UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

The System Dynamics approach as a modelling tool for health care.

by

PETRUS MILLAR NIENABER
26061041

A dissertation in partial fulfilment of the requirements for the degree of

MASTERS OF ENGINEERING (INDUSTRIAL ENGINEERING)

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND INFORMATION TECHNOLOGY

University of Pretoria

South Africa

September 2012

Abstract

The System Dynamics approach as a modelling tool for health care

by

PETRUS MILLAR NIENABER

Supervisor	:	Prof PS Kruger
Co-Supervisor	:	Prof VSS Yadavalli
Department	:	Industrial and Systems Engineering
University	:	University of Pretoria
Degree	:	Masters of Engineering (Industrial and Systems Engineering)

In this dissertation System Dynamics is used as a modelling approach to model health care systems to gain a better understanding of the system's behaviour. This improved understanding can be used to better manage the system and in turn will translate to improved health outcomes.

The characteristics of complex systems were reviewed to define a health system as a complex system. Four appropriate modelling approaches was studied that could be used to model complex systems. These modelling approaches included: Monte Carlo Simulation, Discrete Event Simulation, System Dynamics and Agent Based Modelling. System Dynamics was identified as being the most appropriate modelling methodology to be used for the framework. Before the framework was developed health system performance measurement was reviewed to further the understanding of health system measurement.

The framework was developed according to the insights gained from the previous reviews. Specifically the elements identification was customised to the health care environment based on available health indicators. The framework was applied in a case study where a section of the South Africa health care system was modelled to focus interventions for human immunodeficiency virus (HIV). The outcomes of the case studies delivered an increased understanding of the system behaviour and also showed appropriates of the framework.

Keywords: Health System, Modelling, System Performance, System Dynamics, Health Indicators, Modelling Framework.

Acknowledgments

Firstly I wish to thank my wife, Izel for her continuous support and motivation that was essential for me to complete this dissertation.

I would further like to express my gratitude to my supervisors Prof PS Kruger and Prof VSS Yadavalli for their motivation and confidence.

I am also grateful for BroadReach Healthcare for their support that enabled me to complete this dissertation, especially my colleagues for sharing their insights of the health system's operation.

Finally I would like to express a special thank you to my family and friends for providing inputs and continuous support.

CONTENTS

List of Figures

List of Tables

List of Graphs

List of Abbreviations

ART	Anti-retroviral Treatment
BRHC	Broadreach Healthcare
CD4	Cluster of differentiation 4
DALY	Disability adjusted life years
DES	Discrete event Simulation
DHIS	District health information system
Dmnl	Dimensionless
ES	Engineering System
HIV	Human immunodeficiency virus
INF	Infections
INTEG	Integration
IOM	Institute of Medicine
MDG	Millennium Development Goals
MSC	Monte Carlo Simulation
NHI	National Health Insurance
NR	Number
NSP	National Strategic Plan
OECD	Organisation for Economic and Co-operation and Development
PERS	Persons
PHC	Primary health care
PMTCT	Prevention of mother to child transmission
Rx	Treatment
SD	System Dynamics
WHO	World Health Organisation

Chapter 1

INTRODUCTION

In this dissertation System Dynamics is used as an approach to model health care systems to gain a better understanding of the system's behaviour. This improved understanding can be used to better manage the system and in turn will translate to improved health outcomes.

The characteristics of complex systems were reviewed to define a health system as a complex system. Four modelling approaches were studied that could be used to model complex systems. System Dynamics was identified as being the most appropriate modelling methodology to be used for the framework. Before the framework was developed health system performance measurement was reviewed to further the understanding of health system measurement.

The framework was developed according to the insights gained from the previous reviews. Specifically the elements identification was customised to the health care environment based on available health indicators. A modelling guide was developed to make the framework accessible and practically implementable in the health care environment. The framework was applied through the modelling guide, in a case study where a section of the South African health care system was modelled, to focus interventions for HIV. The outcomes of the case studies delivered an increased understanding of the system behaviour and showed the appropriateness of the framework.

1.1.1 *HEALTH SYSTEM MANAGEMENT*

Research studies done on the clinical aspects, such as epidemiology and pharmacology, as well as operational aspects of the health care system are common, yet it can be stated that health care system performance is poor [1]. Research that answers the challenges of providing care at a system level is less common. Thus there is a need to improve the understanding of the health care environment, as a system level to empower decision makers and managers.

The management of health systems are further complicated by the structure of the system and the operating environment.

The South African public health care system is divided into three levels of care: primary, secondary and tertiary. Primary health care (PHC) is the first line of contact that a patient has with the formal health care system. The first contact point being Primary Health Centres, which are imbedded in the communities they serve, and should be accessible to the patients. Secondary health care facilities are usually regional or district hospitals with around 50 and more beds. Tertiary health care facilities are large hospitals with 100's of beds, including academic hospitals that only accept patients on referrals from secondary level facilities.

1.1.2 *HEALTH SYSTEMS PERFORMANCE*

South African health indicators remain poor for a middle income developing country [2]. This can be explained by the four epidemics; HIV/AIDS, chronic and non-communicable diseases, violence and injury. In addition to the high burden of disease the health system is unable to meet the health care needs for a number of reasons including: lack of integration of health care service providers, a hospital physician curative focus, low professional motivation and inadequate working relationship between public and private sector. In South Africa and in a number of other middle income developing countries the shortcomings of the health care systems can be seen in the burden of disease as well as the inability to achieve the Millennium Development Goals (MDG) for health.

