J. Clin. Virol. 58, 515–521. doi:10.1016/j.jcv.2013.08.013

Haas, C.N., (2014). QUANTITATIVE MICROBIAL RISK ASSESSMENT.

Haas, Charles N; Rose , Joan B; Gerba, C.P., (2014). Introduction : Models for Population, in: Quantitative Microbial Risk Assessment 2nd Edition. pp. 377–398.

Haas, Charles N; Rose, Joan B; Gerba, C.P., (2014a). Risk Assessment Paradigms, in: Quantitative Microbial Risk Assessment. pp. 63–89.

Haas, Charles N; Rose, Joan B; Gerba, C.P., (2014b). Conducting the Dose-Response

Assessment, in: Quantitative Microbial Risk Assessment. pp. 267–321.

Haas, Charles N; Rose, Joan B; Gerba, C.P., (2014a). Analytical Methods and the QMRA Framework: Developing Occurrnece and Exposure Databases, in: Quantitative Microbial Risk Assessment. pp. 129–157.

Haas, Charles N; Rose, Joan B; Gerba, C.P., (2014b). Uncertainty, in: Quantitative Microbial Risk Assessment. pp. 323–375.

Jaynes, E.T., (2003). Probability Theory: The Logic of Science. Cambridge University Press.

Jiménez, B., (2006). Irrigation in Developing Countries Using Wastewater. Int. Rev. Environ. Strateg. 6, 229–250.

Jimenez, B., Asano, T., (2008). Water Reuse: An International Survey of current practice, issues and needs.

Johansen, S., Juselius, K., (2009). MAXIMUM LIKELIHOOD ESTIMATION AND INFERENCE ON COINTEGRATION - WITH APPLICATIONS TO THE DEMAND FOR MONEY. Oxf. Bull. Econ. Stat. 52, 169–210. doi:10.1111/j.1468-0084.1990.mp52002003.x

Karavarsamis, N., Hamilton, a J., (2010). Estimators of annual probability of infection for quantitative microbial risk assessment. J. Water Health 8, 365–73. doi:10.2166/wh.2010.045

Katayama, H., Haramoto, E., Oguma, K., Yamashita, H., Tajima, A., Nakajima, H., Ohgaki, S., (2008). One-year monthly quantitative survey of noroviruses, enteroviruses, and adenoviruses in wastewater collected from six plants in Japan. Water Res. 42, 1441–8. doi:10.1016/j.watres.2007.10.029

Keraita, B; Dreschsel, P; Hulbers, F; Raschid-Sally, L., (2002). Wastewater Use in Informal Irrigation in Urban and Periurban Areas of Kumasi, Ghana | The RUAF Foundation. Urban Agric. Mag. 11–13.

Keramat, M., Kielbasa, R., (1999). Modified Latin Hypercube Sampling Monte Carlo (MLHSMC) Estimation for Average Quality Index 19.

Kermack, W O, McKendrick, a. G., (1927). A contribution to the mathematical theory of epidemics.

Kotz, S., Read, C.B., Balakrishnan, N., Vidakovic, B., Johnson, N.L. (Eds.), (2004).

Encyclopedia of Statistical Sciences. John Wiley & Sons, Inc., Hoboken, NJ, USA. doi:10.1002/0471667196

La Rosa, G., Iaconelli, M., Pourshaban, M., Muscillo, M., (2010). Detection and molecular characterization of noroviruses from five sewage treatment plants in central Italy. Water Res. 44, 1777–84. doi:10.1016/j.watres.2009.11.055

Labite, H., Lunani, I., van der Steen, P., Vairavamoorthy, K., Drechsel, P., Lens, P., (2010). Quantitative Microbial Risk Analysis to evaluate health effects of interventions in the urban water system of Accra, Ghana. J. Water Health 8, 417–30. doi:10.2166/wh.2010.021

Lawhon, P., Schwartz, M., (2006). Linking environmental and economic sustainability in establishing standards for wastewater re-use in Israel. Water Sci. Technol. 53, 203–212.

Li, Y., Ruan, S., Xiao, D., (2011). The Within-Host dynamics of malaria infection with immune response. Math. Biosci. Eng. 8, 999–1018. doi:10.3934/mbe.2011.8.999

Linnemann, C.C., Jaffa, R., Gartside, P.S., Scarpino, P.V., Clark, C.S., (1984). Risk of Infection Associated with Spray Irrigation System Used for Farming. J. Occup. Med. 26.

Lundqvist, P.J., Raschid-Sally, L., (2013). City Waste for Agriculture: Emerging Priorities which Influence Agenda Setting. Aquat. Procedia 1, 88–99. doi:10.1016/j.aqpro.2013.07.008

Machdar, E., van der Steen, N.P., Raschid-Sally, L., Lens, P.N.L., (2013). Application of Quantitative Microbial Risk Assessment to analyze the public health risk from poor drinking water quality in a low income area in Accra, Ghana. Sci. Total Environ. 449, 134–42. doi:10.1016/j.scitotenv.2013.01.048

Maheepal, S., Singh, S., (2014). CORRELATIVE STUDY OF PHYSICO-CHEMICAL AND MICROBIOLOGICAL PARAMETERS OF RADHA AND SHYAM KUND GOVARDHAN , MATHURA (U . P). Exp. 25, 1726–1735.

Mara D. , Sleigh, A., (2010). Estimation of Norovirus and Ascaris infection risks to urban farmers in developing countries using wastewater for Crop irrigation. Water Heal. 8, 572 – 576.

Mara, D., Hamilton, A., (2010). Tools for risk analysis: updating the 2006 WHO

guidelines. astewater … 7.

Mara, D., Hamilton, A., Sleigh, A., Karavarsamis, N., (2010). Options for Updating the 2006 WHO Guidelines.

Mara, D., Sleigh, A., (2010). Estimation of norovirus infection risks to consumers of wastewater-irrigated food crops eaten raw. J. Water Health 8, 39–43. doi:10.2166/wh.2009.140

Mara, D., Sleigh, A., (2009). Understanding and Updating the 2006 WHO Guidelines for the Safe Use of Wastewater in Agriculture 1–2.

Mara, D.D., Sleigh, P. a., Blumenthal, U.J., Carr, R.M., (2007). Health risks in wastewater irrigation: Comparing estimates from quantitative microbial risk analyses and epidemiological studies. J. Water Health 5, 39. doi:10.2166/wh.2006.055

Mason, S.J., Hill, R.R., Mönch, L., Rose, O., Jefferson, T., Fowler, J.W., (2008). Proceedings of the 2008 Winter Simulation Conference S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds. 91–100.

Mattison, K., (2011). Norovirus as a foodborne disease hazard. Adv. Food Nutr. Res. 62, 1–39. doi:10.1016/B978-0-12-385989-1.00001-6

McBride, G.B., (2003). Confidence of compliance: parametric versus nonparametric approaches. Water Res. 37, 3666–71. doi:10.1016/S0043-1354(03)00272-0

Mcbride, G.B., Ellis, J.C., (2001). Confidence of compliance: a bayesian approach for percentile standards. Water Res. 35, 1117–1124. doi:10.1016/S0043-1354(00)00536-4

Medema, G., Ashbolt, N., (2006). QMRA: its value for risk management. Rep. Microrisk Proj. Available online http//www ….

Medlock, J., (2009). Mathematical Analysis of Epidemiological Models II.

Messner, M.J., Berger, P., Nappier, S.P., (2014). Fractional Poisson—A Simple Dose-Response Model for Human Norovirus. Risk Anal. 34, 1820–1829. doi:10.1111/risa.12207

Mickey, R.A.Y., Furumoto, W.A., (1970). Mathematical Analysis of the Inference Phenomenon of Tobacco Mosaic Virus:Theorectical Consideration. Virology 321, 316–321.

Mitakakis, T., (2004). Food safety in family homes in Melbourne, Australia. J. Food Prot. 67, 818 – 822.

Mofarrah, A., Husain, T., (2010). Modeling for Uncertainty Assessment in Human Health Risk Quantification : A Fuzzy Based approach.

Mok, H.-F., Barker, S.F., Hamilton, A.J., (2014). A probabilistic quantitative microbial risk assessment model of norovirus disease burden from wastewater irrigation of vegetables in Shepparton, Australia. Water Res. 54, 347–62. doi:10.1016/j.watres.2014.01.060

Mok, H.-F., Hamilton, A.J., (2014). Exposure factors for wastewater-irrigated Asian vegetables and a probabilistic rotavirus disease burden model for their consumption. Risk Anal. 34, 602–13. doi:10.1111/risa.12178

Mota, A., Mena, K.D., Soto-Beltran, M., Tarwater, P.M., Cháidez, C., (2009). Risk assessment of cryptosporidium and giardia in water irrigating fresh produce in Mexico. J. Food Prot. 72, 2184–2188.

National Geographic Report, O., (2015). Clean Water Crisis, Water Crisis Facts, Water Crisis Resources - National Geographic [WWW Document]. Natl. Geogr. Channel. URL http://environment.nationalgeographic.com/environment/freshwater/freshwater-crisis/ (accessed 6.7.15).

Ndiaye, M.L., Niang, S., Pfeifer, H., Peduzzi, R., Tonolla, M., Dieng, Y., (2011). Effect of irrigation water and processing on the microbial quality of lettuces produced and sold on markets in Dakar (Senegal). Irrig. … 517, 509–517.

Niku, S., Schroeder, E.D., Samaniego, F.J., Edward, D., Samaniego, J., (1979). All use subject to JSTOR Terms and Conditions Performance sludge processes of activated and design reliability-based. Water Pollut. Control Fed. 51, 2841–2857.

Nordgren, J., Kindberg, E., Lindgren, P.E., Matussek, A., Svensson, L., (2010). Norovirus gastroenteritis outbreak with a secretor-independent susceptibility pattern, Sweden. Emerg. Infect. Dis. 16, 81–87. doi:10.3201/eid1601.090633

Ntzoufras, I., (2009). Markov Chain Monte Carlo Algorithms in Bayesian Inference. Bayesian Model. Using WinBUGS.

Oliveira, S.C., Von Sperling, M., (2008). Reliability analysis of wastewater treatment

plants. Water Res. 42, 1182–94. doi:10.1016/j.watres.2007.09.001

Ozawa, K., Oka, T., Takeda, N., Hansman, G.S., (2007). Norovirus infections in symptomatic and asymptomatic food handlers in Japan. J. Clin. Microbiol. 45, 3996–4005. doi:10.1128/JCM.01516-07

Palmer, T.N., Doblas-Reyes, F.J., Hagedorn, R., Weisheimer, A., (2005). Probabilistic prediction of climate using multi-model ensembles: from basics to applications. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 1991–8. doi:10.1098/rstb.2005.1750

Parsons, D.J., Orton, T.G., D'Souza, J., Moore, a, Jones, R., Dodd, C.E.R., (2005). A comparison of three modelling approaches for quantitative risk assessment using the case study of Salmonella spp. in poultry meat. Int. J. Food Microbiol. 98, 35–51. doi:10.1016/j.ijfoodmicro.2004.05.005

Pavione, D.M.S., Bastos, R.K.X., Bevilacqua, P.D., (2013). Quantitative microbial risk assessment applied to irrigation of salad crops with waste stabilization pond effluents. Water Sci. Technol. 67, 1208–15. doi:10.2166/wst.2013.674

Payment, P., Locas, A., (2011). Pathogens in water: value and limits of correlation with microbial indicators. Ground Water 49, 4–11. doi:10.1111/j.1745-6584.2010.00710.x

Petterson, S., (2001). Microbial risks from wastewater irrigation of salad crops: A screening-level risk assessment. WATER Environ. Res. 73, 667 – 672. doi:10.2307/25045556

Petterson, S.R., (2002). Of: Microbial risks from wastewater irrigation of salad crops: a screening-level risk assessment, S. R. Petterson; N. J. Ashbolt; A. Sharma, 73, 667 (2001). Water Environ. Res. 74.

Petterson, S.R., Signor, R.S., Ashbolt, N.J., (2007). Incorporating method recovery uncertainties in stochastic estimates of raw water protozoan concentrations for QMRA. Water Heal. 51–65. doi:10.2166/wh.2007.142

Qadir, M., Wichelns, D., Raschid-Sally, L., McCornick, P.G., Drechsel, P., Bahri, A., Minhas, P.S., (2010). The challenges of wastewater irrigation in developing countries. Agric. Water Manag. 97, 561–568. doi:10.1016/j.agwat.2008.11.004

Rabiner, L.R., (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77, 257–286. doi:10.1109/5.18626

Raschid-sally, L., Jayakody, P., (2008). Drivers and characteristics of wastewater

agriculture in developing countries: results from a global assessment. doi:http://dx.doi.org/10.3910/2009.127

Reid, D.., (2012). Statistical Distributions in Scientific Work: Volume 4 — Models, Structures, and Characterizations, Proceedings of the NATO Advanced Study Institute held at the Università degli Studi di Trieste, Trieste, Italy, July 10 – August 1, 1980. Springer Science & Business Media.

Reinoso, R., Bécares, E., (2008). Environmental inactivation of Cryptosporidium parvum oocysts in waste stabilization ponds. Microb. Ecol. 56, 585–92. doi:10.1007/s00248-008-9378-7

Ribeiro Duarte, A.S., Stockmarr, A., Nauta, M., (2013). Making sense of zeros: impact on human health risk estimates, in: Med-Vet-Net Association International Scientific Conference DTU, Lyngby, Denmark, 24-25 June 2013. p. 55.

Rice, J.A., (2001). Mathematical Statistics and Data Analysis, Thomson Higher Education. doi:10.2307/3619963

Rigaux, C., Ancelet, S., Carlin, F., Nguyen-thé, C., Albert, I., (2013). Inferring an augmented Bayesian network to confront a complex quantitative microbial risk assessment model with durability studies: application to Bacillus cereus on a courgette purée production chain. Risk Anal. 33, 877–92. doi:10.1111/j.1539-6924.2012.01888.x

Rooijen, V.A.N., Rooijen, D. Van, (2008). Analysing productive use of domestic water and wastewater for urban livelihoods of the poor – a study from 4–6.

Rose, J.B., (1997). Environmental ecology of Cryptosporidium and public health implications. Annu. Rev. Public Health 18, 135–61. doi:10.1146/annurev.publhealth.18.1.135

Salmona, S., (2014). Measurement and Measurement Uncertainty. IEEE Instrum. Meas. 39–46.

Sartwell, P.E., (1995). The distribution of incubation periods of infectious disease. 1949. Am. J. Epidemiol. 141, 386–94; discussion 385.

Scheierling, S., Bartone, C., Mara, D., Drechsel, P., (2010). Improving wastewater use in agriculture: an emerging priority.

Scheierling, S.M., Mara, D.D., (2010). Improving Wastewater Use in Agriculture An

Emerging Priority.

Seidu, R., Heistad, A., Amoah, P., Drechsel, P., Jenssen, P.D., Stenström, T.-A., (2008). Quantification of the health risk associated with wastewater reuse in Accra, Ghana: a contribution toward local guidelines. J. Water Health 6, 461–71. doi:10.2166/wh.2008.118

Sentz, K., (2002). Combination of Evidence in Dempster-Shafer Theory. Rep. No SAND2002 835, 96. doi:10.1.1.122.7929

Shuval, H., Lampert, Y., Fattal, B., (1997). Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. Water Sci. Technol. 35, 15–20.

Signor, R.S., Ashbolt, N.J., (2009). Comparing probabilistic microbial risk assessments for drinking water against daily rather than annualised infection probability targets. J. Water Health 7, 535–43. doi:10.2166/wh.2009.101

Silverman, A.I., Akrong, M.O., Amoah, P., Drechsel, P., Nelson, K.L., (2013). Quantification of human norovirus GII, human adenovirus, and fecal indicator organisms in wastewater used for irrigation in Accra, Ghana. J. Water Health 11, 473–88. doi:10.2166/wh.2013.025

Simmons, K., Gambhir, M., Leon, J., Lopman, B., (2013). Duration of immunity to norovirus gastroenteritis. Emerg. Infect. Dis. 19, 1260–1267. doi:10.3201/eid1908.130472

Singh, V., Setup, P., (2003). Principle of Maximum Entropy. Entropy Theory its Appl. Environ. … 85–92. doi:10.1007/s10955-006-9121-z

Smith, E.P., Ye, K., Hughes, C., Shabman, L., (2001). Statistical Assessment of Violations of Water Quality Standards under Section 303(d) of the Clean Water Act. Environ. Sci. Technol. 35, 606–612. doi:10.1021/es001159e

Sofia, A., Duarte, R., (2013). The interpretation of quantitative microbial data: meeting the demands of quantitative microbiological risk assessment.

Soize, C., Capiez-Lernout, E., Durand, J.-F., Fernandez, C., Gagliardini, L., (2008). Probabilistic model identification of uncertainties in computational models for dynamical systems and experimental validation. Comput. Methods Appl. Mech. Eng. 198, 150–163. doi:10.1016/j.cma.2008.04.007

Soller, J., (2008). Quantitative Microbial Risk Assessment Why Use Risk Assessment ?, in: An Introduction to Quantitative Microbial Risk Assessment. Washington D.C, pp. 1–17.

Spanos, A., (1999). Probability Theory and Statistical Inference: Econometric Modeling with Observational Data. Cambridge University Press.

Sparck Jones, K., Walker, S., Robertson, S.., (2000). A probabilistic model of information retrieval: development and comparative experiments. Inf. Process. Manag. 36, 809–840. doi:10.1016/S0306-4573(00)00016-9

State, A., (2010). Development of Guidelines for Wastewater Use in Agriculture 1918-2010 106–144.

Sukhrie, F.H. a, Siebenga, J.J., Beersma, M.F.C., Koopmans, M., (2010). Chronic shedders as reservoir for nosocomial transmission of norovirus. J. Clin. Microbiol. 48, 4303–4305. doi:10.1128/JCM.01308-10

Sukhrie, F.H. a, Teunis, P., Vennema, H., Copra, C., Thijs Beersma, M.F.C., Bogerman, J., Koopmans, M., (2012). Nosocomial transmission of norovirus is mainly caused by symptomatic cases. Clin. Infect. Dis. 54, 931–937. doi:10.1093/cid/cir971

Swart, a N., Tomasi, M., Kretzschmar, M., Havelaar, a H., Diekmann, O., (2012). The protective effects of temporary immunity under imposed infection pressure. Epidemics 4, 43–7. doi:10.1016/j.epidem.2011.12.002

Teunis, P., Chappell, C., Okhuysen, P., (2002). Cryptosporidium Dose-Response Studies: Variation Between Hosts. Risk Anal. 22.

Teunis, P.F., Havelaar, a H., (2000). The Beta Poisson dose-response model is not a single-hit model. Risk Anal. 20, 513–20.

Teunis, P.F., Nagelkerke, N.J., Haas, C.N., (1999). Dose response models for infectious gastroenteritis. Risk Anal. 19, 1251–60.

Teunis, P.F.M., Moe, C.L., Liu, P., Miller, S.E., Lindesmith, L., Baric, R.S., Pendu, J. Le, Calderon, R.L., (2008). Norwalk Virus : How Infectious is It ? 1476, 1468–1476. doi:10.1002/jmv

Teunis, P.F.M., Sukhrie, F.H. a., Vennema, H., Bogerman, J., Beersma, M.F.C., Koopmans, M.P.G., (2014). Shedding of norovirus in symptomatic and asymptomatic infections. Epidemiol. Infect. 1–8. doi:10.1017/S095026881400274X

Tiimub, B., Kuffour, R., Kwarteng, A., (2012). Bacterial Contamination Levels of Lettuce Irrigated with Waste Water in the Kumasi Metropolis. J. Biol. Agric. Healthc. 2, 116–127.

Travis, P.B., Goodman, K.J., O'Rourke, K.M., Groves, F.D., Sinha, D., Nicholas, J.S., VanDerslice, J., Lackland, D., Mena, K.D., (2010). Estimation of Norovirus risk. J. Water Health 8, 192–203. doi:10.2166/wh.2009.040

Tribble, D.R., Baqar, S., Scott, D. a, Oplinger, M.L., Trespalacios, F., Rollins, D., Walker, R.I., Clements, J.D., Walz, S., Gibbs, P., Burg, E.F., Moran, A.P., Applebee, L., Bourgeois, a L., (2010). Assessment of the duration of protection in Campylobacter jejuni experimental infection in humans. Infect. Immun. 78, 1750–9. doi:10.1128/IAI.01021-09

U S EPA, R., (2006). National primary drinking water regulation Long term 2 enhanced surface water treatment rule (LT2ESWTR); final rule. 40CFR Parts 9, 141 and 142,volume 71, Number 654,.

van der Voet, H., de Mul, A., van Klaveren, J.D., (2007). A probabilistic model for simultaneous exposure to multiple compounds from food and its use for risk-benefit assessment. Food Chem. Toxicol. 45, 1496–506. doi:10.1016/j.fct.2007.02.009

van der Voet, H., van der Heijden, G.W. a M., Bos, P.M.J., Bosgra, S., Boon, P.E., Muri, S.D., Brüschweiler, B.J., (2009). A model for probabilistic health impact assessment of exposure to food chemicals. Food Chem. Toxicol. 47, 2926–40. doi:10.1016/j.fct.2008.12.027

van Lieverloo, J.H.M., Blokker, E.J.M., Medema, G., (2007). Quantitative microbial risk assessment of distributed drinking water using faecal indicator incidence and concentrations. J. Water Health 5 Suppl 1, 131–49. doi:10.2166/wh.2007.134

Vega, E., Barclay, L., Gregoricus, N., Williams, K., Lee, D., Vinjé, J., (2011). Novel surveillance network for norovirus gastroenteritis outbreaks, United States. Emerg. Infect. Dis. 17, 1389–1395. doi:10.3201/eid1708.101837

Verdonck, F., (2001). Probabilistic environmental risk assessment. Meded. Rijksuniv. Gent. Fak. Landbouwkd. Toegep. Biol. Wet. 66, 13 – 19.