The health system is under a lot of political pressure to improve; this is motivated by the 10-point plan as well as the National Strategic Plan (NSP) [3] and indicates how changes in the system are being pushed for implementation. The implementation of National Health Insurance (NHI) in South Africa is another indication of the need to re-engineer health systems to improve performance.

The reaction of a health system to intervention is unknown due to complexity of the health system. This reaction is further complicated by the dynamic environment that interventions are implemented in. To ensure the health system is strengthened by management interventions and not weakened by unintentional effects a careful analysis and deeper understanding of these implementations is needed.

1.2 Problem statement

This dissertation focuses on the development and use of a framework to model a health care system that can be used to aid decision makers in the management of health care systems, and in turn improve system performance. Furthermore, the application of the framework should be relevant to the health care environment and its specific challenges. To achieve this goal the following research question should be answered:

> *What would be a reasonable framework to aid decision makers to increase understanding of the health care system through a modelling approach, while being sensitive to the challenges specific to the health care environment, but be generic enough to be useful in various scenarios?*

1.3 Research design

To deal with the above mentioned research question the dissertation focuses on the development of a framework to aid in the building of a simulation model that can assist in managing of health systems, characterised by the challenges associated with complex system management. The framework is tailored to the needs of modelling a health care system, while balancing the level of detail with the capability to be used in multiple situations. Once the framework is developed it is also validatedby means of a case study. The framework is applied to the health care environment to help gain insight into the use of health care interventions.

> *"Health care organisations are large, complex and adaptive. They show both detailed complexity and important dynamic complexity. They have long processes which often cross organisational boundaries and have their own cultures" [1]*

Complex system modelling and analysis tools will be investigated to find an appropriate method of modelling health care system

1.4 Research methodology

The research was structured in the following five phases.

Orientation: A significant section of the orientation of the research has been completed through prior and current exposure to the health care environment. Operational and academic involvement in health care environment have identified areas of need, and highlighted how these needs can realistically be addressed. Preliminary literature studies on the application of modelling in health care environments further augment the orientation.

Problem statement: To formulate the research problem the project aims to address some of the challenges found in applying modelling techniques in the health care environment at a system level.

Information collection: To gain a clear understanding of the health system operation the work of world leaders in health system monitoring was reviewed. To further the understanding of health system operations, academic and operational publications, national and international reports as well as various relevant websites, were utilised. Information collection on modelling approaches as described in literature was used to identify the appropriate modelling approach that should be used in the framework.

Framework development: The development of the framework was based on the understanding of health system operation gained from the collection of information. The framework will not only be developed according to the formal suggestions, but also being sensitive to the practical constraints of the health care environment.

Implementation and monitoring: Testing of the framework will be done in the form of a case study. In the case study the framework will be implemented in a real life application in the health care environment. By implementing the framework, the usefulness of the framework will be assessed.

1.5 Outline of the dissertation

The dissertation is organised as follows: in Chapter 2 on page 5 the modelling of health systems is discussed. The chapter starts with defining a health care system as a complex system and an engineering system. With the health care system defined, potential approaches are reviewed that could be used for health systems. From the modelling approach review, System Dynamics is identified on page 29 as an appropriate modelling approach to use. Chapter 3 further examines health systems by studying the use of health indicators, and how indicators are used in health system performance measurement. The framework is developed in Chapter 4 on page 48 to answer the research question posed in Chapter 1. The development of the framework is based on the insights gained in the previous chapters. Along with the development of the framework a criteria is developed to measure the usefulness of the framework. To test the framework against the criteria, Chapter 5 includes a case study where the framework used in real life application and scrutinised according to the criteria. Chapter 6 concludes the dissertation with closing remarks and some suggestions for future work in the field of health system modelling.

Chapter 2

MODELLING OF A HEALTH CARE SYSTEM

2.1 CHAPTER OVERVIEW

There are various approaches to use for modelling of systems; each can be appropriately applied to successfully model a certain system. To choose the correct modelling approach the system that is modelled and the available tools to perform the modelling needs to be understood. By combining an understanding of the system and its characteristics with a clear purpose of the model an informed choice of modelling approach can be made.

With the research question clearly stated in Chapter 1, the next step is to study the system that is to be modelled. The aim of studying health care systems is to classify the type of system. The next step is to ascertain whether a health care system is a complex system or not. Following the definition of a health care system as a complex system, an attempt to classify a health care system as an engineering system will be made.

When a health care system has been defined clearly, and insight into its characteristics has been gained, a selection of modelling techniques that could be used for health system modelling will be reviewed. In lieu of the system type identified, four appropriate modelling approaches were reviewed in depth: Monte Carlo simulation, Discrete Event Simulation, System Dynamics and Agent Based modelling. From the revision of the four approaches, System Dynamics was identified as being the most appropriate approach to use for this case of modelling.

2.2 HEALTH CARE AS A COMPLEX SYSTEM

The world has evolved towards a state of greater complexity in the last two centuries. This change has had a profound effect; by creating massive, connected and multi-level systems, that defies analytical understanding and "*seems to have a life of their own*" [4].

With the increase in complexity of our environment, there is an increase in the unintuitive and unexpected responses and reactions of systems to change and interventions.

To ultimately understand the behaviour of systems it is necessary to first understand the fundamental characteristics of a complex system. These characteristics will be matched to the characteristics of a health system to classify the system. With the system classified, the modelling methodology for a given system classification can be identified and an appropriate modelling methodology can be chosen for the system being studied.

2.2.1 COMPLEX SYSTEM CHARACTERISTICS

Minai *et al.* suggest that complex systems have fundamental characteristics in common. Perhaps the most important characteristic shared by all complex systems is self-organisation. This self-organisation can be described as the spontaneous appearance of large-scale organisation through limited interaction among simple components. Another form of self-organisation refers to the act of successive changes or improvements made to previously implemented man-made systems [4].