Wackerly, D., Mendenhall, W., Scheaffer, R., 2007. Mathematical Statistics with

Applications. Cengage Learning.

Ward, C., (1993). ENVIRONMENTAL RISK ASSESSMENT. Bank. LAW J. 110, 204 – 215.

White, K.E., Osterholm, M.T., Mariotti, J.A., Korlath, J.A., Lawrence, D.H., Ristinen, T.L., Greenberg, H.B., (1986). A foodborne outbreak of Norwalk virus gastroenteritis. Evidence for post-recovery transmission. Am. J. Epidemiol. 124, 120–6.

WHO, (2015). WHO | Metrics: Disability-Adjusted Life Year (DALY) [WWW Document]. World Heal. Organ. URL http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/ (accessed 7.15.15).

WHO, (2006). WHO Guideline for the Safe Use of Wastewater, Excreta and Greywater, Safe Use of Wastewater in Agriculture. Geneva, Swittzerland.

Wierman, M.J., (2001). Measuring Granularity in Evidence Theory. Int. J. Gen. Syst. 30, 649–660. doi:10.1080/03081070108960739

William, H. Press; Flannery, S.A; Teukolsky, S.A; Vetterling, W.., (1989). Numerical Recipes in Pascal (First Edition) | Numerical Recipes, Cambridge University Press.

Williams, T., (1965). The Distribution of Response Times in a Birth-Death Process on JSTOR. Biometrika 52, 581–585.

World Wildlife Fund Report, W., (2015). International Decade for Action "Water for Life" 2005-2015. Focus Areas: Water scarcity [WWW Document]. worldwildlife.org. URL http://www.un.org/waterforlifedecade/scarcity.shtml (accessed 6.7.15).

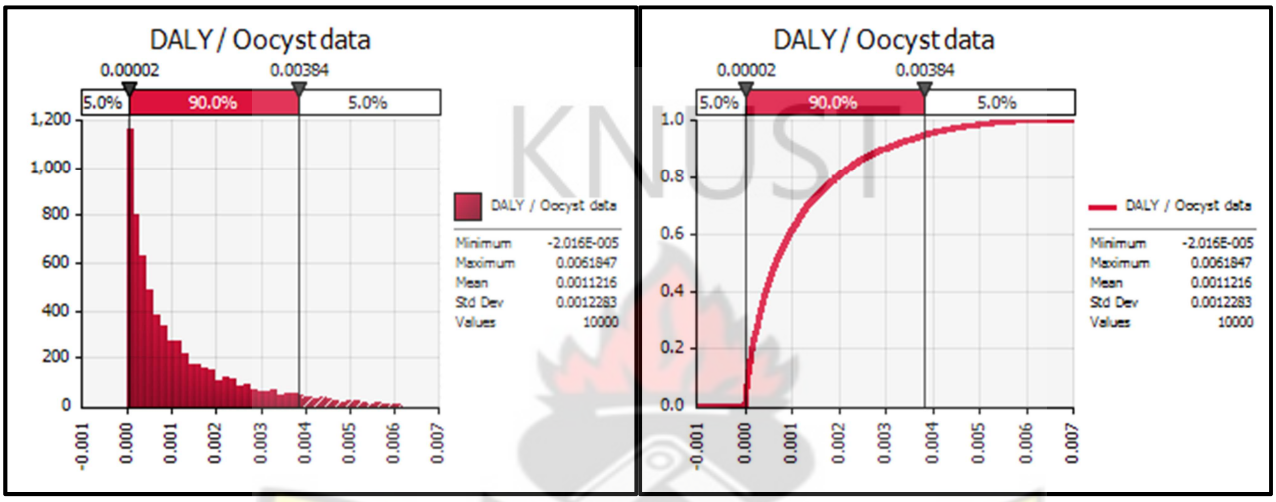
Yusoff, B., Taib, I., Abdullah, L., Wahab, A.F., (2011). A New Similarity Measure on Intuitionistic Fuzzy Sets 23–27.

Zadeh, L., Dubois, D., Prade, H., (1978). Possibility theory Formalization of possibility 2–5.

APPENDICES

APPENDIX A

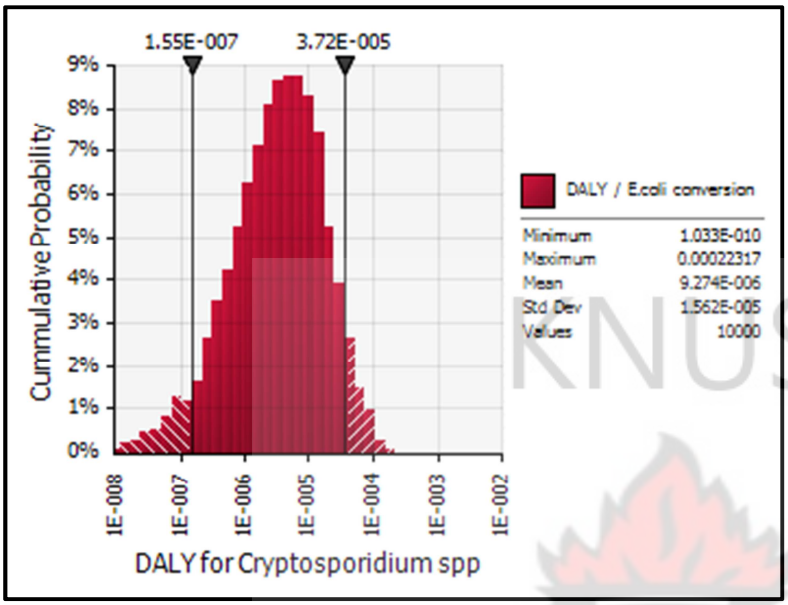
A1 CRYPTOSPORIDIUM OOCYST DATA RISK ASSESSMENT

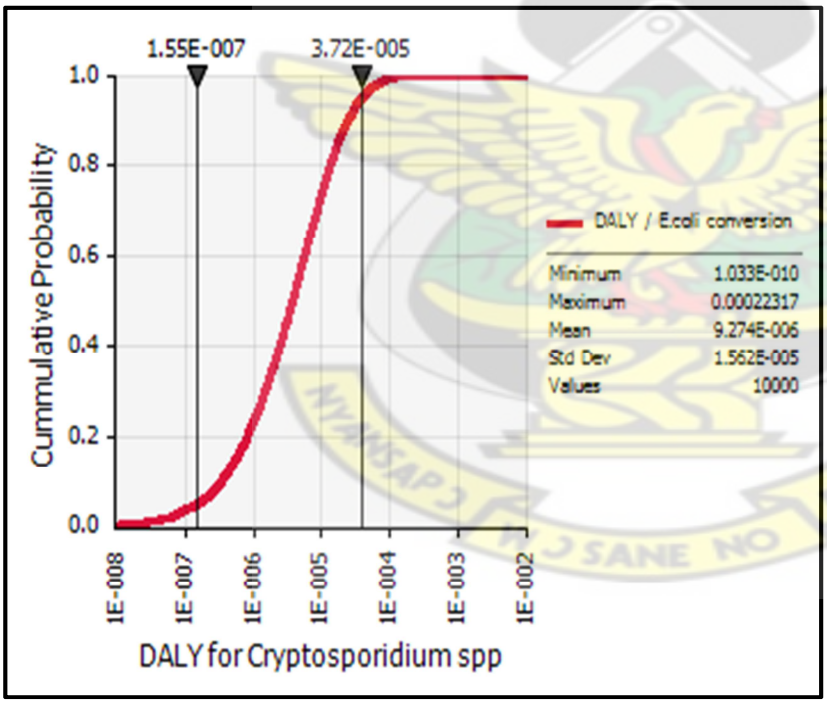


Summary Statistics for DALY / Oocyst data			
Statistics		**Percentile**	
Minimum	-2.01643E-05	**5%**	2.386E-05
Maximum	0.006184724	**10%**	6.123E-05
Mean	0.001121648	**15%**	0.0001058
Std Dev	0.001228328	**20%**	0.0001538
Variance	1.50879E-06	**25%**	0.0002175
Skewness	1.55299248	**30%**	0.000284
Kurtosis	5.003454498	**35%**	0.0003565
Median	0.000655246	**40%**	0.0004397
Mode	1.55852E-05	**45%**	0.0005408
Left X	2.38552E-05	**50%**	0.0006552
Left P	5%	**55%**	0.0007946
Right X	0.003842637	**60%**	0.0009426
Right P	95%	**65%**	0.0011277
Diff X	0.003818782	**70%**	0.00134
Diff P	90%	**75%**	0.0016273
#Errors	0	**80%**	0.0019405
Filter Min	Off	**85%**	0.0023624
Filter Max	Off	**90%**	0.0029489
#Filtered	0	**95%**	0.0038426

Change in Output Statistic for DALY / Oocyst data			
Rank	**Name**	**Lower**	**Upper**
1	PRODUCE DATA	0.0001036	0.0022764
2	Disease Burden	0.0001521	0.0020863
3	Virus reduction / Water real data	0.0003904	0.0021002
4	volume of irrigation water	0.0009586	0.0013629
5	Daily consumption	0.0009059	0.0013071
6	Frequency of consumption	0.0009322	0.0012902
7	Days for withholding water	0.0010679	0.0011718

A2 CRYPTO E.COLI CONVERSION





Summary Statistics for DALY / E.coli conversion			
Statistics		**Percentile**	
Minimum	1.03267E-10	**5%**	1.546E-07
Maximum	0.000223168	**10%**	3.411E-07
Mean	9.27416E-06	**15%**	5.56E-07
Std Dev	1.56191E-05	**20%**	8.103E-07
Variance	2.43956E-10	**25%**	1.12E-06
Skewness	4.041970518	**30%**	1.468E-06
Kurtosis	26.84844898	**35%**	1.894E-06
Median	3.70901E-06	**40%**	2.41E-06
Mode	7.36907E-08	**45%**	2.995E-06
Left X	1.54578E-07	**50%**	3.709E-06
Left P	5%	**55%**	4.515E-06
Right X	3.71838E-05	**60%**	5.547E-06
Right P	95%	**65%**	6.784E-06
Diff X	3.70292E-05	**70%**	8.448E-06
Diff P	90%	**75%**	1.044E-05
#Errors	0	**80%**	1.318E-05
Filter Min	Off	**85%**	1.703E-05
Filter Max	Off	**90%**	2.399E-05
#Filtered	0	**95%**	3.718E-05

Change in Output Statistic for DALY / E.coli conversion			
Rank	**Name**	**Lower**	**Upper**
1	Virus reduction / Water real data	1.678E-06	2.787E-05
2	Time for withholding irrigation	8.475E-07	1.801E-05
3	Disease Burden	1.185E-06	1.652E-05
4	volume of irrigation water (K6)	2.488E-06	1.534E-05
5	Daily consumption	6.326E-06	1.247E-05
6	volume of irrigation water (D6)	6.7E-06	1.188E-05
7	Frequency of consumption	6.836E-06	1.133E-05
8	Days for withholding water	7.839E-06	1.049E-05

APPENDIX B

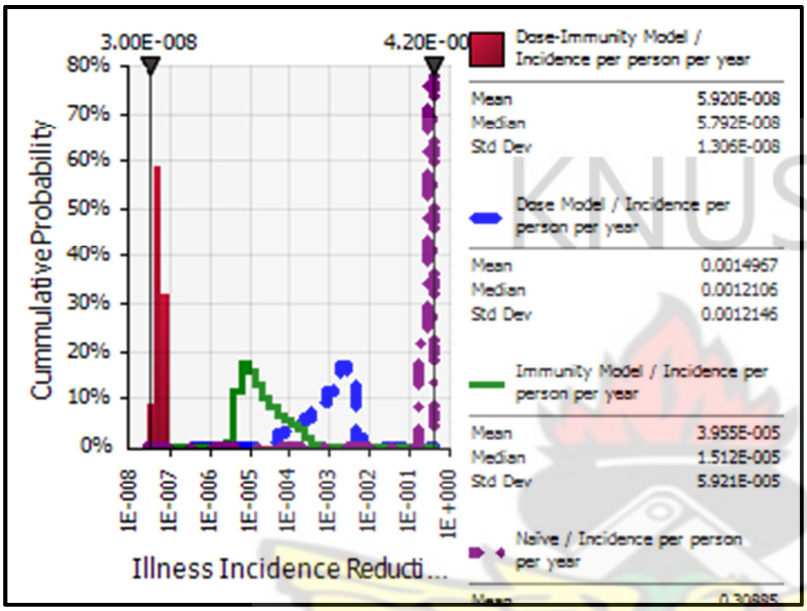
ILLNESS INCIDENCE PER POPULATION PER YEAR

Incidence in the population per year					
Population Percentage	**Population Number**	**Naïve Model**	**Immunity Model**	**Dose Model**	**Dose-immunity Model**
0	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
5	1.25E+06	3.86E+05	1.81E+01	1.54E+03	7.21E-02
10	2.50E+06	7.72E+05	3.63E+01	3.07E+03	1.44E-01
15	3.75E+06	1.16E+06	5.44E+01	4.61E+03	2.16E-01
20	5.00E+06	1.54E+06	7.25E+01	6.14E+03	2.88E-01
25	6.25E+06	1.93E+06	9.07E+01	7.68E+03	3.61E-01
30	7.50E+06	2.32E+06	1.09E+02	9.21E+03	4.33E-01
35	8.75E+06	2.70E+06	1.27E+02	1.07E+04	5.05E-01
40	1.00E+07	3.09E+06	1.45E+02	1.23E+04	5.77E-01
45	1.13E+07	3.47E+06	1.63E+02	1.38E+04	6.49E-01
50	1.25E+07	3.86E+06	1.81E+02	1.54E+04	7.21E-01
55	1.38E+07	4.25E+06	1.99E+02	1.69E+04	7.93E-01
60	1.50E+07	4.63E+06	2.18E+02	1.84E+04	8.65E-01
65	1.63E+07	5.02E+06	2.36E+02	2.00E+04	9.37E-01
70	1.75E+07	5.40E+06	2.54E+02	2.15E+04	1.01E+00
75	1.88E+07	5.79E+06	2.72E+02	2.30E+04	1.08E+00
80	2.00E+07	6.18E+06	2.90E+02	2.46E+04	1.15E+00
85	2.13E+07	6.56E+06	3.08E+02	2.61E+04	1.23E+00
90	2.25E+07	6.95E+06	3.26E+02	2.76E+04	1.30E+00
95	2.38E+07	7.34E+06	3.45E+02	2.92E+04	1.37E+00
100	2.50E+07	7.72E+06	3.63E+02	3.07E+04	1.44E+00

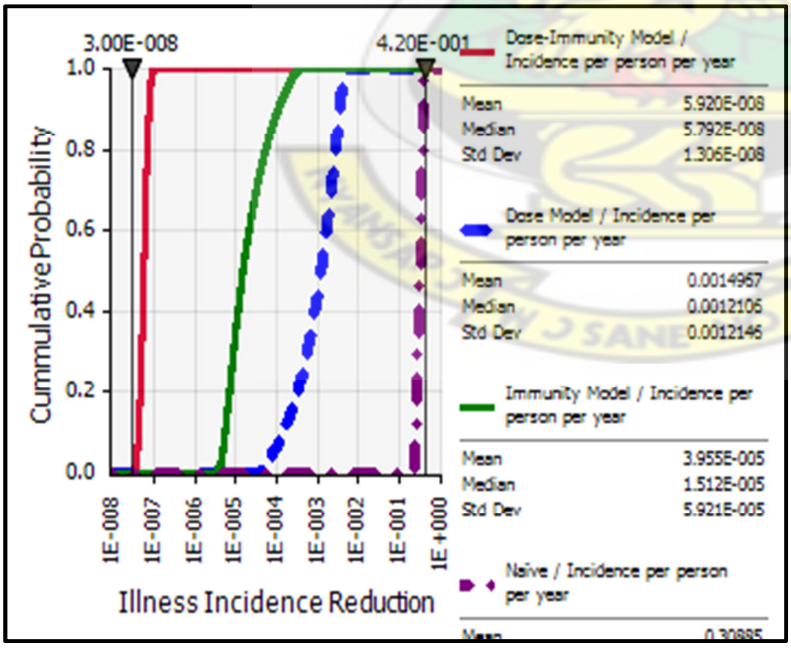
APPENDIX B1 ILLNESS INCIDENCE PER PERSON PER YEAR

B1A. (SYMPTOMATIC INFECTIOUNESS)

PROBABILITY GRAPHS FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



Summary Statistics for Symptomatic Infectiousness Individual

Dose-Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.88E-08	**5%**	4.02E-08
Maximum	1.02E-07	**10%**	4.30E-08
Mean	5.92E-08	**15%**	4.55E-08
Std Dev	1.30E-08	**20%**	4.75E-08
Variance	1.69736E-16	**25%**	4.94E-08
Skewness	0.471635968	**30%**	5.13E-08
Kurtosis	2.832924093	**35%**	5.31E-08
Median	5.81E-08	**40%**	5.47E-08
Mode	5.99E-08	**45%**	5.64E-08
Left X	4.02E-08	**50%**	5.81E-08
Left P	5%	**55%**	5.97E-08
Right X	8.34E-08	**60%**	6.13E-08
Right P	95%	**65%**	6.31E-08
Diff X	4.32E-08	**70%**	6.50E-08
Diff P	90%	**75%**	6.75E-08
#Errors	0	**80%**	7.00E-08
Filter Min	Off	**85%**	7.31E-08
Filter Max	Off	**90%**	7.72E-08
#Filtered	0	**95%**	8.34E-08

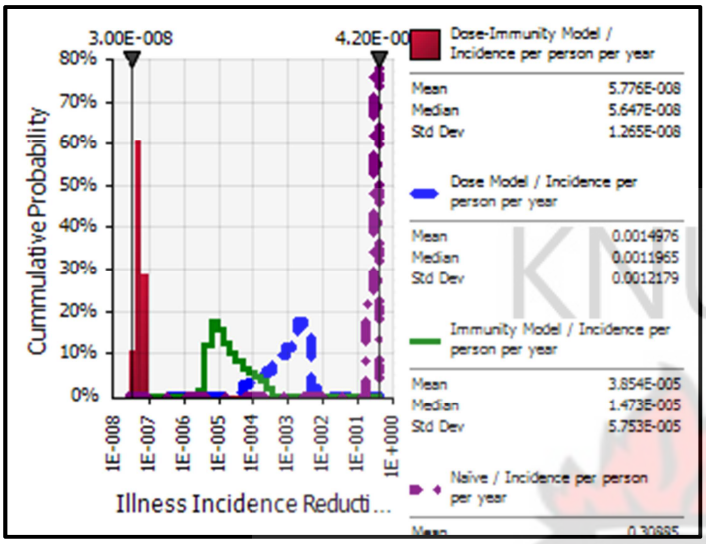
Dose Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.14E-05	**5%**	8.52E-05
Maximum	5.07E-03	**10%**	1.44E-04
Mean	1.50E-03	**15%**	2.22E-04
Std Dev	1.22E-03	**20%**	3.14E-04
Variance	1.48551E-06	**25%**	4.21E-04
Skewness	0.69866711	**30%**	5.41E-04
Kurtosis	2.466732266	**35%**	6.83E-04
Median	1.21E-03	**40%**	8.46E-04
Mode	6.06E-05	**45%**	1.03E-03
Left X	8.52E-05	**50%**	1.21E-03
Left P	5%	**55%**	1.40E-03
Right X	3.85E-03	**60%**	1.63E-03
Right P	95%	**65%**	1.85E-03
Diff X	3.76E-03	**70%**	2.11E-03
Diff P	90%	**75%**	2.37E-03
#Errors	0	**80%**	2.66E-03
Filter Min	Off	**85%**	2.99E-03
Filter Max	Off	**90%**	3.34E-03
#Filtered	0	**95%**	3.85E-03

Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.02E-06	**5%**	4.82E-06
Maximum	4.28E-04	**10%**	5.61E-06
Mean	3.96E-05	**15%**	6.32E-06
Std Dev	5.92E-05	**20%**	7.00E-06
Variance	3.50246E-09	**25%**	7.87E-06
Skewness	2.809171132	**30%**	8.83E-06
Kurtosis	11.6465013	**35%**	9.99E-06
Median	1.51E-05	**40%**	1.13E-05
Mode	6.66E-06	**45%**	1.31E-05
Left X	4.82E-06	**50%**	1.51E-05
Left P	5%	**55%**	1.78E-05
Right X	1.77E-04	**60%**	2.13E-05
Right P	95%	**65%**	2.60E-05
Diff X	1.72E-04	**70%**	3.21E-05
Diff P	90%	**75%**	4.07E-05
#Errors	0	**80%**	5.34E-05
Filter Min	Off	**85%**	7.23E-05
Filter Max	Off	**90%**	1.11E-04
#Filtered	0	**95%**	1.77E-04