Two fundamental insights arise from the concept of self-organisation within complex systems:

- Large scale order can be created through simple processes involving interactions operating locally on simple agents or components.
- Highly complex functional systems can only arise through evolutionary processes of selection in the context of actual tasks.

From the first insight the common assumption that cause and effect operates at the same level is challenged. The concept of non-linearity and the application thereof in complex systems is brought to light. The possibility that the simple processes within the system impact the system characteristics in a much greater way than expected, the concept of non-linearity is highlighted. A change in operation or characteristic of a simple process within the system can roll over to influence the complex system to a much greater (or smaller) extent than the change to the simple process.

From the second insight the fallacy of the system being the sum of its parts is addressed. A complex system is more than the sum of its parts. The characteristics of a complex system cannot fully be described by the summation of the characteristics of the components or basic processes that complete the system.

Complex systems emerge and function in complex and dynamic environments, and their characteristics reflect this reality [4].

2.2.2 Classification of complex systems

Magee and De Weck found that system classifications from the initial general system theory in the 1950's have been unable to differentiate successfully between engineering systems and other systems. Some of the system attributes used in earlier research as a basis for classification are listed below. The number of references approaches identified by Magee and De Weck that uses the attribute for each classification is displayed in the table:

Attribute	Number of references
Functional Type	5
Time dependence	3
Origin	3
Degree of complexity	3
Boundary	3
Branch of economy	1
Human control	1
Human wants	1
Ownership	1
Realm of existence	1
System states	1

Table 1: Attributes of references

It was found that only seven of the 11 attributes were useful in the characterisation of an engineering system from other complex systems, but not when classifying engineering systems [5].

The purpose of the classification scheme proposed by Magee and De Weck was to differentiate between an engineering system and other systems such as natural systems, biological system or complicated systems.

2.2.3 Defining a health care system as a complex system.

To define a health care system as a complex system the definition of complex systems according to Magee and de Weck will be used.

Definition of complex system according to Magee and De Weck:

> *A system with numerous components and interconnections, interactions or interdependencies that is difficult to describe, understand, predict, manage, design and/or change [6].*

2.2.4 *Classification Criteria*

There are clear attributes described in the definition that can be used to measure a health care system as a complex system. Testing of all the mentioned attributes against those of a health care system will indicate how closely a health care system can be classified as a complex system according to Magee and De Weck's definition.

Attributes of a complex system according to Magee and De Weck's definition can be summarised in the following list:

- Numerous components.
- Numerous interconnections.
- Numerous interactions.
- Numerous interdependencies.
- Difficulty to clearly describe above mentioned.
- Difficulty to understand above mentioned.
- Difficulty to predict above mentioned.
- Difficulty to design/change above mentioned.

In the following section it will be shown how a health system exhibits the characteristics of a complex system

2.2.5 *Health Care Fit to the Classification*

To classify a health care system as complex, consider a health care system of a country such as South Africa. The South African health care system has more than 650 hospitals and 5000 lower-level care facilities spread over the entire country, excluding private practices and traditional healers. This is a clear indication of a **number of components** comprising a well-defined health care structure [7].

There are **numerous connections** between these health care facilities. One example of such a connection is the information that is shared between them, combined with patients who are referred between the levels. This includes patient information and information with regards to the state of the hospital; i.e. whether it can accept more patients or not. Furthermore, there is a regulatory process connecting the higher levels of regulation to the provinces, each province having districts and the district having sub-districts with hospital systems contained therein.

Because of the referral approach used, the **interdependencies** between facilities are increased. The capacity and capabilities of facilities also influence the load utilisation and care of the surrounding facilities. There is also an epidemiological dependency of facilities on each other; there is very little stopping patients from utilising more than one facility over time, and so each facility is dependent on the care provided by the others.

Due to the operating environment stated above and the inherent variation in the process of healing and sickness that directly influences the health care system it is very difficult to formulate a **clear description** of the exact relationships and dependencies. To further **predict** the outcome of any case with certainty is very difficult. By aggregation the operations the ability to make predictions improves, but it remains a prediction in an environment with large inherent variation.

The difficulty in designing a health system lies in its complexity regarding size and interdependencies. Changing such a system could be seen as much more difficult due to the large role humans play in the system. Since it is almost entirely operated and used by humans, socio economic, cultural and personal influences cannot be separated from possible change.

2.2.6 Conclude health care system as a complex system

Defining a health care system as a complex system entailed taking the definition of a complex system from Magee and De Weck and identifying the characteristics of a complex system. The characteristics of a health care system were then compared to the identified characteristics.

It was found that a health care system clearly possesses all of the characteristics of a complex system. Therefore henceforth a health care system is defined as being a complex system in this document.

The second step in the classification of health care system is to question if it can be defined as an engineering system.

2.3 A HEALTH CARE SYSTEM AS AN ENGINEERING SYSTEM

2.3.1 CLASSIFICATION OF ENGINEERING SYSTEMS

The classification of health care systems as engineering systems (ES) starts with an overview of previous classification schemes conducted by Magee and De Weck. Characteristics of engineering systems are identified and augmented by the characteristics suggested by Magee. The suggested characteristics are then measured against those of health care systems. Finally, a health care system is defined as an engineering system.

Magee and De Weck reviewed a selection of system classification schemes, starting with the work of Bertalanffy in the 1950's.