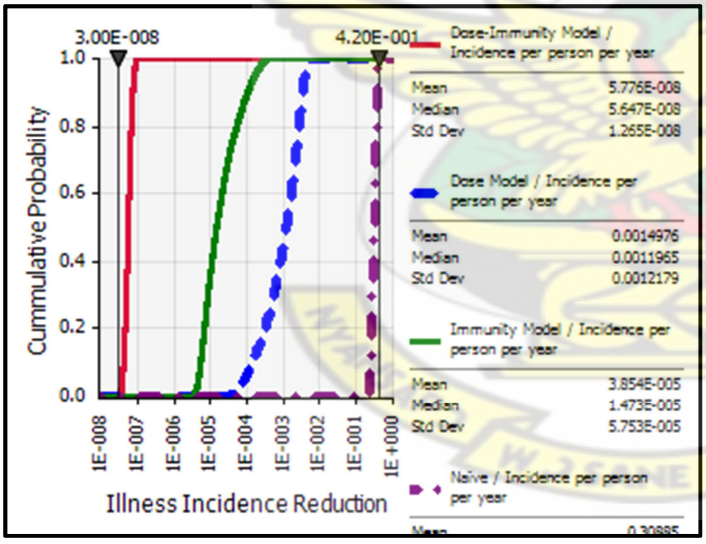
Naive/ Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.24E-01	**5%**	2.33E-01
Maximum	3.93E-01	**10%**	2.41E-01
Mean	3.09E-01	**15%**	2.50E-01
Std Dev	4.89E-02	**20%**	2.58E-01
Variance	0.002387251	**25%**	2.67E-01
Skewness	-2.71097E-06	**30%**	2.75E-01
Kurtosis	1.799996474	**35%**	2.83E-01
Median	3.09E-01	**40%**	2.92E-01
Mode	2.91E-01	**45%**	3.00E-01
Left X	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	3.17E-01
Right X	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	3.34E-01
Diff X	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	3.51E-01
#Errors	0	**80%**	3.60E-01
Filter Min	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	3.77E-01
#Filtered	0	**95%**	3.85E-01

B1B. (PRE-SYMPTOMATIC AND POST SYMPTOMATIC INFECTIOUNESS LOW)

PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAP FOR DR MODELS



Summary Statistics for Pre-Symptomatic and Post-Symptomatic Low

Dose-Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.81E-08	**5%**	3.92E-08
Maximum	1.00E-07	**10%**	4.20E-08
Mean	5.78E-08	**15%**	4.43E-08
Std Dev	1.28E-08	**20%**	4.65E-08
Variance	1.63933E-16	**25%**	4.82E-08
Skewness	0.483785996	**30%**	5.00E-08
Kurtosis	2.839293047	**35%**	5.16E-08
Median	5.65E-08	**40%**	5.32E-08
Mode	5.57E-08	**45%**	5.49E-08
Left X	3.92E-08	**50%**	5.65E-08
Left P	5%	**55%**	5.82E-08
Right X	8.12E-08	**60%**	5.98E-08
Right P	95%	**65%**	6.16E-08
Diff X	4.20E-08	**70%**	6.37E-08
Diff P	90%	**75%**	6.59E-08
#Errors	0	**80%**	6.86E-08
Filter Min	Off	**85%**	7.18E-08
Filter Max	Off	**90%**	7.57E-08
#Filtered	0	**95%**	8.12E-08

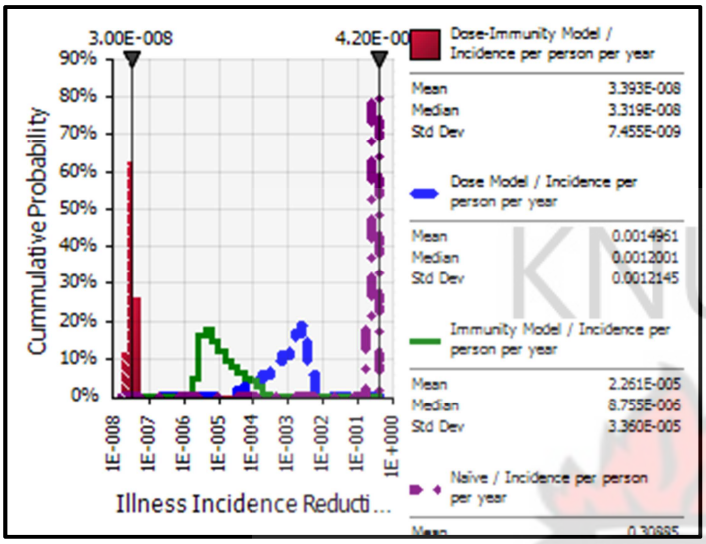
Dose Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.10E-05	**5%**	8.50E-05
Maximum	5.08E-03	**10%**	1.44E-04
Mean	1.50E-03	**15%**	2.23E-04
Std Dev	1.21E-03	**20%**	3.11E-04
Variance	1.46886E-06	**25%**	4.21E-04
Skewness	0.676579428	**30%**	5.47E-04
Kurtosis	2.42652171	**35%**	6.85E-04
Median	1.20E-03	**40%**	8.45E-04
Mode	5.10E-05	**45%**	1.02E-03
Left X	8.50E-05	**50%**	1.20E-03
Left P	5%	**55%**	1.40E-03
Right X	3.77E-03	**60%**	1.63E-03
Right P	95%	**65%**	1.87E-03
Diff X	3.68E-03	**70%**	2.12E-03
Diff P	90%	**75%**	2.38E-03
#Errors	0	**80%**	2.67E-03
Filter Min	Off	**85%**	2.99E-03
Filter Max	Off	**90%**	3.32E-03
#Filtered	0	**95%**	3.77E-03

Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.83E-06	**5%**	4.69E-06
Maximum	4.33E-04	**10%**	5.46E-06
Mean	3.85E-05	**15%**	6.17E-06
Std Dev	5.71E-05	**20%**	6.88E-06
Variance	3.2615E-09	**25%**	7.68E-06
Skewness	2.815156623	**30%**	8.62E-06
Kurtosis	11.81161483	**35%**	9.67E-06
Median	1.48E-05	**40%**	1.11E-05
Mode	6.10E-06	**45%**	1.28E-05
Left X	4.69E-06	**50%**	1.48E-05
Left P	5%	**55%**	1.75E-05
Right X	1.69E-04	**60%**	2.08E-05
Right P	95%	**65%**	2.52E-05
Diff X	1.64E-04	**70%**	3.12E-05
Diff P	90%	**75%**	4.01E-05
#Errors	0	**80%**	5.22E-05
Filter Min	Off	**85%**	7.38E-05
Filter Max	Off	**90%**	1.05E-04
#Filtered	0	**95%**	1.69E-04

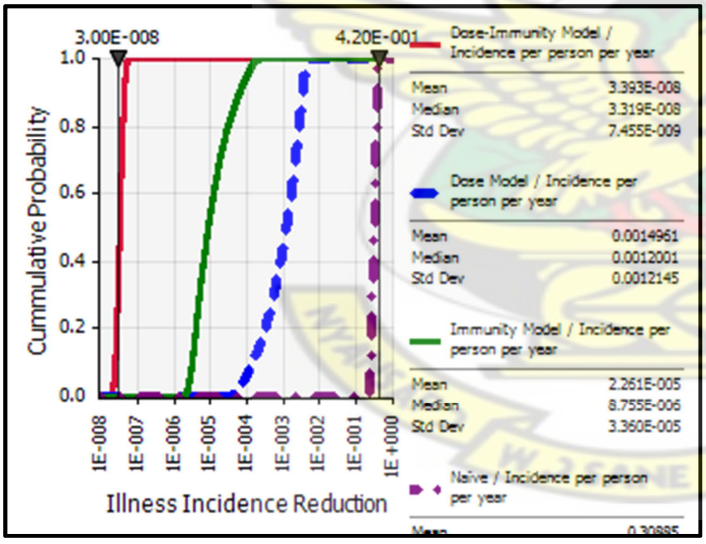
Naive/ Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.24E-01	**5%**	2.33E-01
Maximum	3.93E-01	**10%**	2.41E-01
Mean	3.09E-01	**15%**	2.50E-01
Std Dev	4.89E-02	**20%**	2.58E-01
Variance	0.002387257	**25%**	2.67E-01
Skewness	-7.56816E-07	**30%**	2.75E-01
Kurtosis	1.799994143	**35%**	2.83E-01
Median	3.09E-01	**40%**	2.92E-01
Mode	3.72E-01	**45%**	3.00E-01
Left X	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	3.17E-01
Right X	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	3.34E-01
Diff X	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	3.51E-01
#Errors	0	**80%**	3.60E-01
Filter Min	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	3.77E-01
#Filtered	0	**95%**	3.85E-01

B1C. (PRE-SYMPTOMATIC AND POST SYMPTOMATIC INFECTIOUNESS HIGH)

PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



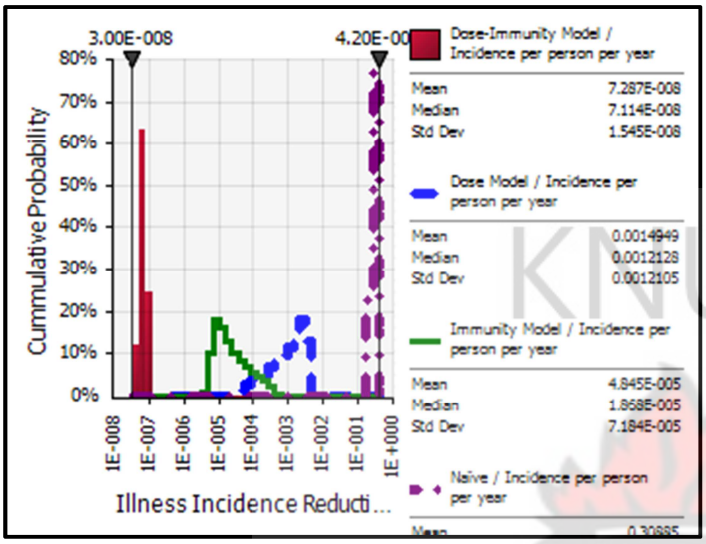
Summary Statistics for Pre-Symptomatic and Post-Symptomatic High

Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	1.67E-08	**5%**	2.32E-08	**Minimum**	3.22E-05	**5%**	8.46E-05
Maximum	5.88E-08	**10%**	2.48E-08	**Maximum**	5.03E-03	**10%**	1.45E-04
Mean	3.39E-08	**15%**	2.62E-08	**Mean**	1.50E-03	**15%**	2.21E-04
Std Dev	7.44E-09	**20%**	2.73E-08	**Std Dev**	1.21E-03	**20%**	3.12E-04
Variance	5.53724E-17	**25%**	2.84E-08	**Variance**	1.47456E-06	**25%**	4.21E-04
Skewness	0.52151757	**30%**	2.93E-08	**Skewness**	0.69355851	**30%**	5.51E-04
Kurtosis	2.874078836	**35%**	3.03E-08	**Kurtosis**	2.468769613	**35%**	6.93E-04
Median	3.31E-08	**40%**	3.13E-08	**Median**	1.21E-03	**40%**	8.45E-04
Mode	2.87E-08	**45%**	3.22E-08	**Mode**	5.10E-05	**45%**	1.02E-03
Left X	2.32E-08	**50%**	3.31E-08	**Left X**	8.46E-05	**50%**	1.21E-03
Left P	5%	**55%**	3.41E-08	**Left P**	5%	**55%**	1.41E-03
Right X	4.77E-08	**60%**	3.50E-08	**Right X**	3.80E-03	**60%**	1.62E-03
Right P	95%	**65%**	3.60E-08	**Right P**	95%	**65%**	1.85E-03
Diff X	2.45E-08	**70%**	3.72E-08	**Diff X**	3.72E-03	**70%**	2.11E-03
Diff P	90%	**75%**	3.86E-08	**Diff P**	90%	**75%**	2.38E-03
#Errors	0	**80%**	4.02E-08	**#Errors**	0	**80%**	2.67E-03
Filter Min	Off	**85%**	4.20E-08	**Filter Min**	Off	**85%**	2.96E-03
Filter Max	Off	**90%**	4.43E-08	**Filter Max**	Off	**90%**	3.33E-03
#Filtered	0	**95%**	4.77E-08	**#Filtered**	0	**95%**	3.80E-03

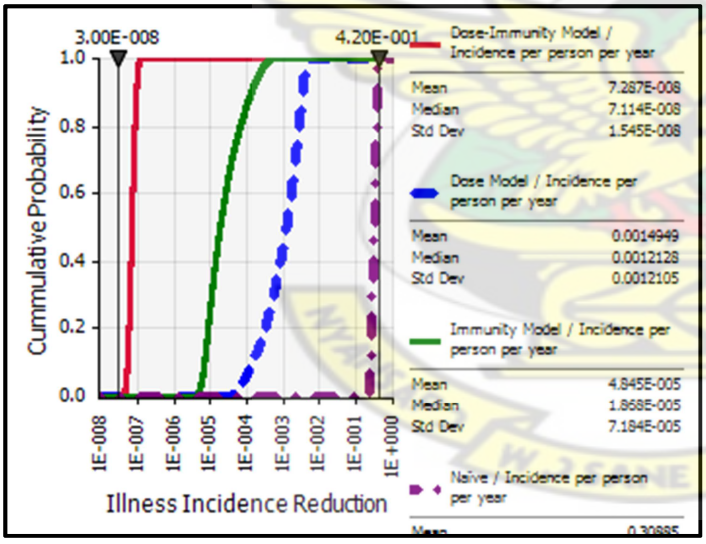
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	1.74E-06	**5%**	2.76E-06	**Minimum**	2.24E-01	**5%**	2.33E-01
Maximum	2.40E-04	**10%**	3.16E-06	**Maximum**	3.93E-01	**10%**	2.41E-01
Mean	2.27E-05	**15%**	3.58E-06	**Mean**	3.09E-01	**15%**	2.50E-01
Std Dev	3.39E-05	**20%**	4.02E-06	**Std Dev**	4.89E-02	**20%**	2.58E-01
Variance	1.14885E-09	**25%**	4.54E-06	**Variance**	0.00238726	**25%**	2.67E-01
Skewness	2.814525713	**30%**	5.10E-06	**Skewness**	-1.16214E-06	**30%**	2.75E-01
Kurtosis	11.62342123	**35%**	5.76E-06	**Kurtosis**	1.799999732	**35%**	2.83E-01
Median	8.79E-06	**40%**	6.54E-06	**Median**	3.09E-01	**40%**	2.92E-01
Mode	3.79E-06	**45%**	7.51E-06	**Mode**	3.60E-01	**45%**	3.00E-01
Left X	2.76E-06	**50%**	8.79E-06	**Left X**	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	1.03E-05	**Left P**	5%	**55%**	3.17E-01
Right X	1.00E-04	**60%**	1.22E-05	**Right X**	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	1.48E-05	**Right P**	95%	**65%**	3.34E-01
Diff X	9.76E-05	**70%**	1.82E-05	**Diff X**	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	2.32E-05	**Diff P**	90%	**75%**	3.51E-01
#Errors	0	**80%**	3.09E-05	**#Errors**	0	**80%**	3.60E-01
Filter Min	Off	**85%**	4.24E-05	**Filter Min**	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	6.21E-05	**Filter Max**	Off	**90%**	3.77E-01
#Filtered	0	**95%**	1.00E-04	**#Filtered**	0	**95%**	3.85E-01

B1D. (INNATE GENETIC RESISTANCE)

PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



Summary Statistics for Innate Genetic Resistance

Dose-Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.57E-08	**5%**	5.05E-08
Maximum	1.23E-07	**10%**	5.41E-08
Mean	7.28E-08	**15%**	5.69E-08
Std Dev	1.53E-08	**20%**	5.92E-08
Variance	2.34802E-16	**25%**	6.15E-08
Skewness	0.487413486	**30%**	6.36E-08
Kurtosis	2.890478312	**35%**	6.56E-08
Median	7.13E-08	**40%**	6.74E-08
Mode	7.33E-08	**45%**	6.94E-08
Left X	5.05E-08	**50%**	7.13E-08
Left P	5%	**55%**	7.32E-08
Right X	1.01E-07	**60%**	7.52E-08
Right P	95%	**65%**	7.74E-08
Diff X	5.07E-08	**70%**	7.98E-08
Diff P	90%	**75%**	8.24E-08
#Errors	0	**80%**	8.55E-08
Filter Min	Off	**85%**	8.95E-08
Filter Max	Off	**90%**	9.42E-08
#Filtered	0	**95%**	1.01E-07

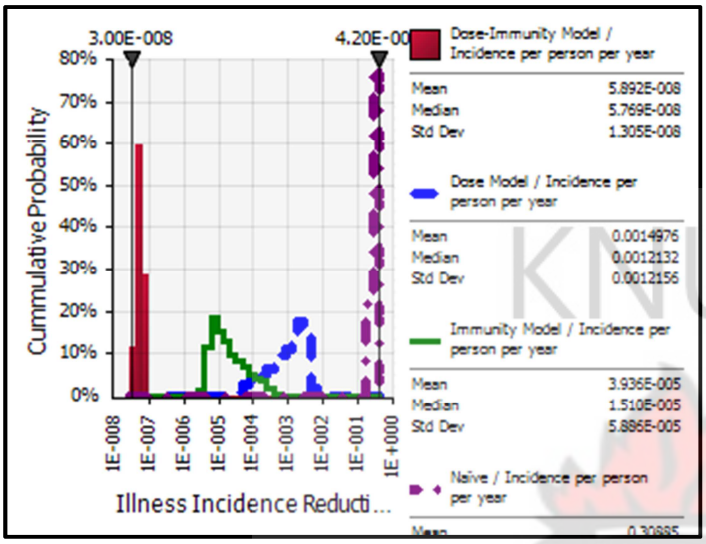
Dose Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.13E-05	**5%**	8.52E-05
Maximum	5.12E-03	**10%**	1.45E-04
Mean	1.49E-03	**15%**	2.22E-04
Std Dev	1.21E-03	**20%**	3.14E-04
Variance	1.45973E-06	**25%**	4.23E-04
Skewness	0.677393831	**30%**	5.48E-04
Kurtosis	2.434814128	**35%**	6.88E-04
Median	1.20E-03	**40%**	8.43E-04
Mode	5.18E-05	**45%**	1.02E-03
Left X	8.52E-05	**50%**	1.20E-03
Left P	5%	**55%**	1.41E-03
Right X	3.78E-03	**60%**	1.64E-03
Right P	95%	**65%**	1.85E-03
Diff X	3.69E-03	**70%**	2.12E-03
Diff P	90%	**75%**	2.38E-03
#Errors	0	**80%**	2.67E-03
Filter Min	Off	**85%**	2.96E-03
Filter Max	Off	**90%**	3.29E-03
#Filtered	0	**95%**	3.78E-03

Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.83E-06	**5%**	6.08E-06
Maximum	5.68E-04	**10%**	6.97E-06
Mean	4.87E-05	**15%**	7.80E-06
Std Dev	7.28E-05	**20%**	8.66E-06
Variance	5.30392E-09	**25%**	9.69E-06
Skewness	2.863963764	**30%**	1.09E-05
Kurtosis	12.17038231	**35%**	1.23E-05
Median	1.87E-05	**40%**	1.40E-05
Mode	8.23E-06	**45%**	1.61E-05
Left X	6.08E-06	**50%**	1.87E-05
Left P	5%	**55%**	2.17E-05
Right X	2.11E-04	**60%**	2.63E-05
Right P	95%	**65%**	3.16E-05
Diff X	2.05E-04	**70%**	3.95E-05
Diff P	90%	**75%**	5.01E-05
#Errors	0	**80%**	6.53E-05
Filter Min	Off	**85%**	9.08E-05
Filter Max	Off	**90%**	1.35E-04
#Filtered	0	**95%**	2.11E-04

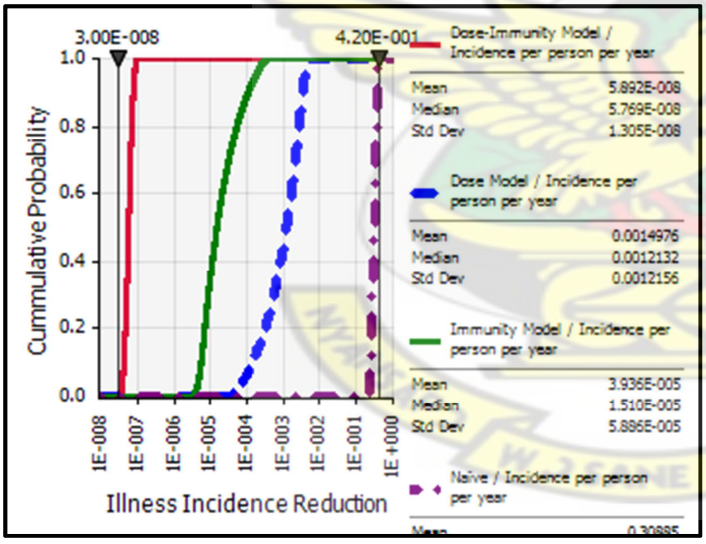
Naive/ Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.24E-01	**5%**	2.33E-01
Maximum	3.93E-01	**10%**	2.41E-01
Mean	3.09E-01	**15%**	2.50E-01
Std Dev	4.89E-02	**20%**	2.58E-01
Variance	0.002387262	**25%**	2.67E-01
Skewness	1.13629E-06	**30%**	2.75E-01
Kurtosis	1.799999743	**35%**	2.83E-01
Median	3.09E-01	**40%**	2.92E-01
Mode	3.67E-01	**45%**	3.00E-01
Left X	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	3.17E-01
Right X	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	3.34E-01
Diff X	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	3.51E-01
#Errors	0	**80%**	3.60E-01
Filter Min	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	3.77E-01
#Filtered	0	**95%**	3.85E-01

B1E. (NO-IMMUNE BOOSTING AFTER ASYMPTOMATIC INFECTIOUNESS)

PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



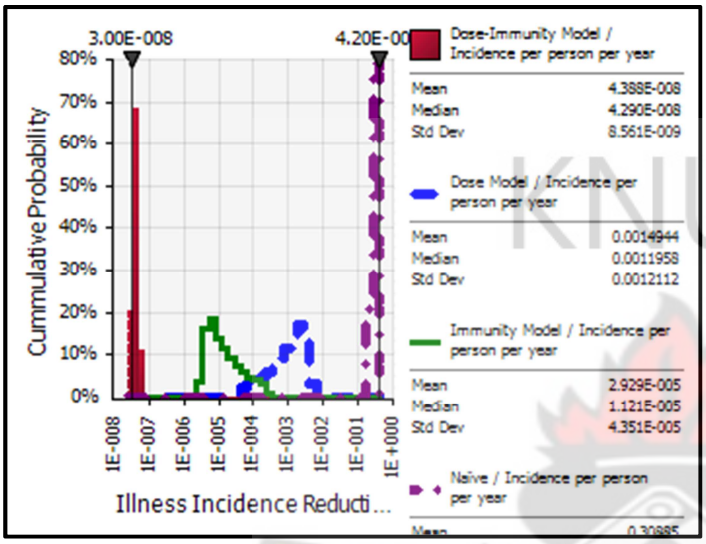
Summary Statistics for No-Immune Boosting after Asymptomatic Infectiousness

Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	2.86E-08	**5%**	4.00E-08	**Minimum**	3.08E-05	**5%**	8.35E-05
Maximum	1.03E-07	**10%**	4.30E-08	**Maximum**	5.10E-03	**10%**	1.44E-04
Mean	5.89E-08	**15%**	4.52E-08	**Mean**	1.50E-03	**15%**	2.21E-04
Std Dev	1.31E-08	**20%**	4.73E-08	**Std Dev**	1.22E-03	**20%**	3.14E-04
Variance	1.7088E-16	**25%**	4.92E-08	**Variance**	1.4928E-06	**25%**	4.22E-04
Skewness	0.511478568	**30%**	5.10E-08	**Skewness**	0.702327797	**30%**	5.51E-04
Kurtosis	2.852008556	**35%**	5.28E-08	**Kurtosis**	2.488539216	**35%**	6.86E-04
Median	5.74E-08	**40%**	5.44E-08	**Median**	1.21E-03	**40%**	8.45E-04
Mode	5.88E-08	**45%**	5.58E-08	**Mode**	5.93E-05	**45%**	1.01E-03
Left X	4.00E-08	**50%**	5.74E-08	**Left X**	8.35E-05	**50%**	1.21E-03
Left P	5%	**55%**	5.89E-08	**Left P**	5%	**55%**	1.41E-03
Right X	8.32E-08	**60%**	6.06E-08	**Right X**	3.81E-03	**60%**	1.62E-03
Right P	95%	**65%**	6.26E-08	**Right P**	95%	**65%**	1.86E-03
Diff X	4.32E-08	**70%**	6.47E-08	**Diff X**	3.73E-03	**70%**	2.11E-03
Diff P	90%	**75%**	6.72E-08	**Diff P**	90%	**75%**	2.39E-03
#Errors	0	**80%**	6.99E-08	**#Errors**	0	**80%**	2.66E-03
Filter Min	Off	**85%**	7.32E-08	**Filter Min**	Off	**85%**	2.98E-03
Filter Max	Off	**90%**	7.73E-08	**Filter Max**	Off	**90%**	3.35E-03
#Filtered	0	**95%**	8.32E-08	**#Filtered**	0	**95%**	3.81E-03

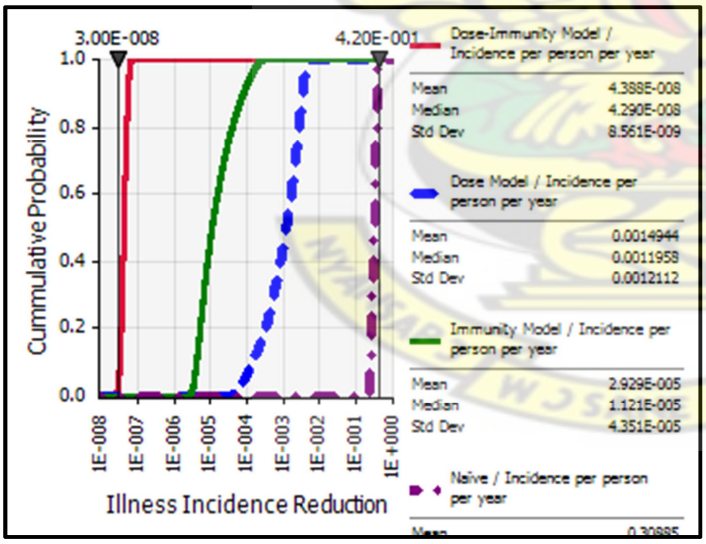
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	2.91E-06	**5%**	4.77E-06	**Minimum**	2.24E-01	**5%**	2.33E-01
Maximum	4.37E-04	**10%**	5.56E-06	**Maximum**	3.93E-01	**10%**	2.41E-01
Mean	3.93E-05	**15%**	6.25E-06	**Mean**	3.09E-01	**15%**	2.50E-01
Std Dev	5.84E-05	**20%**	6.99E-06	**Std Dev**	4.89E-02	**20%**	2.58E-01
Variance	3.41367E-09	**25%**	7.82E-06	**Variance**	0.002387261	**25%**	2.67E-01
Skewness	2.823672209	**30%**	8.79E-06	**Skewness**	2.71019E-06	**30%**	2.75E-01
Kurtosis	11.84958146	**35%**	9.99E-06	**Kurtosis**	1.800002432	**35%**	2.83E-01
Median	1.51E-05	**40%**	1.14E-05	**Median**	3.09E-01	**40%**	2.92E-01
Mode	6.62E-06	**45%**	1.31E-05	**Mode**	3.35E-01	**45%**	3.00E-01
Left X	4.77E-06	**50%**	1.51E-05	**Left X**	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	1.77E-05	**Left P**	5%	**55%**	3.17E-01
Right X	1.72E-04	**60%**	2.12E-05	**Right X**	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	2.57E-05	**Right P**	95%	**65%**	3.34E-01
Diff X	1.68E-04	**70%**	3.21E-05	**Diff X**	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	4.12E-05	**Diff P**	90%	**75%**	3.51E-01
#Errors	0	**80%**	5.28E-05	**#Errors**	0	**80%**	3.60E-01
Filter Min	Off	**85%**	7.34E-05	**Filter Min**	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	1.08E-04	**Filter Max**	Off	**90%**	3.77E-01
#Filtered	0	**95%**	1.72E-04	**#Filtered**	0	**95%**	3.85E-01

B1F. (GENEGROUP II TYPE 4)

PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



Summary Statistics for Genogroup II Type 4

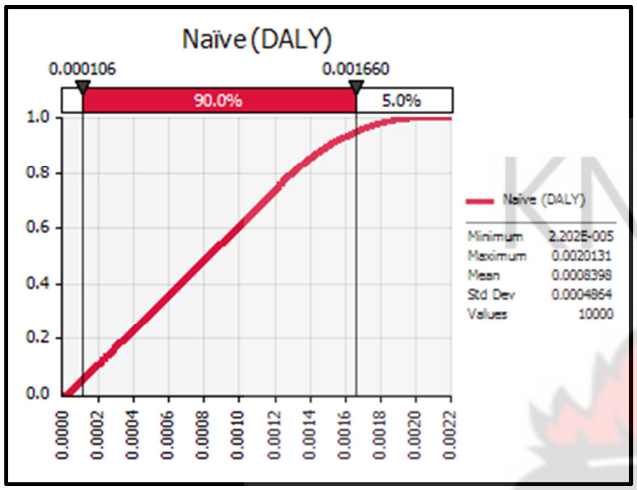
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	2.30E-08	**5%**	3.14E-08	**Minimum**	3.04E-05	**5%**	8.36E-05
Maximum	7.19E-08	**10%**	3.34E-08	**Maximum**	5.04E-03	**10%**	1.45E-04
Mean	4.39E-08	**15%**	3.50E-08	**Mean**	1.50E-03	**15%**	2.21E-04
Std Dev	8.54E-09	**20%**	3.63E-08	**Std Dev**	1.21E-03	**20%**	3.16E-04
Variance	7.28569E-17	**25%**	3.75E-08	**Variance**	1.47255E-06	**25%**	4.20E-04
Skewness	0.465230605	**30%**	3.86E-08	**Skewness**	0.672747896	**30%**	5.47E-04
Kurtosis	2.799305973	**35%**	3.97E-08	**Kurtosis**	2.414534399	**35%**	6.87E-04
Median	4.30E-08	**40%**	4.08E-08	**Median**	1.21E-03	**40%**	8.43E-04
Mode	4.14E-08	**45%**	4.19E-08	**Mode**	5.99E-05	**45%**	1.02E-03
Left X	3.14E-08	**50%**	4.30E-08	**Left X**	8.36E-05	**50%**	1.21E-03
Left P	5%	**55%**	4.41E-08	**Left P**	5%	**55%**	1.41E-03
Right X	5.95E-08	**60%**	4.52E-08	**Right X**	3.78E-03	**60%**	1.62E-03
Right P	95%	**65%**	4.65E-08	**Right P**	95%	**65%**	1.87E-03
Diff X	2.81E-08	**70%**	4.79E-08	**Diff X**	3.70E-03	**70%**	2.13E-03
Diff P	90%	**75%**	4.94E-08	**Diff P**	90%	**75%**	2.41E-03
#Errors	0	**80%**	5.11E-08	**#Errors**	0	**80%**	2.68E-03
Filter Min	Off	**85%**	5.32E-08	**Filter Min**	Off	**85%**	2.97E-03
Filter Max	Off	**90%**	5.59E-08	**Filter Max**	Off	**90%**	3.32E-03
#Filtered	0	**95%**	5.95E-08	**#Filtered**	0	**95%**	3.78E-03

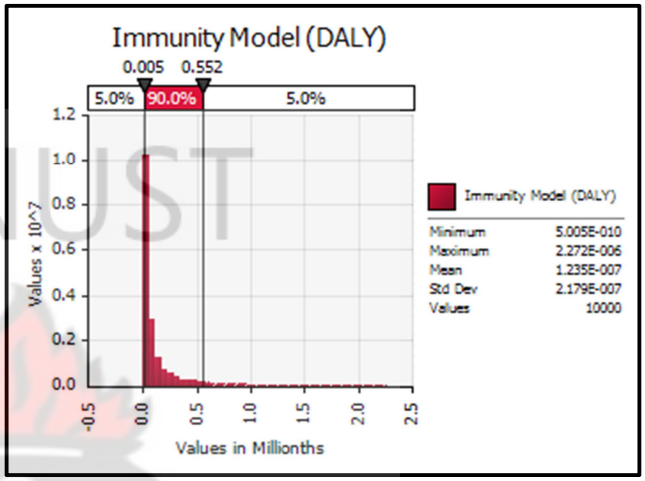
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	2.39E-06	**5%**	3.61E-06	**Minimum**	2.24E-01	**5%**	2.33E-01
Maximum	3.33E-04	**10%**	4.20E-06	**Maximum**	3.93E-01	**10%**	2.41E-01
Mean	2.93E-05	**15%**	4.74E-06	**Mean**	3.09E-01	**15%**	2.50E-01
Std Dev	4.35E-05	**20%**	5.25E-06	**Std Dev**	4.89E-02	**20%**	2.58E-01
Variance	1.89012E-09	**25%**	5.83E-06	**Variance**	0.002387257	**25%**	2.67E-01
Skewness	2.766709876	**30%**	6.57E-06	**Skewness**	-2.07672E-06	**30%**	2.75E-01
Kurtosis	11.25084402	**35%**	7.45E-06	**Kurtosis**	1.799997377	**35%**	2.83E-01
Median	1.14E-05	**40%**	8.55E-06	**Median**	3.09E-01	**40%**	2.92E-01
Mode	4.58E-06	**45%**	9.85E-06	**Mode**	3.69E-01	**45%**	3.00E-01
Left X	3.61E-06	**50%**	1.14E-05	**Left X**	2.33E-01	**50%**	3.09E-01
Left P	5%	**55%**	1.31E-05	**Left P**	5%	**55%**	3.17E-01
Right X	1.32E-04	**60%**	1.58E-05	**Right X**	3.85E-01	**60%**	3.26E-01
Right P	95%	**65%**	1.90E-05	**Right P**	95%	**65%**	3.34E-01
Diff X	1.28E-04	**70%**	2.37E-05	**Diff X**	1.52E-01	**70%**	3.43E-01
Diff P	90%	**75%**	3.02E-05	**Diff P**	90%	**75%**	3.51E-01
#Errors	0	**80%**	4.02E-05	**#Errors**	0	**80%**	3.60E-01
Filter Min	Off	**85%**	5.45E-05	**Filter Min**	Off	**85%**	3.68E-01
Filter Max	Off	**90%**	8.13E-05	**Filter Max**	Off	**90%**	3.77E-01
#Filtered	0	**95%**	1.32E-04	**#Filtered**	0	**95%**	3.85E-01

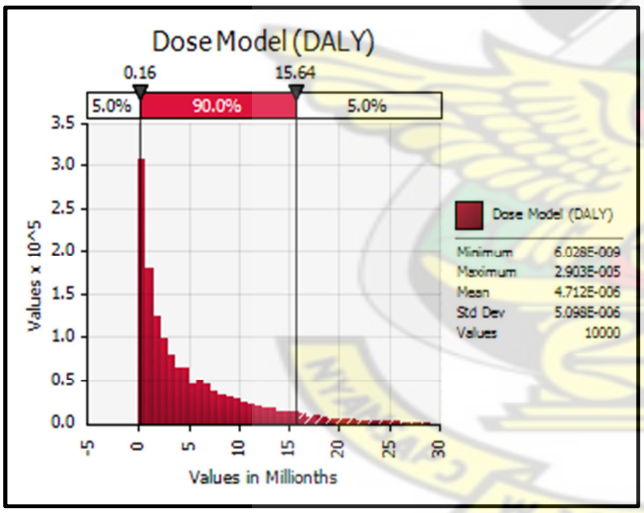
APPENDIX B2 DAILY ADJUSTED LIFE YEARS

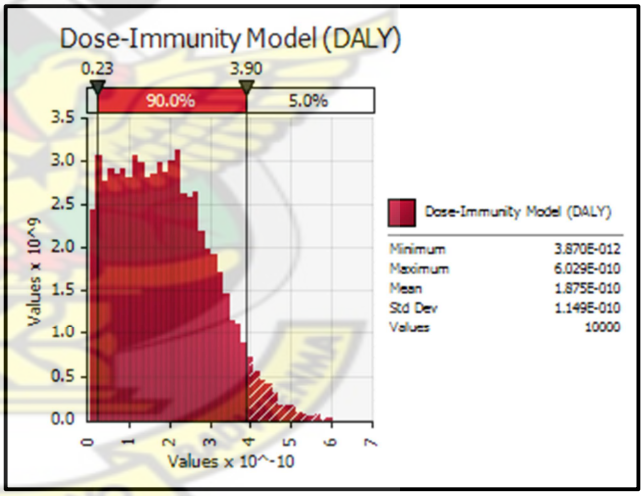
B2A. (SYMPTOMATIC INFECTIOUNESS)

PROBABILITY GRAPHS FOR DR MODELS

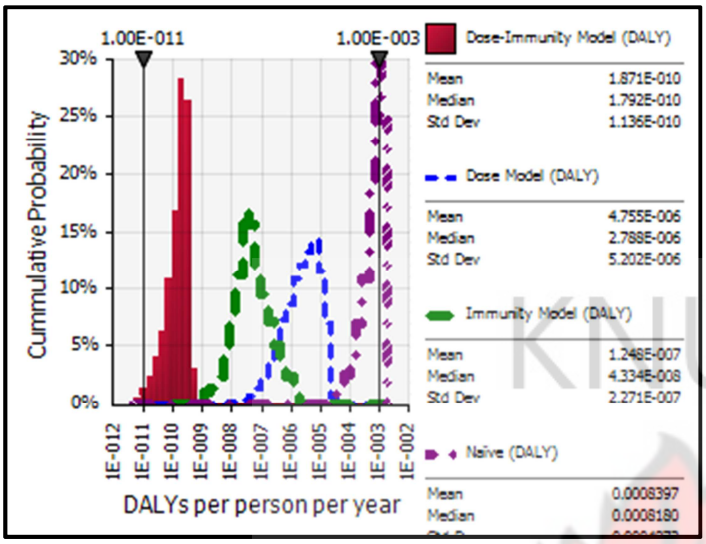




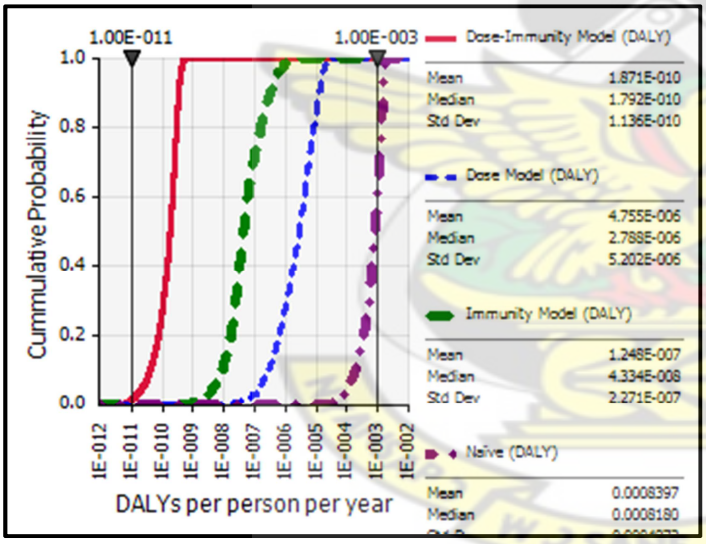




PROBABILITY GRAPH FOR DALY SYMPTOMATIC INFECTIOUNESS



CUMMULATIVE GRAPH FOR DALY SYMPTOMATIC INFECTIOUNESS

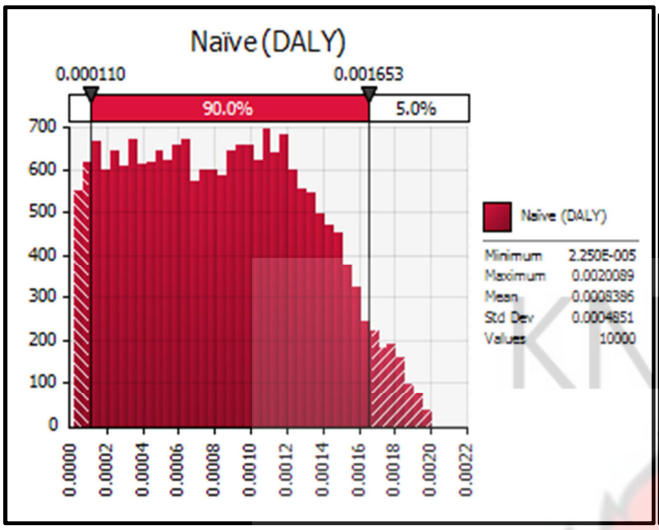


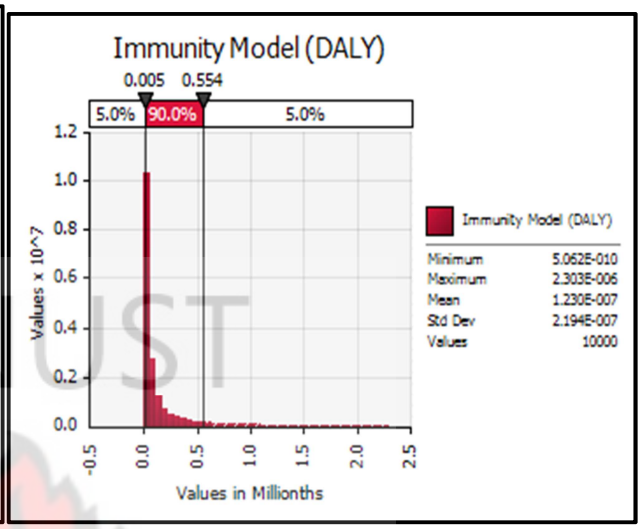
Summary Statistics for Symptomatic Infectiousness Individual