Bertalanffy's classification scheme has a strong relation to biological classifications. The scheme starts with static structures, moves to organisms and ends with Symbolic systems. Each scheme increases in complexity [8]. Secondly Paynter's framework as used in his MIT course was reviewed. Paynter considered four system types [9]:

1. Services and Utilities
2. Structures
3. Instruments
4. Vehicles

Paynter did not consider the human factor or any cross boundary system characteristics in his framework.

Thirdly Hubka was mentioned as he had a more complete approach, considering not only the technical aspects of the system but also the human and environmental aspects. Hubka considered: Function, branch of the economy, type of operand, physical principles of importance, product use, production method, materials etc [10].

Magee and De Weck suggests that an engineering system can be distinguished from other complex systems by the following attributes; human designed, high technical complexity, and high human complexity. Engineering systems are real, open, dynamic, with hybrid system states, and have autonomous and human-in-loop subsystems or elements [5].

Due to the comprehensive overview of complex system literature and relative recent nature of the proposed scheme of Magee and De Weck it will be used to measure the health care system against to establish if a health care system falls within the bounds of an engineering system.

2.3.2 *CHARACTERISTICS OF ENGINEERING SYSTEMS*

Magee and De Weck identified the following attributes as being that of an engineering system:

- Human design.
- High technical complexity.
- High human complexity.
- Real.
- Open.
- Dynamic.
- Hybrid system states.
- Autonomous subsystems.
- Human-in-loop subsystem.

Each of the above mentioned attributes will be discussed to establish to what degree health care systems fit into the engineering system classification.

2.3.3 *ASSESSMENT OF HEALTH CARE SYSTEMS AS ENGINEERING SYSTEMS*

The attributes are discussed by following the structure of first defining a *working definition*, secondly by indicating *how a health care system exhibits the applicable attribute*, thirdly if the attribute is present in a health care system, and finally *to what degree the attribute is present* in a health care system.

Human designed

Human designed systems are designed and implemented by humans to perform a specific function. Unlike natural systems where natural accruing systems exists without any human intervention or interaction, a human designed system exists for human benefit.

A health care system is implemented by humans to facilitate the provision of health care, irrespective of private or public providers of services or suppliers products. The health care system would not exist if humans did not exist to use it or design it for the specific purpose.

The health care system is humanly designed to a high degree. There is almost no aspect of the health care system that cannot be directly linked to human design. From the most basic subsystems such as a single system to allocate nursing staff rosters in a small clinic to a complex subsystem such as the national tissue bank of a country, the human design can be seen throughout the health care system.

There are elements within the health care system that are not of human design. The most obvious is the human factor of the clients or patients within the system. The clients or users of the system are of a biological nature.

Secondly, the product or aim of the health care system is to provide health care. The provision of health care can be seen as the service of caring and supporting operations, but the health that is the aim of the care is not of human design. Disease or burden of disease that the health care system is to alleviate, is also an intricate part of the health care system, not of human design, but of a biological nature. Even though disease is such an integral part of the health care system, the health care system is designed to combat disease and react to disease rather than disease being part of the health care system.

High technical complexity

Technical complexity can be described as the intricacy of the technical aspects of the system. The technical nature of operations can relate to the specific skill, often scientific in nature, with which the operation is to be completed. Usually the operation needs to be completed with a certain level of the appropriate skill and competence to be successful.

The technical attributes of a health care system is manifested in a variety of areas within the system. At an operational level the technical complexity is the most obvious. Technical complexity is not only displayed in the procedures that are performed on patients, but also by the technical nature of equipment and resources used during the operational aspects of a health care system.

At higher levels of operation, technical complexity is less visible but effectively increases. The technical complexity to manage the operations within the system can be attributed to the highly stochastic nature of the operating environment and the high level of priority that is needed in some cases. From the management of theatre time and scheduling of staff to the administration of patient records and management of emergency response systems, there is an intrinsic level of technical complexity.

At the highest level of health care system technical complexity can be seen in the interaction of the system with other systems and within itself. There is a definite technical component to the management of the system. In areas such as funding allocation to the system components, system-wide measures, and analysis are but a few of the technically complex manifestations within the health care system.

The health care system displays a definite tendency towards technically complex operations at various levels through the system hierarchy.

High human complexity

High human complexity is apparent in a health care system. Due to the various roles humans play in the health care system, the attribute of human complexity will be discussed from four points of view. The four points of view are as follows: humans as raw material, humans as a tool, humans as the client and humans as the operators.

The human as raw material used within the health care system.

The health care system uses a lot of other materials, but the main "product" of the system, which is health (human in the state of non-ill health), is supplied by using a human (in a state of ill health) as the raw material. The complexity associated with a human as the raw material is not only limited to the biological aspect of human complexity, such as complicated physiology and neurology. In addition to these factors, there is a complex social structure associated with humans. Both the physical and non-physical attributes augment the complexity of humans as a raw material.

The human as tool in the health care system

Humans that provide the care within the health care system fill the same role as a device or tool in manufacturing plant. This not only refers to the health care professionals but also the supporting staff and other humans performing duties to facilitate care. The human complexity from this point of view is less in a physical sense but more of a social and interaction complexity. The interaction is between the humans acting as tools within the system and with the humans acting as raw materials/clients. The complexity of the interaction between the humans performing the tool roles can be described by the different roles each group has to play within the system. A system of hierarchy and specific skills being supplied by a selection of suppliers (humans acting as tools) illustrates the complex nature of interpersonal complexity between humans acting as tools.

The social factor adds to the complexity of the relation between the humans, and as a result, the human complexity within the health care system.