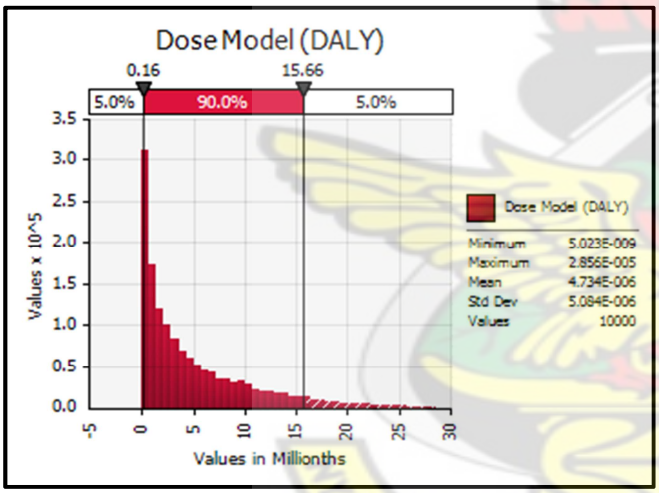
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	3.87E-12	**5%**	2.32E-11	**Minimum**	6.03E-09	**5%**	1.59E-07
Maximum	6.03E-10	**10%**	4.08E-11	**Maximum**	2.90E-05	**10%**	2.89E-07
Mean	1.87E-10	**15%**	5.77E-11	**Mean**	4.71E-06	**15%**	4.41E-07
Std Dev	1.15E-10	**20%**	7.57E-11	**Std Dev**	5.10E-06	**20%**	6.29E-07
Variance	1.32061E-20	**25%**	9.25E-11	**Variance**	2.59851E-11	**25%**	8.61E-07
Skewness	0.452798307	**30%**	1.10E-10	**Skewness**	1.521360609	**30%**	1.12E-06
Kurtosis	2.592060219	**35%**	1.27E-10	**Kurtosis**	5.037160683	**35%**	1.44E-06
Median	1.79E-10	**40%**	1.44E-10	**Median**	2.78E-06	**40%**	1.82E-06
Mode	2.08E-10	**45%**	1.62E-10	**Mode**	2.23E-07	**45%**	2.25E-06
Left X	2.32E-11	**50%**	1.79E-10	**Left X**	1.59E-07	**50%**	2.78E-06
Left P	5%	**55%**	1.96E-10	**Left P**	5%	**55%**	3.35E-06
Right X	3.90E-10	**60%**	2.13E-10	**Right X**	1.56E-05	**60%**	4.05E-06
Right P	95%	**65%**	2.29E-10	**Right P**	95%	**65%**	4.88E-06
Diff X	3.67E-10	**70%**	2.48E-10	**Diff X**	1.55E-05	**70%**	5.92E-06
Diff P	90%	**75%**	2.67E-10	**Diff P**	90%	**75%**	7.00E-06
#Errors	0	**80%**	2.89E-10	**#Errors**	0	**80%**	8.29E-06
Filter Min	Off	**85%**	3.15E-10	**Filter Min**	Off	**85%**	1.00E-05
Filter Max	Off	**90%**	3.45E-10	**Filter Max**	Off	**90%**	1.22E-05
#Filtered	0	**95%**	3.90E-10	**#Filtered**	0	**95%**	1.56E-05

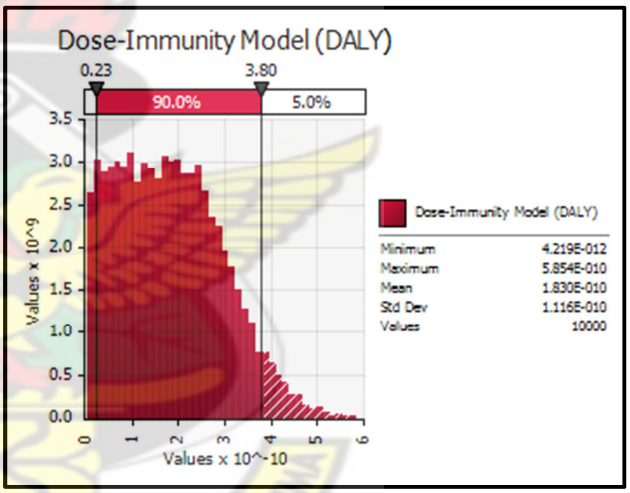
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	5.01E-10	**5%**	5.01E-09	**Minimum**	2.20E-05	**5%**	1.06E-04
Maximum	2.27E-06	**10%**	8.69E-09	**Maximum**	2.01E-03	**10%**	1.83E-04
Mean	1.23E-07	**15%**	1.23E-08	**Mean**	8.40E-04	**15%**	2.67E-04
Std Dev	2.18E-07	**20%**	1.59E-08	**Std Dev**	4.86E-04	**20%**	3.46E-04
Variance	4.74898E-14	**25%**	1.98E-08	**Variance**	2.36544E-07	**25%**	4.26E-04
Skewness	3.73522302	**30%**	2.39E-08	**Skewness**	0.171402279	**30%**	5.08E-04
Kurtosis	20.47696808	**35%**	2.79E-08	**Kurtosis**	2.040761429	**35%**	5.89E-04
Median	4.36E-08	**40%**	3.24E-08	**Median**	8.26E-04	**40%**	6.62E-04
Mode	6.09E-09	**45%**	3.75E-08	**Mode**	8.62E-04	**45%**	7.43E-04
Left X	5.01E-09	**50%**	4.36E-08	**Left X**	1.06E-04	**50%**	8.26E-04
Left P	5%	**55%**	5.17E-08	**Left P**	5%	**55%**	9.06E-04
Right X	5.52E-07	**60%**	6.17E-08	**Right X**	1.66E-03	**60%**	9.87E-04
Right P	95%	**65%**	7.51E-08	**Right P**	95%	**65%**	1.07E-03
Diff X	5.47E-07	**70%**	9.28E-08	**Diff X**	1.55E-03	**70%**	1.15E-03
Diff P	90%	**75%**	1.17E-07	**Diff P**	90%	**75%**	1.22E-03
#Errors	0	**80%**	1.55E-07	**#Errors**	0	**80%**	1.30E-03
Filter Min	Off	**85%**	2.14E-07	**Filter Min**	Off	**85%**	1.39E-03
Filter Max	Off	**90%**	3.22E-07	**Filter Max**	Off	**90%**	1.51E-03
#Filtered	0	**95%**	5.52E-07	**#Filtered**	0	**95%**	1.66E-03

B2B. (PRE-SYMPTOMATIC AND POST SYMPTOMATIC INFECTIOUNESS LOW)

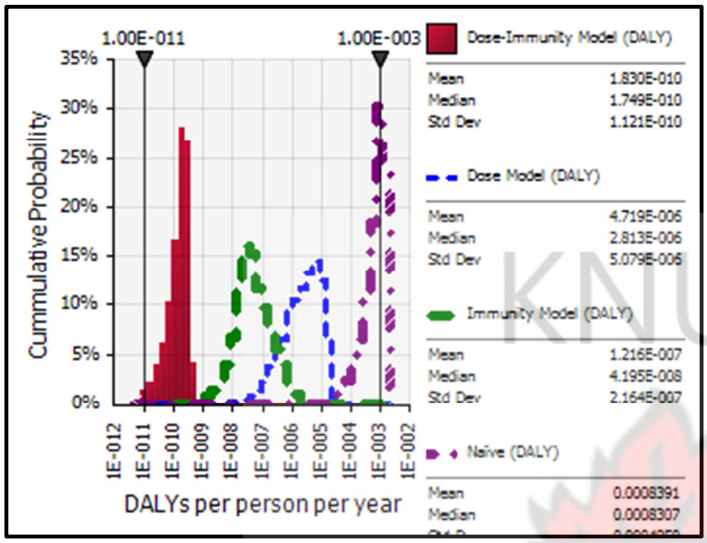




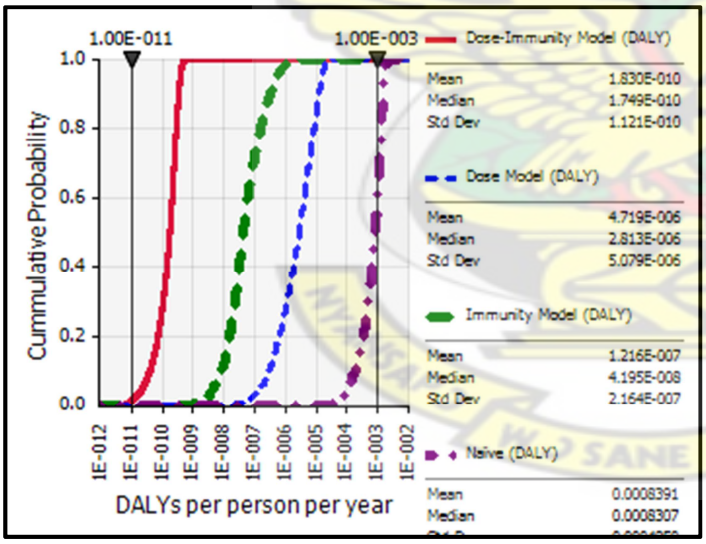




PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAP FOR DR MODELS

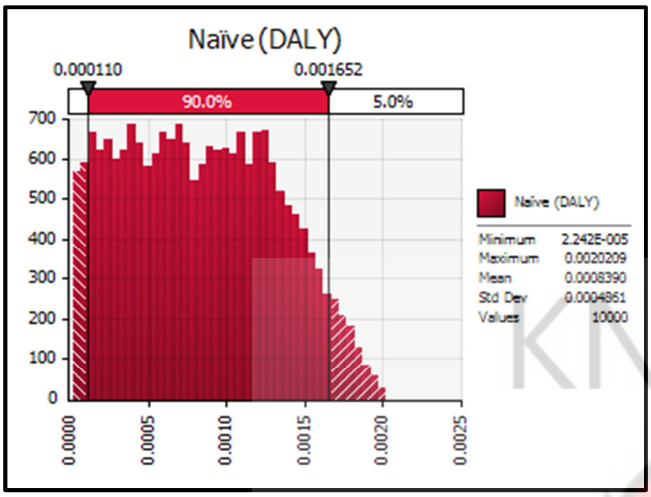


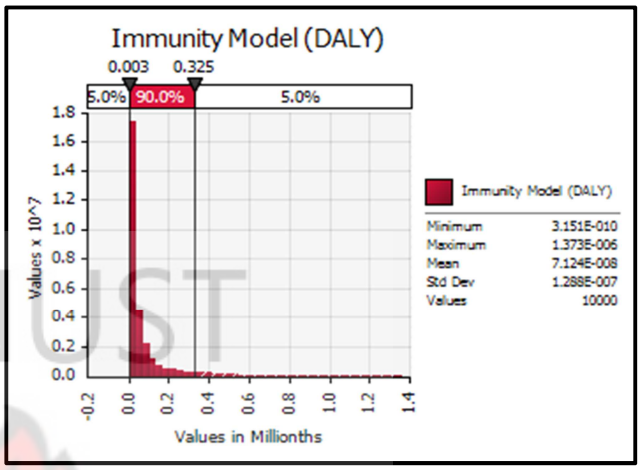
Summary Statistics for Pre-Symptomatic and Post-Symptomatic Low

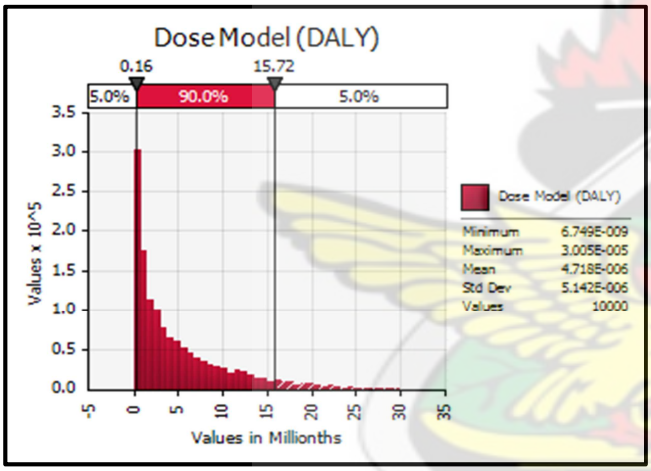
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	4.22E-12	**5%**	2.28E-11	**Minimum**	5.02E-09	**5%**	1.56E-07
Maximum	5.85E-10	**10%**	3.97E-11	**Maximum**	2.86E-05	**10%**	2.83E-07
Mean	1.83E-10	**15%**	5.72E-11	**Mean**	4.73E-06	**15%**	4.34E-07
Std Dev	1.12E-10	**20%**	7.30E-11	**Std Dev**	5.08E-06	**20%**	6.14E-07
Variance	1.24636E-20	**25%**	9.01E-11	**Variance**	2.58503E-11	**25%**	8.41E-07
Skewness	0.426850051	**30%**	1.06E-10	**Skewness**	1.469588941	**30%**	1.13E-06
Kurtosis	2.549796025	**35%**	1.24E-10	**Kurtosis**	4.780574204	**35%**	1.46E-06
Median	1.76E-10	**40%**	1.41E-10	**Median**	2.81E-06	**40%**	1.84E-06
Mode	9.85E-11	**45%**	1.58E-10	**Mode**	2.23E-07	**45%**	2.32E-06
Left X	2.28E-11	**50%**	1.76E-10	**Left X**	1.56E-07	**50%**	2.81E-06
Left P	5%	**55%**	1.92E-10	**Left P**	5%	**55%**	3.42E-06
Right X	3.80E-10	**60%**	2.09E-10	**Right X**	1.57E-05	**60%**	4.10E-06
Right P	95%	**65%**	2.26E-10	**Right P**	95%	**65%**	4.92E-06
Diff X	3.58E-10	**70%**	2.43E-10	**Diff X**	1.55E-05	**70%**	5.90E-06
Diff P	90%	**75%**	2.61E-10	**Diff P**	90%	**75%**	7.05E-06
#Errors	0	**80%**	2.82E-10	**#Errors**	0	**80%**	8.46E-06
Filter Min	Off	**85%**	3.05E-10	**Filter Min**	Off	**85%**	1.01E-05
Filter Max	Off	**90%**	3.35E-10	**Filter Max**	Off	**90%**	1.24E-05
#Filtered	0	**95%**	3.80E-10	**#Filtered**	0	**95%**	1.57E-05

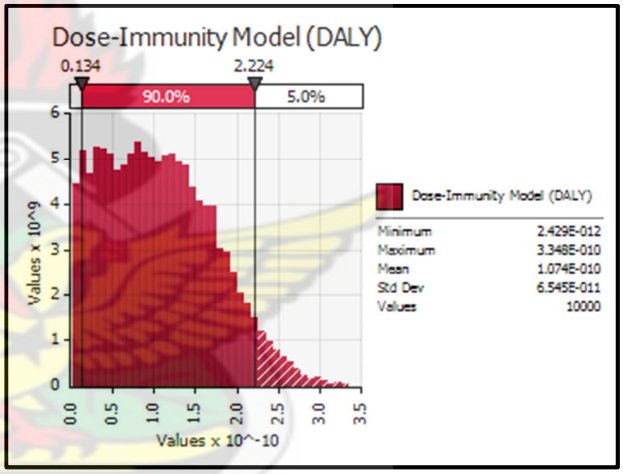
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	5.06E-10	**5%**	4.57E-09	**Minimum**	2.25E-05	**5%**	1.10E-04
Maximum	2.30E-06	**10%**	8.37E-09	**Maximum**	2.01E-03	**10%**	1.87E-04
Mean	1.23E-07	**15%**	1.21E-08	**Mean**	8.39E-04	**15%**	2.67E-04
Std Dev	2.19E-07	**20%**	1.60E-08	**Std Dev**	4.85E-04	**20%**	3.44E-04
Variance	4.81227E-14	**25%**	1.94E-08	**Variance**	2.35281E-07	**25%**	4.24E-04
Skewness	3.73975682	**30%**	2.33E-08	**Skewness**	0.173813979	**30%**	5.05E-04
Kurtosis	20.58032945	**35%**	2.70E-08	**Kurtosis**	2.047171235	**35%**	5.85E-04
Median	4.18E-08	**40%**	3.12E-08	**Median**	8.28E-04	**40%**	6.60E-04
Mode	1.63E-08	**45%**	3.61E-08	**Mode**	6.37E-04	**45%**	7.45E-04
Left X	4.57E-09	**50%**	4.18E-08	**Left X**	1.10E-04	**50%**	8.28E-04
Left P	5%	**55%**	4.97E-08	**Left P**	5%	**55%**	9.11E-04
Right X	5.54E-07	**60%**	5.94E-08	**Right X**	1.65E-03	**60%**	9.86E-04
Right P	95%	**65%**	7.23E-08	**Right P**	95%	**65%**	1.06E-03
Diff X	5.49E-07	**70%**	9.05E-08	**Diff X**	1.54E-03	**70%**	1.14E-03
Diff P	90%	**75%**	1.16E-07	**Diff P**	90%	**75%**	1.21E-03
#Errors	0	**80%**	1.53E-07	**#Errors**	0	**80%**	1.30E-03
Filter Min	Off	**85%**	2.16E-07	**Filter Min**	Off	**85%**	1.39E-03
Filter Max	Off	**90%**	3.27E-07	**Filter Max**	Off	**90%**	1.50E-03
#Filtered	0	**95%**	5.54E-07	**#Filtered**	0	**95%**	1.65E-03

B2C. (PRE-SYMPTOMATIC AND POST SYMPTOMATIC INFECTIOUNESS HIGH)

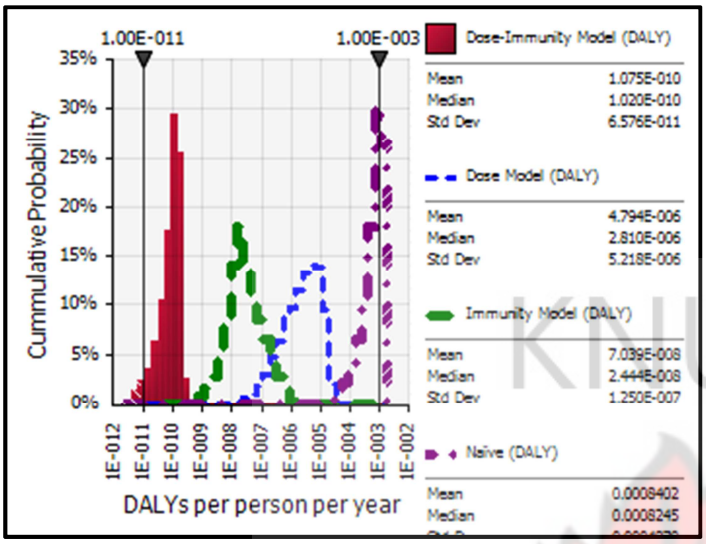




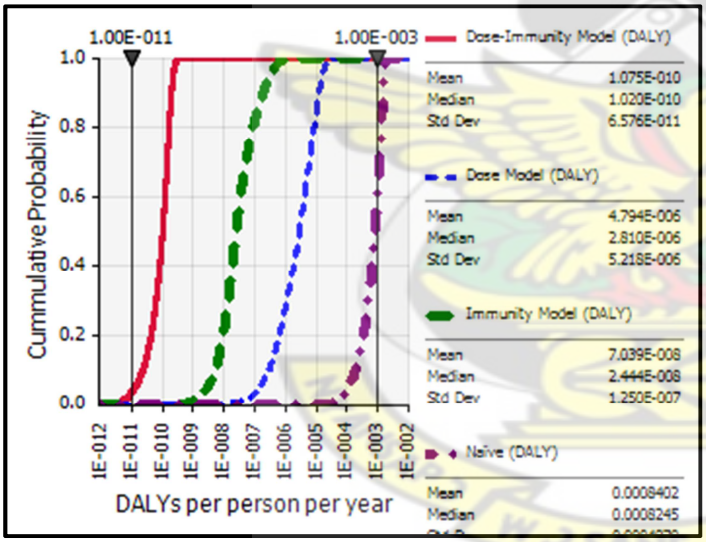




PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS

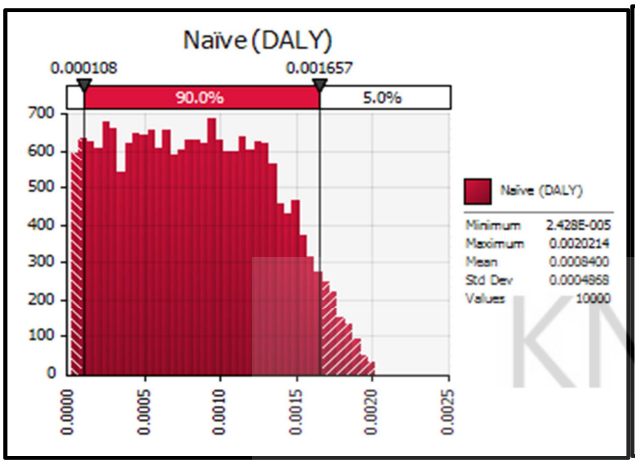


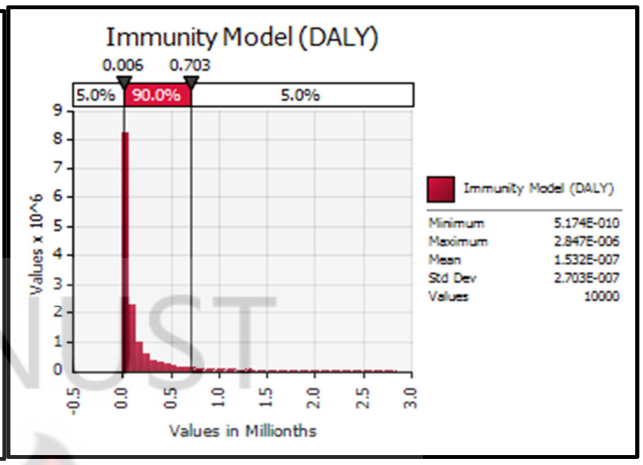
Summary Statistics for Pre-Symptomatic and Post-Symptomatic High

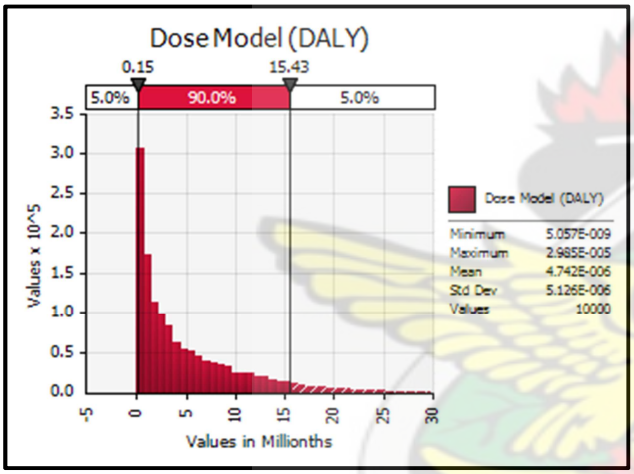
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	2.43E-12	**5%**	1.34E-11	**Minimum**	6.75E-09	**5%**	1.61E-07
Maximum	3.35E-10	**10%**	2.34E-11	**Maximum**	3.00E-05	**10%**	2.98E-07
Mean	1.07E-10	**15%**	3.34E-11	**Mean**	4.72E-06	**15%**	4.38E-07
Std Dev	6.55E-11	**20%**	4.28E-11	**Std Dev**	5.14E-06	**20%**	6.33E-07
Variance	4.28405E-21	**25%**	5.28E-11	**Variance**	2.64408E-11	**25%**	8.53E-07
Skewness	0.418612994	**30%**	6.33E-11	**Skewness**	1.584434156	**30%**	1.12E-06
Kurtosis	2.500428674	**35%**	7.35E-11	**Kurtosis**	5.382651481	**35%**	1.44E-06
Median	1.02E-10	**40%**	8.26E-11	**Median**	2.78E-06	**40%**	1.85E-06
Mode	4.20E-11	**45%**	9.25E-11	**Mode**	3.09E-07	**45%**	2.30E-06
Left X	1.34E-11	**50%**	1.02E-10	**Left X**	1.61E-07	**50%**	2.78E-06
Left P	5%	**55%**	1.12E-10	**Left P**	5%	**55%**	3.39E-06
Right X	2.22E-10	**60%**	1.22E-10	**Right X**	1.57E-05	**60%**	4.08E-06
Right P	95%	**65%**	1.32E-10	**Right P**	95%	**65%**	4.88E-06
Diff X	2.09E-10	**70%**	1.42E-10	**Diff X**	1.56E-05	**70%**	5.78E-06
Diff P	90%	**75%**	1.54E-10	**Diff P**	90%	**75%**	6.87E-06
#Errors	0	**80%**	1.67E-10	**#Errors**	0	**80%**	8.23E-06
Filter Min	Off	**85%**	1.80E-10	**Filter Min**	Off	**85%**	1.00E-05
Filter Max	Off	**90%**	1.97E-10	**Filter Max**	Off	**90%**	1.23E-05
#Filtered	0	**95%**	2.22E-10	**#Filtered**	0	**95%**	1.57E-05

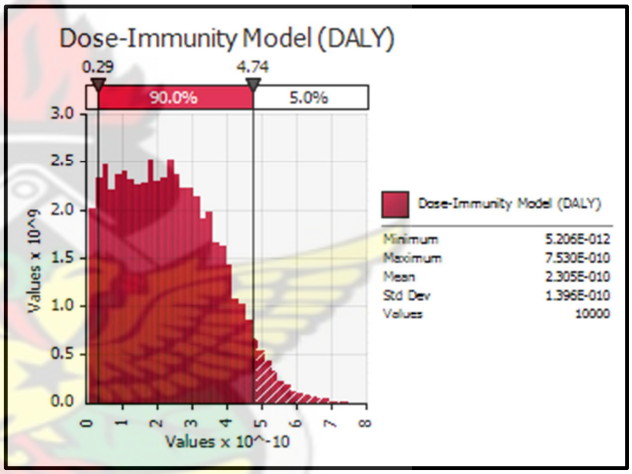
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	3.15E-10	**5%**	2.76E-09	**Minimum**	2.24E-05	**5%**	1.10E-04
Maximum	1.37E-06	**10%**	5.01E-09	**Maximum**	2.02E-03	**10%**	1.87E-04
Mean	7.12E-08	**15%**	7.08E-09	**Mean**	8.39E-04	**15%**	2.67E-04
Std Dev	1.29E-07	**20%**	9.40E-09	**Std Dev**	4.86E-04	**20%**	3.47E-04
Variance	1.65967E-14	**25%**	1.17E-08	**Variance**	2.3629E-07	**25%**	4.22E-04
Skewness	3.955094049	**30%**	1.40E-08	**Skewness**	0.176484417	**30%**	5.04E-04
Kurtosis	23.09409582	**35%**	1.62E-08	**Kurtosis**	2.032922131	**35%**	5.86E-04
Median	2.53E-08	**40%**	1.86E-08	**Median**	8.22E-04	**40%**	6.62E-04
Mode	7.75E-09	**45%**	2.16E-08	**Mode**	7.15E-04	**45%**	7.38E-04
Left X	2.76E-09	**50%**	2.53E-08	**Left X**	1.10E-04	**50%**	8.22E-04
Left P	5%	**55%**	2.93E-08	**Left P**	5%	**55%**	9.03E-04
Right X	3.25E-07	**60%**	3.47E-08	**Right X**	1.65E-03	**60%**	9.84E-04
Right P	95%	**65%**	4.17E-08	**Right P**	95%	**65%**	1.07E-03
Diff X	3.22E-07	**70%**	5.25E-08	**Diff X**	1.54E-03	**70%**	1.14E-03
Diff P	90%	**75%**	6.66E-08	**Diff P**	90%	**75%**	1.22E-03
#Errors	0	**80%**	8.74E-08	**#Errors**	0	**80%**	1.30E-03
Filter Min	Off	**85%**	1.21E-07	**Filter Min**	Off	**85%**	1.39E-03
Filter Max	Off	**90%**	1.84E-07	**Filter Max**	Off	**90%**	1.51E-03
#Filtered	0	**95%**	3.25E-07	**#Filtered**	0	**95%**	1.65E-03

B2D. (INNATE GENETIC RESISTANCE)

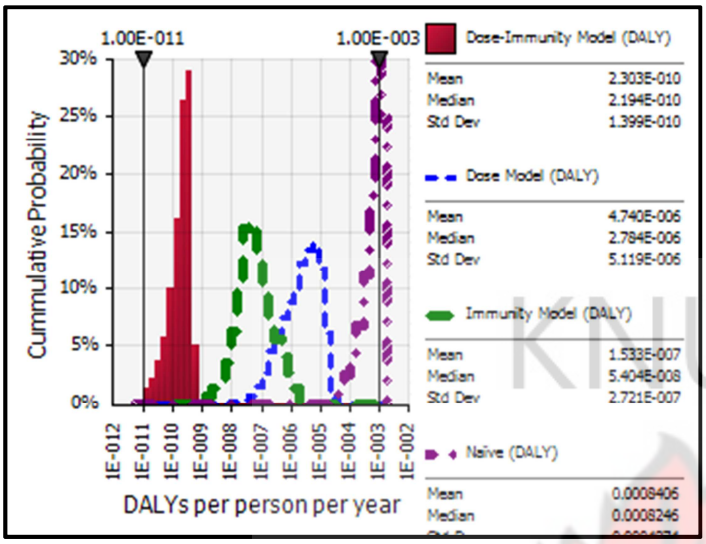




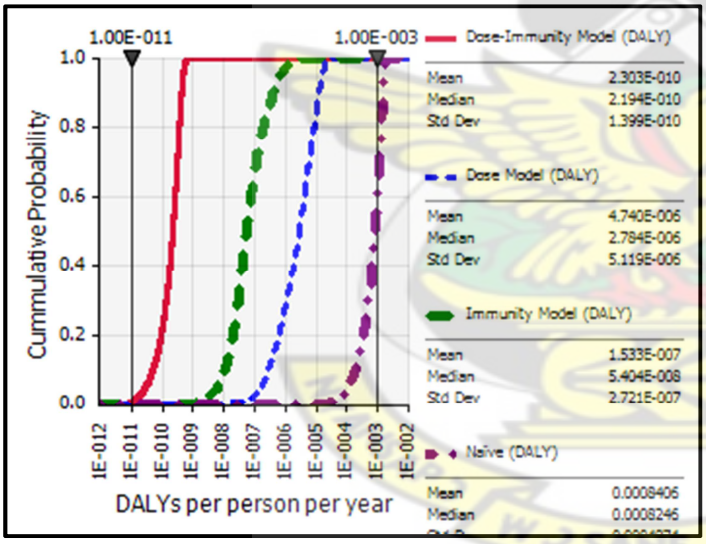




PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS

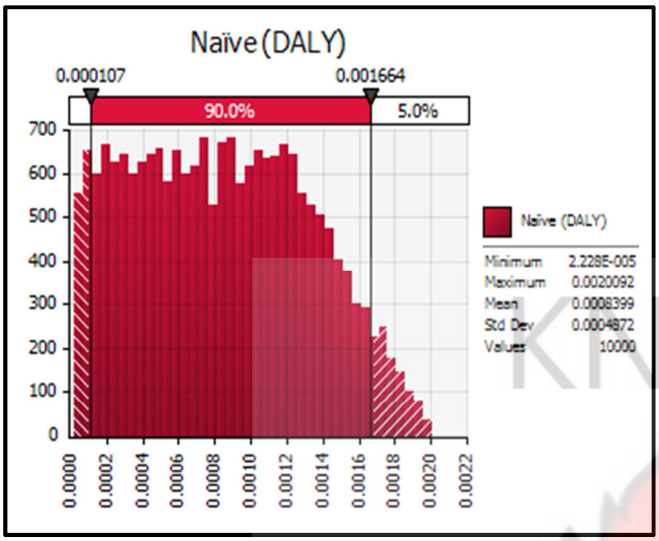


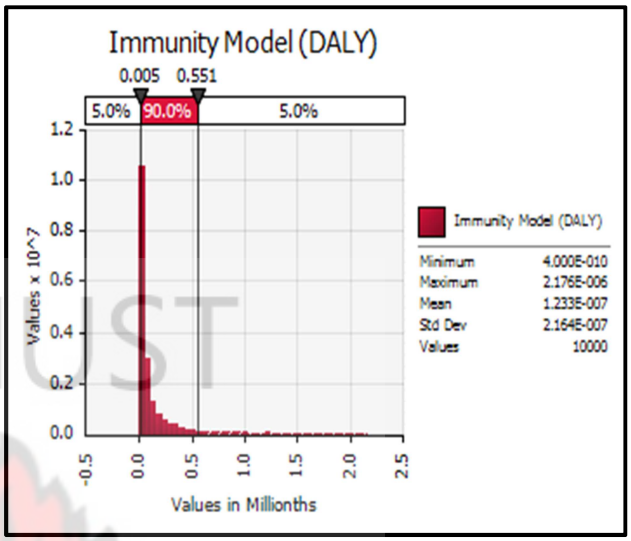
Summary Statistics for Innate Genetic Resistance

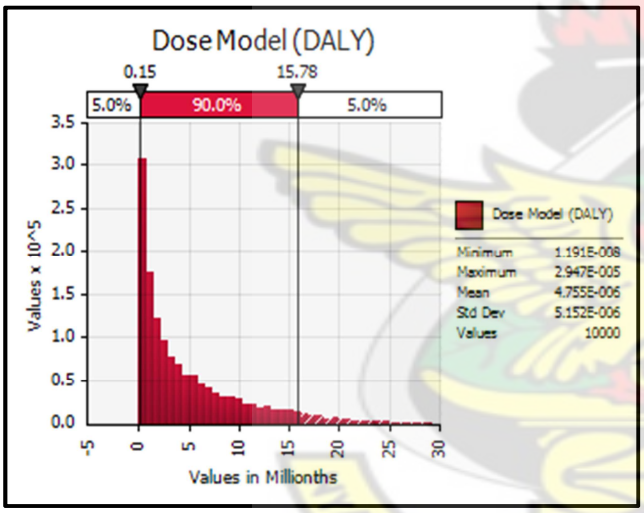
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	5.21E-12	**5%**	2.85E-11	**Minimum**	5.06E-09	**5%**	1.54E-07
Maximum	7.53E-10	**10%**	4.97E-11	**Maximum**	2.99E-05	**10%**	2.87E-07
Mean	2.30E-10	**15%**	7.10E-11	**Mean**	4.74E-06	**15%**	4.37E-07
Std Dev	1.40E-10	**20%**	9.29E-11	**Std Dev**	5.13E-06	**20%**	6.26E-07
Variance	1.9494E-20	**25%**	1.14E-10	**Variance**	2.62801E-11	**25%**	8.63E-07
Skewness	0.394049907	**30%**	1.36E-10	**Skewness**	1.5297722	**30%**	1.12E-06
Kurtosis	2.47116647	**35%**	1.57E-10	**Kurtosis**	5.145107801	**35%**	1.43E-06
Median	2.21E-10	**40%**	1.79E-10	**Median**	2.80E-06	**40%**	1.85E-06
Mode	2.44E-10	**45%**	1.99E-10	**Mode**	2.18E-07	**45%**	2.30E-06
Left X	2.85E-11	**50%**	2.21E-10	**Left X**	1.54E-07	**50%**	2.80E-06
Left P	5%	**55%**	2.42E-10	**Left P**	5%	**55%**	3.39E-06
Right X	4.74E-10	**60%**	2.62E-10	**Right X**	1.54E-05	**60%**	4.08E-06
Right P	95%	**65%**	2.84E-10	**Right P**	95%	**65%**	4.91E-06
Diff X	4.45E-10	**70%**	3.07E-10	**Diff X**	1.53E-05	**70%**	5.91E-06
Diff P	90%	**75%**	3.31E-10	**Diff P**	90%	**75%**	7.05E-06
#Errors	0	**80%**	3.56E-10	**#Errors**	0	**80%**	8.45E-06
Filter Min	Off	**85%**	3.85E-10	**Filter Min**	Off	**85%**	1.01E-05
Filter Max	Off	**90%**	4.20E-10	**Filter Max**	Off	**90%**	1.23E-05
#Filtered	0	**95%**	4.74E-10	**#Filtered**	0	**95%**	1.54E-05

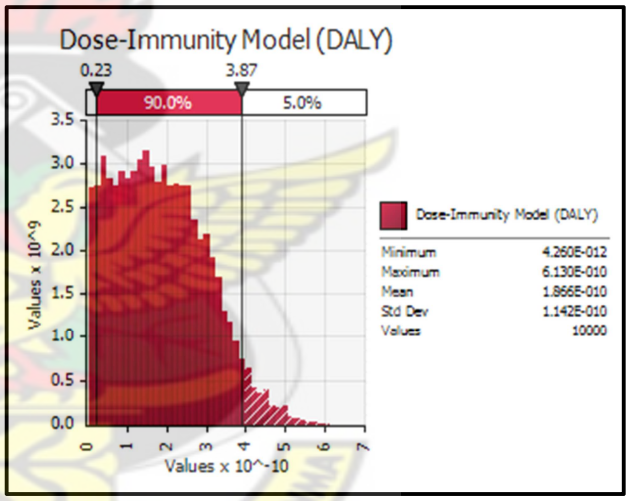
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	5.17E-10	**5%**	6.22E-09	**Minimum**	2.43E-05	**5%**	1.08E-04
Maximum	2.85E-06	**10%**	1.10E-08	**Maximum**	2.02E-03	**10%**	1.86E-04
Mean	1.53E-07	**15%**	1.55E-08	**Mean**	8.40E-04	**15%**	2.65E-04
Std Dev	2.70E-07	**20%**	2.03E-08	**Std Dev**	4.87E-04	**20%**	3.45E-04
Variance	7.30434E-14	**25%**	2.47E-08	**Variance**	2.36977E-07	**25%**	4.29E-04
Skewness	3.697836184	**30%**	2.95E-08	**Skewness**	0.17210223	**30%**	5.05E-04
Kurtosis	20.40807467	**35%**	3.42E-08	**Kurtosis**	2.032512639	**35%**	5.83E-04
Median	5.25E-08	**40%**	3.92E-08	**Median**	8.26E-04	**40%**	6.62E-04
Mode	2.43E-08	**45%**	4.54E-08	**Mode**	4.99E-04	**45%**	7.45E-04
Left X	6.22E-09	**50%**	5.25E-08	**Left X**	1.08E-04	**50%**	8.26E-04
Left P	5%	**55%**	6.20E-08	**Left P**	5%	**55%**	9.08E-04
Right X	7.03E-07	**60%**	7.53E-08	**Right X**	1.66E-03	**60%**	9.83E-04
Right P	95%	**65%**	9.11E-08	**Right P**	95%	**65%**	1.06E-03
Diff X	6.97E-07	**70%**	1.13E-07	**Diff X**	1.55E-03	**70%**	1.14E-03
Diff P	90%	**75%**	1.44E-07	**Diff P**	90%	**75%**	1.23E-03
#Errors	0	**80%**	1.92E-07	**#Errors**	0	**80%**	1.31E-03
Filter Min	Off	**85%**	2.67E-07	**Filter Min**	Off	**85%**	1.40E-03
Filter Max	Off	**90%**	4.04E-07	**Filter Max**	Off	**90%**	1.51E-03
#Filtered	0	**95%**	7.03E-07	**#Filtered**	0	**95%**	1.66E-03

B2E. (NO-IMMUNE BOOSTING AFTER ASYMPTOMATIC INFECTIOUNESS)

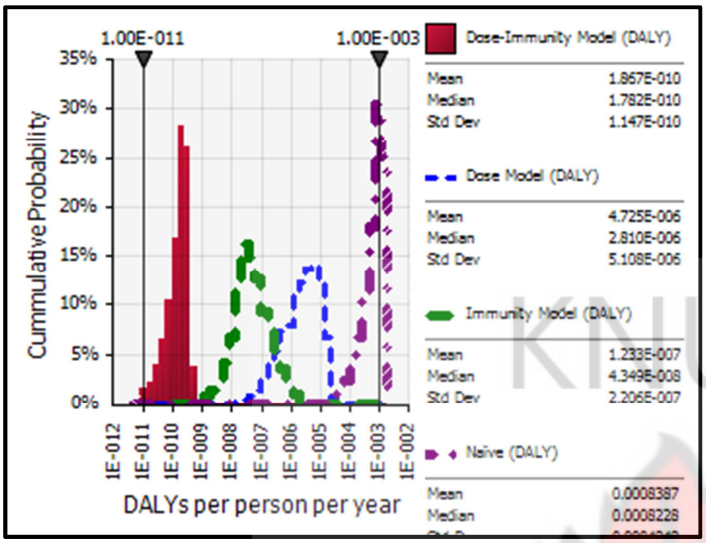




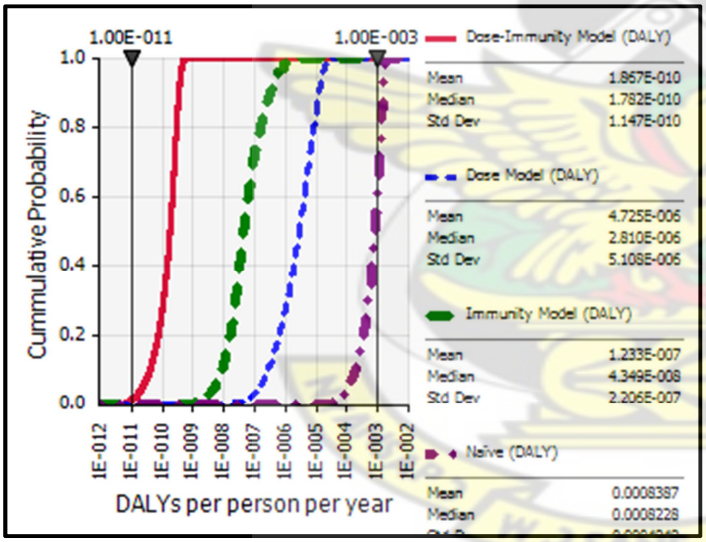




PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS

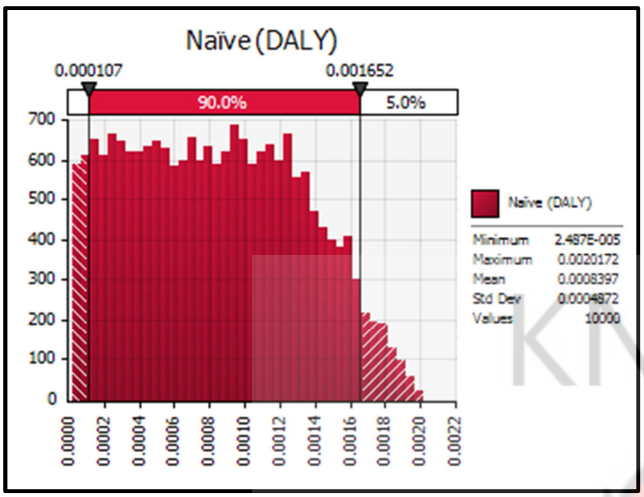


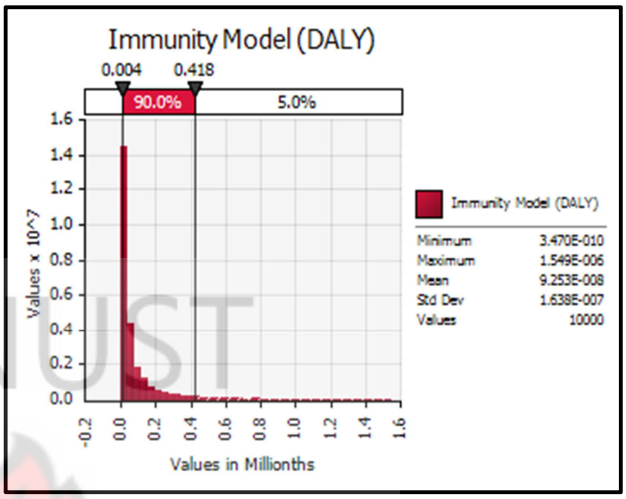
Summary Statistics for No-Immune Boosting after Asymptomatic Infectiousness