The human performing the role of the client of the health care system can be vastly expanded. The stakeholders to a health care system are very wide, and cross social boundaries. Political and industrial boundaries are also crossed. In fact, almost every person on earth is a stakeholder to at least one health care system.

The number of different stakeholder groups increases the complexity of a health care system due to the vast number of expectations from the health care system and often the opposing nature thereof.

The political stakeholders want an effective health care system that promotes their priorities.

The clients (patients) of the health care system want the best service possible from the health care system for their personal needs.

The community wants the health care system to address problems of health with larger impact than a single household or person, but to the maximum advantage of that community.

The providers want the health care system to provide a viable and sustainable working environment for them to practice their trade and earn a stable income.

Humans as the operator of the system

With humans operating the health care system, there is a definite effect of human complexity for which there needs to be accounted for. The operations are complex (technically and humanly) but the management thereof is subject to all of the complexity and limitations of human factors. Personal interest, conflicting priorities, incompetence, exceptional expertise and pride are but a few of the elements that drive the complexity of humans as the operators of the health care system.

Real

A real system is an existing system that has a tangible effect on the surroundings or environment it intends to influence.

The health care system is a real system, with tangible and intangible components. The system provides a real service with real products and expertise flowing through it. There is a large intangible component within the health care system. The gaining and transfer of knowledge is an example. This has a real effect on the rest of the health care environment.

Open

In an open system there is a flow through the system, as opposed to a closed system where all of the elements within the system stays in the system.

The flow of elements (patients, knowledge, resources) in the health care system could have a notion of a closed system with patients that are chronically ill or resources being used continuously but still there is a point where the elements exit the system never to return and new elements enter.

A second characteristic of open systems is the degree to which there is interaction with the surrounding systems. With the health care system being a system with such an influence on the higher priority of human needs, the interaction with surrounding systems is active to a large extent. Further the level of interaction with surrounding systems is increased by a large number of stakeholders.

Dynamic

Dynamic systems can change between system states. For a system to change between states it should be possible for the system to exist in a range of states. The dynamic system is able to function and operate at various states. The dynamic system reacts to its environment. The system is not static and unaffected by the environment it is in, on the contrary it actively engages to adapt to the change of state of itself and of its environment.

The health care system is in a continuous change of state. This continuous change of state can largely be described by the highly stochastic environment it operates in and the need to adapt to that stochastic nature of the requirements of the system.

The health care system is a system that adapts to the changes in its environment. The health care system reacts to the changes in it environment to succeed in its purpose of facilitating care.

The health care system is a truly dynamic system. With the dynamic reaction of the system to itself and its own state and the reaction of the health care system to its ever changing environment the health care system is ever changing.

Hybrid system states

With hybrid system states in this context the assumption is made that the system has both continuous and discrete properties. This property of hybrid system states is usually the reason for modelling of such system with a selection of differential equations.

The way a health system exhibits the continuous properties is largely through the epidemiological impact on a health system as well as the flow of resources, patients and information through the system. The discrete properties exhibited by the system can be most obvious in the total absence of a health system but can also be observed on process level. For example, when a facility is closed / or a position is open, the care is standing still for that part of the system at that point in time.

Autonomous and Human-in-loop subsystem

Autonomy refers to sub-systems or components that are not reliant on direct human interaction. A human-in-loop subsystem refers to the interaction and dependence of human intervention in the subsystem.

Human-in-loop subsystems are highly visible throughout the entire health system. From high level or national management to process level at the facilities, human involvement can be displayed.

Autonomous subsystems in the health system are more readily displayed in the supporting functions of care. Routing systems for emergency response or transferring of information between hospitals or hospital systems are a few examples.

The emphasis in health care systems is on human-in-loop subsystems. This again emphasises the inherent variable and reliance of health systems on humans. The characteristic of both autonomous and human-in-loop subsystems is evident in health care systems, although it is evident to a lesser degree compared to other characteristics discussed earlier.

2.3.4 CONCLUDE HEALTH CARE SYSTEM AS AN ENGINEERING SYSTEM

In the above analysis, typical attributes of engineering systems are compared to attributes of a health care system. The attributes Magee and De Weck suggested includes human design, high technical complexity, high human complexity, being a real, open and dynamic system with hybrid system states and autonomous and human-in-loop subsystems. The attributes of a health care system closely mimic these attributes suggested. Only the attributes of autonomous subsystems are displayed to a slightly lesser extend compare to the rest to the attributes.

With this reasoning from this point forward in this document, a health care system is defined as an engineering system.

2.4 MODELLING AND SIMULATION OF ENGINEERING SYSTEMS

Simulation is arguably the most commonly used Operational Research technique and has widely been used in the health care domain. This is mainly due to its flexibility, ability to deal with variability and uncertainty, and to facilitate communication between different professions [11].

Due the uniqueness of the different approaches, some approaches are more effective in certain applications than other approaches.

It is therefore necessary to identify the most appropriate approach for the purpose of this study. Before choosing a modelling approach the possible approaches should be reviewed and their characteristics identified.

2.5 MODELLING AND SIMULATION OF HEALTH CARE SYSTEMS

With the possibility of modelling the South African Health Care system the question follows: how is modelling used in health care?

Mustafee *et al.* answers the question by reviewing 201 publications from high quality journals published between 1970 and 2007. The papers that were reviewed have been categorised under different simulation techniques. Further the health care application for the given technique is investigated.

The four relevant simulation techniques that have been identified through the review of the papers are: Monte Carlo simulation, Discrete event simulation, Systems Dynamics, and Agent-based simulation [12].

An overview of the simulation techniques identified by Mustafee *et al.* will be discussed as follows:

- Definition and description
- Characteristics
- Information needed
- Outputs
- Primary use
- Limitations

2.5.1 *Monte Carlo Simulation*

Definition and description of Monte Carlo Simulation:

Monte Carlo simulation (MCS) was chosen as a code name by John von Neumann and Stanislaw M. Ulam for their work in 1946 [13]. Due to the sensitive nature of their work (Manhattan Project) they had to choose a code name. They chose the name to reflect the modelling approach's similarity to gambling

MCS is a statistical technique that uses the stochastic qualities of an element and simulates processes accordingly. Random numbers are created according to a defined statistical distribution. The random numbers are then manipulated to form values that are used in the simulation and calculated in a discrete fashion.

Large number of iterations of the discrete calculations is done with the stochastic inputs, to give a stochastic representation of the simulation. This information is then aggregated to illustrate the simulation. The replication that is run is a way of sampling the possible reality that is modelled.

As an example, the duration of a journey to a destination may be simulated as follows; each portion of the journey would be identified and the corresponding times for each portion and the potential outcomes for a given portion would be given a probability of occurring. The duration of each portion of the journey will be defined by a statistical distribution. The expected time for a given journey will be calculated by combining all of the portions of the journey.

Monte Carlo replicates the scenario many times to establish which statistical attributes the system has. By having an idea of the statistical nature of the system, better informed decisions can be made regarding the system.

Information needed:

The process under investigation is to be broken down into steps or portions that can be described by a single distribution or statistical characteristics.

This distribution has to represent the reality to a sufficient degree. With the known statistical characteristics of the steps or portions, the steps can be modelled in sequence.

Primary uses:

MCS is widely used to model stochastic environments where discrete solutions cannot be obtained or where such a solution would be an infeasible option. The sampling of MCS makes it a useful approach as an alternative means of obtaining solutions.

One of the popular applications of MCS is the analysis of risk associated with a project or process. Due to the statistical nature of the output from a simulation the calculation of risk, whether it is the risk of long processing times or financial risk of a project, can be calculated.

MCS can be used in sensitivity analysis. With the change of one parameter at a time the simulation can give an idea of the sensitivity of a scenario or process to the elements impacting on it.

Limitations:

The use of MCS is highly dependent on the statistical accuracy of the input data. Where limited data is available the application of MCS can be limited.

With the input data, the approach used to model the reality can have a large impact on the output from the model. It is thus important to make an informed decision on the way the model represents reality.

As MCS utilises replications of specific cases, the modelling of continuous systems is thus limited.

2.5.2 DISCRETE EVENT SIMULATION

Definition and description of Discrete Event Simulation:

Discrete Event Simulation (DES) is one of the most commonly used operations research techniques in practice [11].The development of DES cannot be claimed by any one person, but it is rather a result of continuous growth of simulation as described by Sammet and Nance [14] [15]

DES is based on the concept of events or activities happening to change the state of the system over time. There is a flow of entities throughout the model which are processed by the events. The logic of the simulation model is captured relative to the relevant time line. An event is scheduled to happen at a specific point in time. As the event occurs, at the specified point in time, a variable is changed, adapted or influenced. The event can represent any activity or process within the model. A typical event would be a patient that arrives at a facility, or a form that is processed within that facility.

The event of arrival or processing would impact the variables defined within the system. Variables are defined in order to represent the reality in the model and measure the changes within the model. According to Ross there are generally three variables that are utilized in DES: Time variable, Counter variable and System State variable.

The time variable is defined to guide the model with regard to the time axis that is used as a reference for events to occur. As an event happens, the following event is scheduled to happen at the appropriate time. With the completion of an event, the simulation "jumps" to the time that the next event that is scheduled, and stops as that event occurs. On completion the model jumps again to the next event. The event doesn't need to follow the initial event, but could have been scheduled following the completion of another event that occurred. All of the events scheduled to happen are "placed" on the same time scale. It is thus not possible for two events to occur at the same instance in time.

The method used to schedule events allows that the stochastic nature of reality can be incorporatedin the simulation. The times between events, arrivals or process time can be defined as stochastic times with predetermined distributions.

The counter variable is used to give information with regard to the elements within the system. The elements enter the model, go through a selection of events or processes, and then exit the model. The counter variables count the elements within the model. This could refer to the current number of people in a queue or the number of patients that moved through a facility in a given amount of time. By using counter variables, trends could also be drawn and analysed.

The system state variable gives an indication of the state of the system. With a resource being either active or not active the state of the resource can be used to calculate the utilization of that specific resource. The state variables are defined as needed in the model to portray information from the simulation in a useful manner.

Information needed for DES

- The process logic and steps, as well as the resources utilized, are to be clearly defined.
- The data with regard to the corresponding process times has to represent the reality that is modelled.
- Utilized resources capabilities and capacities.

Outputs from simulation:

- Detailed description of waiting times
- Trends of queue lengths

Typical applications:

Discrete event simulation is primarily used for analysis and simulation of operational level processes. The simulation of clients arriving at a bank and being served by a given amount of tellers would be an example of such a simulation. In health care the use of an operating theatre or emergency room would be text book examples of discrete event simulation. The average waiting time in the queue, utilization of resources, probability of open operating theatre and cost could be analysed with such a simulation.

Limitations DES approach:

- Due to high level of detail the complexity of the model increases greatly as the system or process that is modelled increases in complexity.
- The process that is modelled should be known in detail to build the necessary logic into the model

2.5.3 SYSTEM DYNAMICS

Definition and description of System Dynamics:

Professor J Forrester, who has a background in engineering, created the System Dynamics (SD) approach in the 1950's in an attempt to understand why and how cooperation succeeded or failed [16].