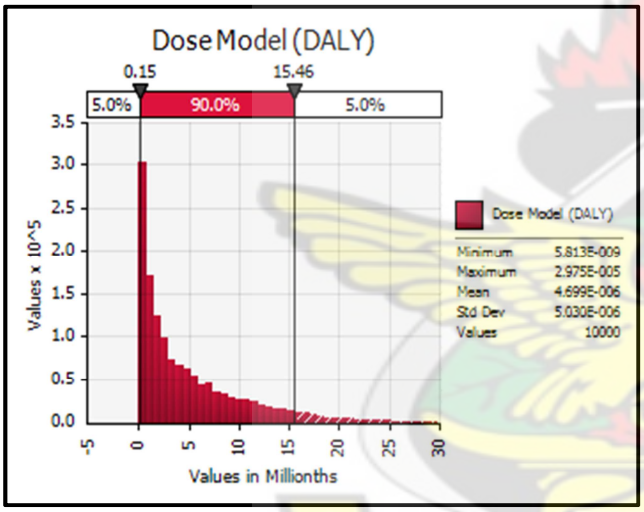
Dose-Immunity Model / Incidence per person per year				Dose Model / Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	4.26E-12	**5%**	2.31E-11	**Minimum**	1.19E-08	**5%**	1.54E-07
Maximum	6.13E-10	**10%**	4.00E-11	**Maximum**	2.95E-05	**10%**	2.88E-07
Mean	1.87E-10	**15%**	5.70E-11	**Mean**	4.76E-06	**15%**	4.45E-07
Std Dev	1.14E-10	**20%**	7.50E-11	**Std Dev**	5.15E-06	**20%**	6.32E-07
Variance	1.30332E-20	**25%**	9.25E-11	**Variance**	2.65423E-11	**25%**	8.61E-07
Skewness	0.438631402	**30%**	1.10E-10	**Skewness**	1.503421922	**30%**	1.12E-06
Kurtosis	2.556541133	**35%**	1.27E-10	**Kurtosis**	4.939202974	**35%**	1.45E-06
Median	1.77E-10	**40%**	1.44E-10	**Median**	2.80E-06	**40%**	1.81E-06
Mode	1.48E-10	**45%**	1.60E-10	**Mode**	1.65E-07	**45%**	2.27E-06
Left X	2.31E-11	**50%**	1.77E-10	**Left X**	1.54E-07	**50%**	2.80E-06
Left P	5%	**55%**	1.94E-10	**Left P**	5%	**55%**	3.38E-06
Right X	3.87E-10	**60%**	2.12E-10	**Right X**	1.58E-05	**60%**	4.08E-06
Right P	95%	**65%**	2.31E-10	**Right P**	95%	**65%**	4.91E-06
Diff X	3.64E-10	**70%**	2.48E-10	**Diff X**	1.56E-05	**70%**	5.85E-06
Diff P	90%	**75%**	2.67E-10	**Diff P**	90%	**75%**	7.01E-06
#Errors	0	**80%**	2.89E-10	**#Errors**	0	**80%**	8.44E-06
Filter Min	Off	**85%**	3.12E-10	**Filter Min**	Off	**85%**	1.02E-05
Filter Max	Off	**90%**	3.42E-10	**Filter Max**	Off	**90%**	1.26E-05
#Filtered	0	**95%**	3.87E-10	**#Filtered**	0	**95%**	1.58E-05

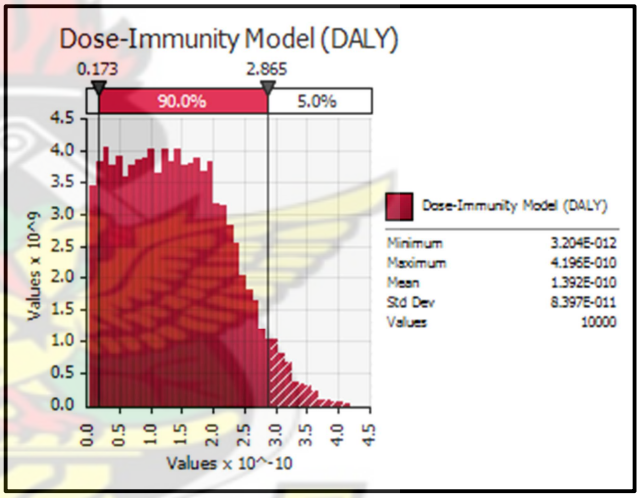
Immunity Model / Incidence per person per year				Naive/ Incidence per person per year			
Statistics		**Percentage**		**Statistics**		**Percentage**	
Minimum	4.00E-10	**5%**	4.73E-09	**Minimum**	2.23E-05	**5%**	1.07E-04
Maximum	2.18E-06	**10%**	8.40E-09	**Maximum**	2.01E-03	**10%**	1.87E-04
Mean	1.23E-07	**15%**	1.23E-08	**Mean**	8.40E-04	**15%**	2.65E-04
Std Dev	2.16E-07	**20%**	1.62E-08	**Std Dev**	4.87E-04	**20%**	3.45E-04
Variance	4.68413E-14	**25%**	2.00E-08	**Variance**	2.37398E-07	**25%**	4.27E-04
Skewness	3.678275458	**30%**	2.38E-08	**Skewness**	0.178388897	**30%**	5.00E-04
Kurtosis	19.86742882	**35%**	2.78E-08	**Kurtosis**	2.043520821	**35%**	5.83E-04
Median	4.39E-08	**40%**	3.24E-08	**Median**	8.27E-04	**40%**	6.65E-04
Mode	1.35E-08	**45%**	3.71E-08	**Mode**	7.38E-04	**45%**	7.44E-04
Left X	4.73E-09	**50%**	4.39E-08	**Left X**	1.07E-04	**50%**	8.27E-04
Left P	5%	**55%**	5.08E-08	**Left P**	5%	**55%**	9.04E-04
Right X	5.51E-07	**60%**	6.07E-08	**Right X**	1.66E-03	**60%**	9.86E-04
Right P	95%	**65%**	7.38E-08	**Right P**	95%	**65%**	1.06E-03
Diff X	5.46E-07	**70%**	9.22E-08	**Diff X**	1.56E-03	**70%**	1.14E-03
Diff P	90%	**75%**	1.18E-07	**Diff P**	90%	**75%**	1.22E-03
#Errors	0	**80%**	1.56E-07	**#Errors**	0	**80%**	1.30E-03
Filter Min	Off	**85%**	2.21E-07	**Filter Min**	Off	**85%**	1.40E-03
Filter Max	Off	**90%**	3.24E-07	**Filter Max**	Off	**90%**	1.51E-03
#Filtered	0	**95%**	5.51E-07	**#Filtered**	0	**95%**	1.66E-03

B2F. (GENEGROUP II TYPE 4)

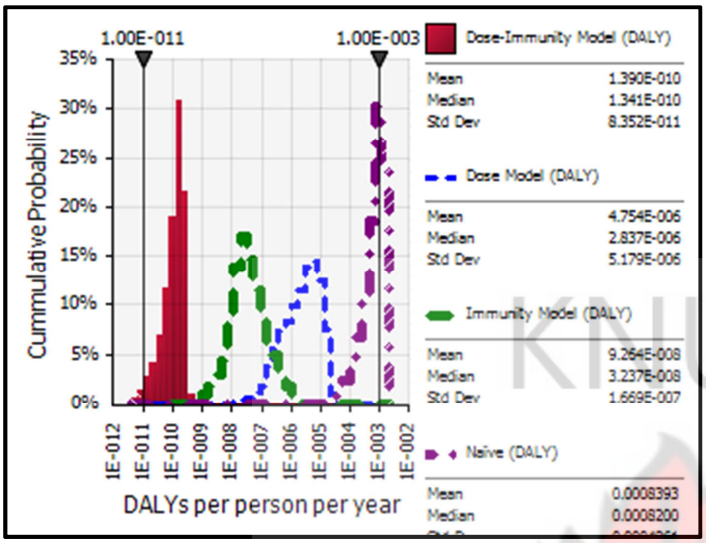




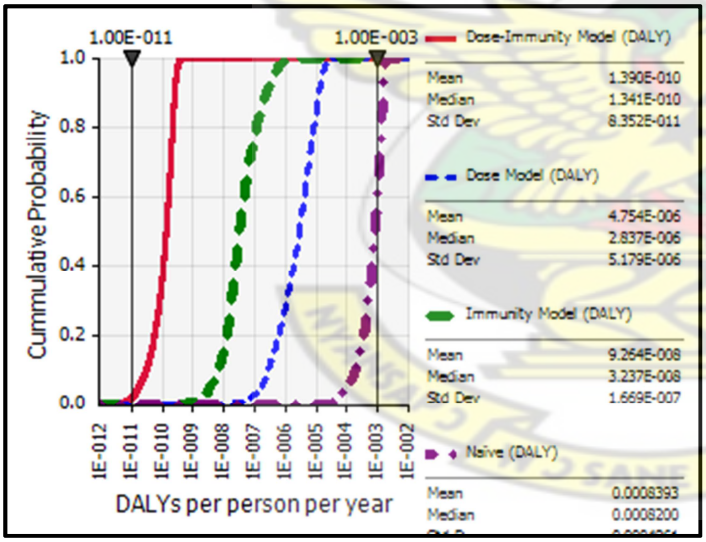




PROBABILITY GRAPH FOR DR MODELS



CUMMULATIVE GRAPH FOR DR MODELS



Summary Statistics for Genogroup II Type 4

Dose-Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.20E-12	**5%**	1.73E-11
Maximum	4.20E-10	**10%**	3.02E-11
Mean	1.39E-10	**15%**	4.28E-11
Std Dev	8.40E-11	**20%**	5.57E-11
Variance	7.05068E-21	**25%**	6.97E-11
Skewness	0.371280581	**30%**	8.28E-11
Kurtosis	2.42294695	**35%**	9.55E-11
Median	1.34E-10	**40%**	1.08E-10
Mode	1.61E-10	**45%**	1.21E-10
Left X	1.73E-11	**50%**	1.34E-10
Left P	5%	**55%**	1.47E-10
Right X	2.87E-10	**60%**	1.60E-10
Right P	95%	**65%**	1.73E-10
Diff X	2.69E-10	**70%**	1.86E-10
Diff P	90%	**75%**	1.99E-10
#Errors	0	**80%**	2.15E-10
Filter Min	Off	**85%**	2.32E-10
Filter Max	Off	**90%**	2.54E-10
#Filtered	0	**95%**	2.87E-10

Dose Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	5.81E-09	**5%**	1.54E-07
Maximum	2.97E-05	**10%**	2.86E-07
Mean	4.70E-06	**15%**	4.33E-07
Std Dev	5.03E-06	**20%**	6.39E-07
Variance	2.52973E-11	**25%**	8.69E-07
Skewness	1.498333705	**30%**	1.13E-06
Kurtosis	5.001455157	**35%**	1.47E-06
Median	2.81E-06	**40%**	1.86E-06
Mode	1.18E-07	**45%**	2.30E-06
Left X	1.54E-07	**50%**	2.81E-06
Left P	5%	**55%**	3.42E-06
Right X	1.55E-05	**60%**	4.16E-06
Right P	95%	**65%**	4.98E-06
Diff X	1.53E-05	**70%**	5.84E-06
Diff P	90%	**75%**	6.97E-06
#Errors	0	**80%**	8.29E-06
Filter Min	Off	**85%**	1.01E-05
Filter Max	Off	**90%**	1.21E-05
#Filtered	0	**95%**	1.55E-05

Immunity Model / Incidence per person per year			
Statistics		**Percentage**	
Minimum	3.47E-10	**5%**	3.67E-09
Maximum	1.55E-06	**10%**	6.53E-09
Mean	9.25E-08	**15%**	9.48E-09
Std Dev	1.64E-07	**20%**	1.23E-08
Variance	2.68224E-14	**25%**	1.52E-08
Skewness	3.613497548	**30%**	1.78E-08
Kurtosis	18.74992932	**35%**	2.08E-08
Median	3.23E-08	**40%**	2.39E-08
Mode	1.76E-08	**45%**	2.79E-08
Left X	3.67E-09	**50%**	3.23E-08
Left P	5%	**55%**	3.76E-08
Right X	4.18E-07	**60%**	4.48E-08
Right P	95%	**65%**	5.45E-08
Diff X	4.15E-07	**70%**	6.68E-08
Diff P	90%	**75%**	8.65E-08
#Errors	0	**80%**	1.16E-07
Filter Min	Off	**85%**	1.58E-07
Filter Max	Off	**90%**	2.44E-07
#Filtered	0	**95%**	4.18E-07

Naive/ Incidence per person per year			
Statistics		**Percentage**	
Minimum	2.49E-05	**5%**	1.07E-04
Maximum	2.02E-03	**10%**	1.86E-04
Mean	8.40E-04	**15%**	2.65E-04
Std Dev	4.87E-04	**20%**	3.42E-04
Variance	2.37332E-07	**25%**	4.24E-04
Skewness	0.171673405	**30%**	5.01E-04
Kurtosis	2.021870674	**35%**	5.83E-04
Median	8.25E-04	**40%**	6.66E-04
Mode	2.86E-04	**45%**	7.44E-04
Left X	1.07E-04	**50%**	8.25E-04
Left P	5%	**55%**	9.10E-04
Right X	1.65E-03	**60%**	9.82E-04
Right P	95%	**65%**	1.06E-03
Diff X	1.55E-03	**70%**	1.14E-03
Diff P	90%	**75%**	1.22E-03
#Errors	0	**80%**	1.30E-03
Filter Min	Off	**85%**	1.40E-03
Filter Max	Off	**90%**	1.52E-03
#Filtered	0	**95%**	1.65E-03

APPENDIX C1: Reference Chart of Compliance of Mean Effluent Discharge of BOD_5, TN and TSS for 50mg/L and Trichloroethylene. Benzene for 50 µg/l

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	50.00	50.25	50.99	52.20	53.85	55.90	58.31	61.03	64.03	67.27	70.71	78.10	86.02	94.34	102.96	111.80	134.63	158.11	182.00	206.16
60%	50.00	49.00	48.50	48.46	48.85	49.60	50.68	52.02	53.59	55.36	57.28	61.50	66.09	70.94	75.96	81.11	94.30	107.71	121.19	134.66
70%	50.00	47.69	45.95	44.75	43.99	43.62	43.58	43.81	44.26	44.89	45.67	47.57	49.78	52.21	54.78	57.44	64.30	71.28	78.26	85.19
80%	50.00	46.20	43.16	40.77	38.93	37.56	36.56	35.86	35.42	35.17	35.08	35.26	35.78	36.53	37.42	38.42	41.16	44.06	47.02	49.97
90%	50.00	44.22	39.56	35.83	32.86	30.51	28.64	27.16	25.99	25.06	24.32	23.27	22.63	22.25	22.05	21.98	22.16	22.60	23.18	23.83
92%	50.00	43.68	38.61	34.56	31.34	28.79	26.75	25.13	23.84	22.79	21.95	20.72	19.91	19.37	19.02	18.81	18.63	18.75	19.02	19.37
95%	50.00	42.64	36.81	32.21	28.57	25.70	23.42	21.60	20.13	18.95	17.98	16.52	15.50	14.78	14.26	13.87	13.29	13.03	12.93	12.93
98%	50.00	40.94	33.95	28.56	24.41	21.19	18.67	16.68	15.10	13.83	12.79	11.22	10.12	9.32	8.72	8.26	7.47	7.00	6.70	6.50
99%	50.00	39.84	32.17	26.37	21.98	18.63	16.05	14.05	12.47	11.21	10.20	8.68	7.63	6.86	6.29	5.85	5.10	4.64	4.33	4.11
99.9%	50.00	36.92	27.65	21.07	16.38	12.99	10.51	8.67	7.29	6.22	5.40	4.22	3.44	2.90	2.51	2.22	1.74	1.45	1.27	1.14

APPENDIX C2: Reference Chart of Compliance of Mean Effluent Discharge of TP for 2.0mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	2.00	2.01	2.04	2.09	2.15	2.24	2.33	2.44	2.56	2.69	2.83	3.12	3.44	3.77	4.12	4.47	5.39	6.32	7.28	8.25
60%	2.00	1.96	1.94	1.94	1.95	1.98	2.03	2.08	2.14	2.21	2.29	2.46	2.64	2.84	3.04	3.24	3.77	4.31	4.85	5.39
70%	2.00	1.91	1.84	1.79	1.76	1.74	1.74	1.75	1.77	1.80	1.83	1.90	1.99	2.09	2.19	2.30	2.57	2.85	3.13	3.41
80%	2.00	1.85	1.73	1.63	1.56	1.50	1.46	1.43	1.42	1.41	1.40	1.41	1.43	1.46	1.50	1.54	1.65	1.76	1.88	2.00
90%	2.00	1.77	1.58	1.43	1.31	1.22	1.15	1.09	1.04	1.00	0.97	0.93	0.91	0.89	0.88	0.88	0.89	0.90	0.93	0.95
92%	2.00	1.75	1.54	1.38	1.25	1.15	1.07	1.01	0.95	0.91	0.88	0.83	0.80	0.77	0.76	0.75	0.75	0.75	0.76	0.77
95%	2.00	1.71	1.47	1.29	1.14	1.03	0.94	0.86	0.81	0.76	0.72	0.66	0.62	0.59	0.57	0.55	0.53	0.52	0.52	0.52
98%	2.00	1.64	1.36	1.14	0.98	0.85	0.75	0.67	0.60	0.55	0.51	0.45	0.40	0.37	0.35	0.33	0.30	0.28	0.27	0.26
99%	2.00	1.59	1.29	1.05	0.88	0.75	0.64	0.56	0.50	0.45	0.41	0.35	0.31	0.27	0.25	0.23	0.20	0.19	0.17	0.16
99.9%	2.00	1.48	1.11	0.84	0.66	0.52	0.42	0.35	0.29	0.25	0.22	0.17	0.14	0.12	0.10	0.09	0.07	0.06	0.05	0.05

APPENDIX C3: Reference Chart of Compliance of Mean Effluent Discharge of Soluble Arsenic. Lead and Silver of 0.1mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	0.10	0.10	0.10	0.10	0.11	0.11	0.12	0.12	0.13	0.13	0.14	0.16	0.17	0.19	0.21	0.22	0.27	0.32	0.36	0.41
60%	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.11	0.11	0.11	0.12	0.13	0.14	0.15	0.16	0.19	0.22	0.24	0.27
70%	0.10	0.10	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.10	0.10	0.10	0.11	0.11	0.13	0.14	0.16	0.17
80%	0.10	0.09	0.09	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.09	0.09	0.10
90%	0.10	0.09	0.08	0.07	0.07	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.05	0.05
92%	0.10	0.09	0.08	0.07	0.06	0.06	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
95%	0.10	0.09	0.07	0.06	0.06	0.05	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
98%	0.10	0.08	0.07	0.06	0.05	0.04	0.04	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01
99%	0.10	0.08	0.06	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01
99.9%	0.10	0.07	0.06	0.04	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00

APPENDIX C4: Reference Chart of Compliance of Mean Effluent Discharge for Temperature of 30^0C

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	30.00	30.15	30.59	31.32	32.31	33.54	34.99	36.62	38.42	40.36	42.43	46.86	51.61	56.60	61.77	67.08	80.78	94.87	109.20	123.69
60%	30.00	29.40	29.10	29.08	29.31	29.76	30.41	31.21	32.16	33.21	34.37	36.90	39.66	42.56	45.58	48.66	56.58	64.62	72.71	80.80
70%	30.00	28.61	27.57	26.85	26.39	26.17	26.15	26.29	26.56	26.94	27.40	28.54	29.87	31.33	32.87	34.46	38.58	42.77	46.96	51.12
80%	30.00	27.72	25.90	24.46	23.36	22.53	21.93	21.52	21.25	21.10	21.05	21.16	21.47	21.92	22.45	23.05	24.69	26.44	28.21	29.98
90%	30.00	26.53	23.73	21.50	19.72	18.31	17.19	16.30	15.59	15.03	14.59	13.96	13.58	13.35	13.23	13.19	13.29	13.56	13.91	14.30
92%	30.00	26.21	23.16	20.74	18.81	17.27	16.05	15.08	14.30	13.68	13.17	12.43	11.94	11.62	11.41	11.29	11.18	11.25	11.41	11.62
95%	30.00	25.59	22.09	19.32	17.14	15.42	14.05	12.96	12.08	11.37	10.79	9.91	9.30	8.87	8.55	8.32	7.98	7.82	7.76	7.76
98%	30.00	24.56	20.37	17.14	14.64	12.71	11.20	10.01	9.06	8.30	7.67	6.73	6.07	5.59	5.23	4.95	4.48	4.20	4.02	3.90
99%	30.00	23.91	19.30	15.82	13.19	11.18	9.63	8.43	7.48	6.73	6.12	5.21	4.58	4.12	3.77	3.51	3.06	2.78	2.60	2.47
99.9%	30.00	22.15	16.59	12.64	9.83	7.79	6.31	5.20	4.37	3.73	3.24	2.53	2.06	1.74	1.51	1.33	1.04	0.87	0.76	0.68

APPENDIX C5: Reference Chart of Compliance of Mean Effluent Discharge ofTC of 400 MPN/100ml