In the modelling of complex systems SD is a very powerful methodology. The use of feedback loops and delays makes it possible to model a highly complex and dynamic system at high level over time.

According to Randers the four stages of modelling are [17]:

1) Conceptualization
2) Formulation
3) Testing
4) Feedback

SD is an analytical modelling methodology that has two distinct aspects: qualitative and quantitative [11].

The qualitative aspect of SD modelling is concerned with the development of the influence diagram. The influence diagram is a representation of the system that is modelled. The elements within the influence diagram are identified as the elements within the system that has the greatest impact (at the relevant level) on the system. In the influence diagram the impact of the elements on each other is described as positive and negative influences. The loops that are created by connecting the elements can then be defined as stabilizing or vicious cycles.

A stabilizing cycle governs the levels within the cycle. An example of this would be the food in ones stomach over a 24 hour period. When the level of food in your stomach is low you become hungry and eat something. By eating something the level of food in your stomach increases and you feel full again, when full you to stop.

A vicious cycle can be described by extending the period of the earlier example. If you continue to eat a bit more than is needed your stomach stretches, this will increase the volume of your stomach and as a result you will be able to eat more before feeling full. As you eat more over time your stomach will stretch even more and you will eat even more as a result.

The quantitative aspect of SD is based on a concept of stocks and flows. The influence diagram developed in the qualitative part of the SD approach is converted into a stock and flow diagram. The stock and flow diagram is modelled by linking the influences of the different levels and elements within the model with differential equations. In the stock and flow model stocks flow to and from different positions. As stocks increase or decrease it is indicated as a level. Stocks do not only represent physical entities such as number of patients. A level of intangible properties can also be indicated, such as health or happiness.

The levels are linked by valves that govern the flow between the levels. This is the flow between levels that is modelled as differential equations. Not only is the flow of mass represented, but also the flow of information within the system.

The state or level of the following time unit is a function of the current level plus the rate of inflow multiplied with the time unit minus the rate of outflow multiplied with the time unit passed. To simplify; the level can be seen as the total flow into, minus the total flow out off at time *n.*

$$\text{Level}_n = \sum_0^n Flow_{In} - \sum_0^n Flow_{Out}$$

The information governing the system is captured in the structure of the model. The information and levels of the stocks are combined to describe the logic or operational aspects of the system that is modelled. Information flows in feedback loops in the model and through the influenced that impacts the elements.

Information needed:

- Holistic system knowledge to identify the necessary elements defining the system.
- Knowledge based on the interaction of the elements within the system.
- Clear view of what is to be measured.

Outputs from SD simulation:

- High level trends of stocks within the model
- Indication of the flow between elements
- Indication of the interaction of the elements on each other
- Indication of the effects of changes of the elements on the system.

Limitations:

The limitations associated with SD modelling are strongly associated with the fact that high level modelling of complex systems is done with much uncertainty. This uncertainty can be found in the data used in the model, the assumptions made during modelling and the difficulty in modelling human behaviour. Even though various steps can be taken to ensure accuracy of the model DC Lane states that SD models are never more than 40% accurate [18].

The continuous nature of a SD model limits the application of SD modelling in situations where it is needed to model a single entity, or discrete amounts.

2.5.4 *AGENT BASED MODELLING*

Definition and description of ABM:

Agent-based modelling (ABM) is used to model the interaction of various agents or elements with each other and their environment. Each of these agents is driven by simple rules and a given goal. They have the ability to learn and interact accordingly with the environment and other agents. Gilbert explains in his book *Agent Based Modes* that ABM allows for:

1. Creating various agents within a model.
2. A mix of attributes and driving rules can be unevenly assigned to agents.
3. A representation of learning mechanisms.
4. Modelling of agent interaction with each other and their environment [19].

By utilizing the above mentioned capabilities of ABM the modelling of a single simple agent can result in highly complex models of systems. By reproducing the single agent and letting each agent interact with its environment the complexity of reality is mimicked.

The interaction of people or animals among themselves and their environment, such as movement around a train station or swarming colonies, can be modelled.

An example of use for such a modelling technique is the modelling of traffic. Each agent wants to arrive at their allocated destination on time by spending the least amount of time in traffic. To avoid traffic, the agent then learns what time and route to travel to minimise the time spent in traffic. The simulated road network can then be analysed to assess the impact a new road, or toll gates will have on the users and infrastructure.

Outputs:

- The outputs of ABM simulations usually have a strong visual component. This visualisation helps with interpretation of the modelling results. The characteristics of each agent within the simulation can be stored as information for that agent
- The information of each of the agents can be compiled to build a picture of the characteristics of the simulation as a whole.

Information needed:

- For an ABM simulation the logic of interaction of the agents with themselves and their environment should be known.
- The level of learning by the elements should be clearly defined and validated.

Primary uses:

ABM is mainly used to model the interaction of agents with the environment for example traffic analysis, social structures, spreading of information or disease within populations. Another application of ABM is within the operations research applications. Optimization heuristics is one example of such an application.

Limitations:

The high intensity of computing power needed for ABM limited its application at the time of Thomas Schelling's early models in the 1970's. Only in the 1990 did the methodology of ABM start to grow, largely due to the increase in availability of computing power.

Increased automation of the ABM methodology means that verification and validation of the model plays and very important role in the modelling. There are definite difficulties in verification and validation of new models with unintuitive results.

2.5.5 MOST COMMON AREAS OF APPLICATION IN THE HEALTH CARE ENVIRONMENT

Mustafee *et al.* did not only identify the simulation modelling technique used but also the area of application within health care. The areas of application are ranked from most frequent to the least frequent.