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	400.00	402.00	407.92	417.61	430.81	447.21	466.48	488.26	512.25	538.14	565.69	624.82	688.19	754.72	823.65	894.43	1077.03	1264.91	1456.02	1649.24
60%	400.00	391.98	387.99	387.72	390.80	396.84	405.42	416.17	428.75	442.86	458.24	492.02	528.74	567.51	607.69	648.86	754.36	861.65	969.49	1077.30
70%	400.00	381.48	367.64	357.96	351.92	348.99	348.66	350.49	354.09	359.15	365.38	380.55	398.28	417.70	438.23	459.50	514.42	570.27	626.12	681.56
80%	400.00	369.61	345.27	326.16	311.47	300.45	292.45	286.90	283.33	281.34	280.63	282.09	286.27	292.23	299.39	307.35	329.27	352.51	376.14	399.74
90%	400.00	353.74	316.46	286.63	262.90	244.07	229.14	217.30	207.91	200.46	194.55	186.17	181.01	177.99	176.43	175.88	177.25	180.80	185.43	190.60
92%	400.00	349.43	308.84	276.47	250.73	230.29	214.03	201.06	190.68	182.34	175.62	165.75	159.24	154.96	152.18	150.47	149.08	150.02	152.16	154.96
95%	400.00	341.16	294.50	257.66	228.59	205.61	187.36	172.79	161.06	151.57	143.81	132.14	124.02	118.24	114.05	110.97	106.34	104.23	103.46	103.46
98%	400.00	327.52	271.59	228.51	195.27	169.49	149.34	133.46	120.80	110.61	102.31	89.80	80.99	74.58	69.76	66.05	59.80	56.03	53.61	51.97
99%	400.00	318.75	257.35	210.97	175.84	149.05	128.43	112.39	99.77	89.70	81.57	69.45	61.01	54.89	50.31	46.77	40.78	37.09	34.62	32.88
99.9%	400.00	295.36	221.21	168.59	131.01	103.89	84.08	69.38	58.29	49.80	43.18	33.75	27.53	23.21	20.08	17.74	13.91	11.63	10.14	9.09

APPENDIX C6: Reference Chart of Compliance of Mean Effluent Discharge ofTotal Arsenic. Total Chromium and Nickel of 0.5mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	0.50	0.50	0.51	0.52	0.54	0.56	0.58	0.61	0.64	0.67	0.71	0.78	0.86	0.94	1.03	1.12	1.35	1.58	1.82	2.06
60%	0.50	0.49	0.48	0.48	0.49	0.50	0.51	0.52	0.54	0.55	0.57	0.62	0.66	0.71	0.76	0.81	0.94	1.08	1.21	1.35
70%	0.50	0.48	0.46	0.45	0.44	0.44	0.44	0.44	0.44	0.45	0.46	0.48	0.50	0.52	0.55	0.57	0.64	0.71	0.78	0.85
80%	0.50	0.46	0.43	0.41	0.39	0.38	0.37	0.36	0.35	0.35	0.35	0.35	0.36	0.37	0.37	0.38	0.41	0.44	0.47	0.50
90%	0.50	0.44	0.40	0.36	0.33	0.31	0.29	0.27	0.26	0.25	0.24	0.23	0.23	0.22	0.22	0.22	0.22	0.23	0.23	0.24
92%	0.50	0.44	0.39	0.35	0.31	0.29	0.27	0.25	0.24	0.23	0.22	0.21	0.20	0.19	0.19	0.19	0.19	0.19	0.19	0.19
95%	0.50	0.43	0.37	0.32	0.29	0.26	0.23	0.22	0.20	0.19	0.18	0.17	0.16	0.15	0.14	0.14	0.13	0.13	0.13	0.13
98%	0.50	0.41	0.34	0.29	0.24	0.21	0.19	0.17	0.15	0.14	0.13	0.11	0.10	0.09	0.09	0.08	0.07	0.07	0.07	0.06
99%	0.50	0.40	0.32	0.26	0.22	0.19	0.16	0.14	0.12	0.11	0.10	0.09	0.08	0.07	0.06	0.06	0.05	0.05	0.04	0.04
99.9%	0.50	0.37	0.28	0.21	0.16	0.13	0.11	0.09	0.07	0.06	0.05	0.04	0.03	0.03	0.03	0.02	0.02	0.01	0.01	0.01

APPENDIX C7: Reference Chart of Compliance of Mean Effluent Discharge for COD. Chloride and Total residual chlorine of 250 mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	250.00	251.25	254.95	261.01	269.26	279.51	291.55	305.16	320.16	336.34	353.55	390.51	430.12	471.70	514.78	559.02	673.15	790.57	910.01	1030.78
60%	250.00	244.99	242.49	242.32	244.25	248.02	253.39	260.10	267.97	276.79	286.40	307.51	330.46	354.69	379.81	405.54	471.48	538.53	605.93	673.32
70%	250.00	238.43	229.77	223.73	219.95	218.12	217.91	219.05	221.31	224.47	228.36	237.85	248.92	261.06	273.90	287.19	321.51	356.42	391.32	425.97
80%	250.00	231.01	215.79	203.85	194.67	187.78	182.78	179.31	177.08	175.84	175.39	176.31	178.92	182.65	187.12	192.09	205.79	220.32	235.09	249.83
90%	250.00	221.09	197.79	179.15	164.31	152.54	143.21	135.81	129.95	125.29	121.60	116.36	113.13	111.25	110.27	109.92	110.78	113.00	115.89	119.13
92%	250.00	218.39	193.03	172.79	156.71	143.93	133.77	125.66	119.18	113.96	109.76	103.60	99.53	96.85	95.11	94.04	93.17	93.76	95.10	96.85
95%	250.00	213.22	184.07	161.04	142.87	128.50	117.10	107.99	100.67	94.73	89.88	82.59	77.51	73.90	71.28	69.36	66.46	65.14	64.66	64.66
98%	250.00	204.70	169.74	142.82	122.04	105.93	93.34	83.41	75.50	69.13	63.94	56.12	50.62	46.61	43.60	41.28	37.37	35.02	33.50	32.48
99%	250.00	199.22	160.84	131.86	109.90	93.16	80.27	70.25	62.35	56.06	50.98	43.41	38.13	34.31	31.44	29.23	25.49	23.18	21.64	20.55
99.9%	250.00	184.60	138.26	105.37	81.88	64.93	52.55	43.36	36.43	31.12	26.99	21.10	17.20	14.50	12.55	11.09	8.70	7.27	6.34	5.68

APPENDIX C8: Reference Chart of Compliance of Mean Effluent Discharge of Conductivity of 1500$\mu S/cm$

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	1500.00	1507.48	1529.71	1566.05	1615.55	1677.05	1749.29	1830.98	1920.94	2018.04	2121.32	2343.07	2580.70	2830.19	3088.69	3354.10	4038.87	4743.42	5460.08	6184.66
60%	1500.00	1469.91	1454.95	1453.95	1465.51	1488.14	1520.31	1560.63	1607.80	1660.72	1718.42	1845.07	1982.79	2128.16	2278.84	2433.23	2828.86	3231.18	3635.59	4039.89
70%	1500.00	1430.57	1378.65	1342.37	1319.72	1308.70	1307.46	1314.32	1327.84	1346.80	1370.19	1427.07	1493.55	1566.36	1643.37	1723.14	1929.07	2138.51	2347.93	2555.85
80%	1500.00	1386.04	1294.76	1223.08	1168.00	1126.70	1096.70	1075.88	1062.47	1055.02	1052.35	1057.84	1073.50	1095.87	1122.71	1152.57	1234.75	1321.92	1410.53	1499.01
90%	1500.00	1326.52	1186.71	1074.88	985.88	915.25	859.26	814.88	779.68	751.74	729.57	698.14	678.80	667.47	661.60	659.54	664.70	678.02	695.36	714.76
92%	1500.00	1310.34	1158.15	1036.76	940.25	863.59	802.61	753.98	715.06	683.79	658.56	621.58	597.16	581.08	570.68	564.25	559.03	562.58	570.61	581.09
95%	1500.00	1279.35	1104.39	966.23	857.22	771.03	702.60	647.95	603.99	568.37	539.28	495.51	465.07	443.39	427.68	416.14	398.78	390.86	387.96	387.97
98%	1500.00	1228.20	1018.47	856.91	732.25	635.57	560.03	500.46	453.00	414.77	383.65	336.74	303.72	279.66	261.60	247.69	224.25	210.13	201.03	194.90
99%	1500.00	1195.33	965.05	791.15	659.40	558.94	481.62	421.48	374.12	336.37	305.90	260.45	228.78	205.84	188.65	175.40	152.92	139.07	129.83	123.31
99.9%	1500.00	1107.61	829.55	632.20	491.27	389.61	315.30	260.16	218.60	186.74	161.94	126.58	103.22	87.02	75.31	66.54	52.18	43.63	38.02	34.08

APPENDIX C9: Reference Chart of Compliance of Mean Effluent Discharge of TDS of 1000 mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	1000.00	1004.99	1019.80	1044.03	1077.03	1118.03	1166.19	1220.66	1280.62	1345.36	1414.21	1562.05	1720.47	1886.80	2059.13	2236.07	2692.58	3162.28	3640.05	4123.11
60%	1000.00	979.94	969.97	969.30	977.01	992.09	1013.54	1040.42	1071.87	1107.14	1145.61	1230.05	1321.86	1418.77	1519.23	1622.15	1885.91	2154.12	2423.73	2693.26
70%	1000.00	953.71	919.10	894.91	879.81	872.47	871.64	876.22	885.23	897.87	913.46	951.38	995.70	1044.24	1095.58	1148.76	1286.05	1425.67	1565.29	1703.90
80%	1000.00	924.03	863.17	815.39	778.67	751.13	731.13	717.25	708.31	703.34	701.57	705.23	715.67	730.58	748.47	768.38	823.17	881.28	940.35	999.34
90%	1000.00	884.35	791.14	716.59	657.25	610.17	572.84	543.25	519.78	501.16	486.38	465.43	452.53	444.98	441.07	439.70	443.13	452.01	463.57	476.51
92%	1000.00	873.56	772.10	691.17	626.84	575.73	535.07	502.65	476.71	455.86	439.04	414.38	398.11	387.39	380.45	376.17	372.69	375.05	380.41	387.39
95%	1000.00	852.90	736.26	644.15	571.48	514.02	468.40	431.97	402.66	378.92	359.52	330.34	310.04	295.59	285.12	277.43	265.85	260.57	258.64	258.65
98%	1000.00	818.80	678.98	571.28	488.17	423.71	373.35	333.64	302.00	276.52	255.77	224.49	202.48	186.44	174.40	165.12	149.50	140.09	134.02	129.94
99%	1000.00	796.88	643.37	527.43	439.60	372.62	321.08	280.99	249.42	224.25	203.94	173.63	152.52	137.22	125.77	116.94	101.95	92.71	86.55	82.20
99.9%	1000.00	738.41	553.03	421.47	327.51	259.74	210.20	173.44	145.73	124.49	107.96	84.38	68.82	58.02	50.21	44.36	34.78	29.08	25.35	22.72

APPENDIX C10: Reference Chart of Compliance of Mean Effluent Discharge of DO. Total Cyanide. Phenol . Selenium and Ammonia of 1.0mg/L

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	1.00	1.00	1.02	1.04	1.08	1.12	1.17	1.22	1.28	1.35	1.41	1.56	1.72	1.89	2.06	2.24	2.69	3.16	3.64	4.12
60%	1.00	0.98	0.97	0.97	0.98	0.99	1.01	1.04	1.07	1.11	1.15	1.23	1.32	1.42	1.52	1.62	1.89	2.15	2.42	2.69
70%	1.00	0.95	0.92	0.89	0.88	0.87	0.87	0.88	0.89	0.90	0.91	0.95	1.00	1.04	1.10	1.15	1.29	1.43	1.57	1.70
80%	1.00	0.92	0.86	0.82	0.78	0.75	0.73	0.72	0.71	0.70	0.70	0.71	0.72	0.73	0.75	0.77	0.82	0.88	0.94	1.00
90%	1.00	0.88	0.79	0.72	0.66	0.61	0.57	0.54	0.52	0.50	0.49	0.47	0.45	0.44	0.44	0.44	0.44	0.45	0.46	0.48
92%	1.00	0.87	0.77	0.69	0.63	0.58	0.54	0.50	0.48	0.46	0.44	0.41	0.40	0.39	0.38	0.38	0.37	0.38	0.38	0.39
95%	1.00	0.85	0.74	0.64	0.57	0.51	0.47	0.43	0.40	0.38	0.36	0.33	0.31	0.30	0.29	0.28	0.27	0.26	0.26	0.26
98%	1.00	0.82	0.68	0.57	0.49	0.42	0.37	0.33	0.30	0.28	0.26	0.22	0.20	0.19	0.17	0.17	0.15	0.14	0.13	0.13
99%	1.00	0.80	0.64	0.53	0.44	0.37	0.32	0.28	0.25	0.22	0.20	0.17	0.15	0.14	0.13	0.12	0.10	0.09	0.09	0.08
99.9%	1.00	0.74	0.55	0.42	0.33	0.26	0.21	0.17	0.15	0.12	0.11	0.08	0.07	0.06	0.05	0.04	0.03	0.03	0.03	0.02

APPENDIX C11: Reference Chart of Compliance of Mean Effluent Discharge of Turbidity of 75 NTU

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	75.00	75.37	76.49	78.30	80.78	83.85	87.46	91.55	96.05	100.90	106.07	117.15	129.03	141.51	154.43	167.71	201.94	237.17	273.00	309.23
60%	75.00	73.50	72.75	72.70	73.28	74.41	76.02	78.03	80.39	83.04	85.92	92.25	99.14	106.41	113.94	121.66	141.44	161.56	181.78	201.99
70%	75.00	71.53	68.93	67.12	65.99	65.44	65.37	65.72	66.39	67.34	68.51	71.35	74.68	78.32	82.17	86.16	96.45	106.93	117.40	127.79
80%	75.00	69.30	64.74	61.15	58.40	56.33	54.84	53.79	53.12	52.75	52.62	52.89	53.68	54.79	56.14	57.63	61.74	66.10	70.53	74.95
90%	75.00	66.33	59.34	53.74	49.29	45.76	42.96	40.74	38.98	37.59	36.48	34.91	33.94	33.37	33.08	32.98	33.23	33.90	34.77	35.74
92%	75.00	65.52	57.91	51.84	47.01	43.18	40.13	37.70	35.75	34.19	32.93	31.08	29.86	29.05	28.53	28.21	27.95	28.13	28.53	29.05
95%	75.00	63.97	55.22	48.31	42.86	38.55	35.13	32.40	30.20	28.42	26.96	24.78	23.25	22.17	21.38	20.81	19.94	19.54	19.40	19.40
98%	75.00	61.41	50.92	42.85	36.61	31.78	28.00	25.02	22.65	20.74	19.18	16.84	15.19	13.98	13.08	12.38	11.21	10.51	10.05	9.75
99%	75.00	59.77	48.25	39.56	32.97	27.95	24.08	21.07	18.71	16.82	15.30	13.02	11.44	10.29	9.43	8.77	7.65	6.95	6.49	6.17
99.9%	75.00	55.38	41.48	31.61	24.56	19.48	15.76	13.01	10.93	9.34	8.10	6.33	5.16	4.35	3.77	3.33	2.61	2.18	1.90	1.70

APPENDIX C12: Reference Chart of Compliance of Mean Effluent Discharge of E.coli of 10 MPN/100ml

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	10.00	10.05	10.20	10.44	10.77	11.18	11.66	12.21	12.81	13.45	14.14	15.62	17.20	18.87	20.59	22.36	26.93	31.62	36.40	41.23
60%	10.00	9.80	9.70	9.69	9.77	9.92	10.14	10.40	10.72	11.07	11.46	12.30	13.22	14.19	15.19	16.22	18.86	21.54	24.24	26.93
70%	10.00	9.54	9.19	8.95	8.80	8.72	8.72	8.76	8.85	8.98	9.13	9.51	9.96	10.44	10.96	11.49	12.86	14.26	15.65	17.04
80%	10.00	9.24	8.63	8.15	7.79	7.51	7.31	7.17	7.08	7.03	7.02	7.05	7.16	7.31	7.48	7.68	8.23	8.81	9.40	9.99
90%	10.00	8.84	7.91	7.17	6.57	6.10	5.73	5.43	5.20	5.01	4.86	4.65	4.53	4.45	4.41	4.40	4.43	4.52	4.64	4.77
92%	10.00	8.74	7.72	6.91	6.27	5.76	5.35	5.03	4.77	4.56	4.39	4.14	3.98	3.87	3.80	3.76	3.73	3.75	3.80	3.87
95%	10.00	8.53	7.36	6.44	5.71	5.14	4.68	4.32	4.03	3.79	3.60	3.30	3.10	2.96	2.85	2.77	2.66	2.61	2.59	2.59
98%	10.00	8.19	6.79	5.71	4.88	4.24	3.73	3.34	3.02	2.77	2.56	2.24	2.02	1.86	1.74	1.65	1.49	1.40	1.34	1.30
99%	10.00	7.97	6.43	5.27	4.40	3.73	3.21	2.81	2.49	2.24	2.04	1.74	1.53	1.37	1.26	1.17	1.02	0.93	0.87	0.82
99.9%	10.00	7.38	5.53	4.21	3.28	2.60	2.10	1.73	1.46	1.24	1.08	0.84	0.69	0.58	0.50	0.44	0.35	0.29	0.25	0.23

APPENDIX C13: Reference Chart of Compliance of Mean Effluent Discharge for pH
Lower bound for pH for 6 (interval equation 21)

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	6.00	6.03	6.12	6.26	6.46	6.71	7.00	7.32	7.68	8.07	8.49	9.37	10.32	11.32	12.35	13.42	16.16	18.97	21.84	24.74
60%	6.00	5.88	5.82	5.82	5.86	5.95	6.08	6.24	6.43	6.64	6.87	7.38	7.93	8.51	9.12	9.73	11.32	12.92	14.54	16.16
70%	6.00	5.72	5.51	5.37	5.28	5.23	5.23	5.26	5.31	5.39	5.48	5.71	5.97	6.27	6.57	6.89	7.72	8.55	9.39	10.22
80%	6.00	5.54	5.18	4.89	4.67	4.51	4.39	4.30	4.25	4.22	4.21	4.23	4.29	4.38	4.49	4.61	4.94	5.29	5.64	6.00
90%	6.00	5.31	4.75	4.30	3.94	3.66	3.44	3.26	3.12	3.01	2.92	2.79	2.72	2.67	2.65	2.64	2.66	2.71	2.78	2.86
92%	6.00	5.24	4.63	4.15	3.76	3.45	3.21	3.02	2.86	2.74	2.63	2.49	2.39	2.32	2.28	2.26	2.24	2.25	2.28	2.32
95%	6.00	5.12	4.42	3.86	3.43	3.08	2.81	2.59	2.42	2.27	2.16	1.98	1.86	1.77	1.71	1.66	1.60	1.56	1.55	1.55
98%	6.00	4.91	4.07	3.43	2.93	2.54	2.24	2.00	1.81	1.66	1.53	1.35	1.21	1.12	1.05	0.99	0.90	0.84	0.80	0.78
99%	6.00	4.78	3.86	3.16	2.64	2.24	1.93	1.69	1.50	1.35	1.22	1.04	0.92	0.82	0.75	0.70	0.61	0.56	0.52	0.49
99.9%	6.00	4.43	3.32	2.53	1.97	1.56	1.26	1.04	0.87	0.75	0.65	0.51	0.41	0.35	0.30	0.27	0.21	0.17	0.15	0.14

Upper bound for pH for 9 (interval equation 21)

	Coefficient of Variation																			
COR	**0**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**	**1.2**	**1.4**	**1.6**	**1.8**	**2**	**2.5**	**3**	**3.5**	**4**
50%	9.00	9.04	9.18	9.40	9.69	10.06	10.50	10.99	11.53	12.11	12.73	14.06	15.48	16.98	18.53	20.12	24.23	28.46	32.76	37.11
60%	9.00	8.82	8.73	8.72	8.79	8.93	9.12	9.36	9.65	9.96	10.31	11.07	11.90	12.77	13.67	14.60	16.97	19.39	21.81	24.24
70%	9.00	8.58	8.27	8.05	7.92	7.85	7.84	7.89	7.97	8.08	8.22	8.56	8.96	9.40	9.86	10.34	11.57	12.83	14.09	15.34
80%	9.00	8.32	7.77	7.34	7.01	6.76	6.58	6.46	6.37	6.33	6.31	6.35	6.44	6.58	6.74	6.92	7.41	7.93	8.46	8.99
90%	9.00	7.96	7.12	6.45	5.92	5.49	5.16	4.89	4.68	4.51	4.38	4.19	4.07	4.00	3.97	3.96	3.99	4.07	4.17	4.29
92%	9.00	7.86	6.95	6.22	5.64	5.18	4.82	4.52	4.29	4.10	3.95	3.73	3.58	3.49	3.42	3.39	3.35	3.38	3.42	3.49
95%	9.00	7.68	6.63	5.80	5.14	4.63	4.22	3.89	3.62	3.41	3.24	2.97	2.79	2.66	2.57	2.50	2.39	2.35	2.33	2.33
98%	9.00	7.37	6.11	5.14	4.39	3.81	3.36	3.00	2.72	2.49	2.30	2.02	1.82	1.68	1.57	1.49	1.35	1.26	1.21	1.17
99%	9.00	7.17	5.79	4.75	3.96	3.35	2.89	2.53	2.24	2.02	1.84	1.56	1.37	1.24	1.13	1.05	0.92	0.83	0.78	0.74
99.9%	9.00	6.65	4.98	3.79	2.95	2.34	1.89	1.56	1.31	1.12	0.97	0.76	0.62	0.52	0.45	0.40	0.31	0.26	0.23	0.20