Monte Carlo simulation:

- Assessment of health risk.
- Cost effectiveness evaluation of different strategies or technologies.
- Diagnostics on assessment of certain medical interventions
- Methodological improvement and comparison

Discrete Event Simulation

- Health economic costs.
- Policy and strategy evaluation models.
- Review papers
- Methodological improvements and comparison.
- Assessment of health care management models.
- Accident and Emergency department modelling.
- Public response to contagious diseases and bio terrorism.

System Dynamics

- Public health policy evaluation
- Training tool to illustrate dynamics of epidemiology
- Modelling of health care systems.
- Infrastructure modelling
- Health economics modelling

Agent Based Modelling

The application of agent based modelling is still too widely spread to accurately identify specific areas of application in health care [12]. Only the following two areas were identified.

- Dynamics of hallmarks of cancer.
- Accuracy of geological health analysis.

Mustafee *et al.* discusses some of the variables related to the papers published under each of the modelling areas, the most productive authors are mentioned, the amount of publications per year, the source titles and subject classification.

There is a steady increase in the number of papers published in each category. Irrespective of modelling category the most journal papers have been published by the *Journal of Operations Research Society*. The subject classification ranges from

Environmental Sciences, Operations Research and Management Sciences to Management [12].

By looking at the profiling study of Mustafee *et al.* there is a clear application of simulation in health care. Even though only high impact journals and papers were considered, the characteristics of the modelling areas are clearly displayed.

Mustafee *et al.* concludes the paper with the following phrase:

> *Within the profile it appears that simulation is 'healthy' and represents a fascinating range of research and application [12].*

2.6 Modelling of complex systems

With the modelling of a complex system being different from the modelling of a non-complex system or a process, special care should be taken, as a health care system is a complex system. For this study which aims to model a health care system, a modelling approach is needed that can describe a complex system realistically and capture the characteristics of a health care system.

One possible approach to modelling complex engineering systems can be to completely model the whole system, including the smallest parts to the highest possible level of detail. With a complete model approach very large amounts of very accurate data is needed for the model to represent the reality of the system. All of the interactions of the system being modelled should be known and represented in the model.

An attempt to model a complex engineering system with the above-mentioned approach is flawed due to the characteristics of complex systems. For example, human complexity is highly unlikely to be modelled correctly to realistically predict the complexity of human nature in a dynamic environment. It is a great challenge in its own right.

The characteristics of open systems are another obvious problem. To model a system that is open to the finest detail, a model should be developed to be so comprehensive that the size will increase up to the point that the system modelled is no longer an open system but a closed system. This model will then consist of various other systems, possibly becoming a model of a system of systems.

It is clear that the approach of modelling a complex engineering system to the finest detail is flawed. An alternative approach to modelling complex engineering systems is needed. An approach that aggregates the system to a manageable size whilst keeping to its essence, is necessary.

One such modelling approach is system dynamics.

By the very nature of system dynamics approach, the maps and models are pitched at a high level of aggregation and concentrates on the dynamics of the various systems being investigated [20].

The system dynamics methodology addresses the challenges of modelling a complex system in some of the following ways.

- Feedback loops
- Delays
- Non linearity

By incorporating these attributes into the modelling a complex system can be aggregated and modelled while keeping the system characteristics.

2.7 Conclusion

In this chapter a health care system was first defined as a complex system. Secondly a health care system was defined as an engineering system. Using these definitions possible modelling approaches were identified. Monte Carlo simulation, Discrete Event simulation, System Dynamics and Agent-based modelling were studied to assess their characteristics as an approach. Keeping in mind the purpose of the study, the most appropriate modelling approach was identified as being System Dynamics.

The ability of the SD approach to model high level complex systems was then verified with the type of applications it was used for previously, as identified from literature.

Closer inspection of the SD approach will now be made to develop the modelling framework.

Chapter 3

HEALTH INDICATORS AND PERFORMANCE MEASURES

3.1 CHAPTER OVERVIEW

The purpose of the review of health indicators and performance measures are two fold, firstly to gain an understanding of how the performance of a health system is measured, that will be used to inform model building. Secondly generic types of elements are identified that will be used in the modelling framework. In the chapter the application of health indicators and performance measures are discussed, to form an identification of the considerations to take when choosing indicators. By comparing the approaches of two world leaders in health system monitoring the generic types of considerations important to health system performance were identified and used to identify system elements that are used used in the framework.

3.1.1 *BACKGROUND ON HEALTH INDICATORS AND PERFORMANCE MEASURES*

Health indicators and performance measures are widely used terminologies, with their meaning often depending on the context in which they are used. For further use in this document working definitions of the terms health indicators and performance measures are defined as follows:

- Health indicators (HI): A representation of the state of an entity, or collection of entities, that is measured by means of a standard, rating or index.
- Performance measure (PM): Use of the state of the entity (indicator) that is measured relative to some comparative benchmark, to establish how the entity performs

Analysis of health indicators makes it possible to measure the performance of a single entity or collection of entities. It is essential that health indicators used for performance measure are chosen wisely.

By understanding David M Eddy's properties of a good measure it will be possible to differentiate between a good measure and a bad measure.

Eddy identified the following factors as being important for good measures and indicators [21];

i. Purpose
ii. Entity
iii. Dimension
iv. Type
v. Audience

Each of the factors will now be discussed in further detail.

i. The purpose of the measure.

The purpose of the measure is the first attribute that has to be defined. If the purpose of the measure is not clearly defined there will be ambiguity in the use of the measure, and in the application of the measure. If the measure is blindly followed and used for the wrong performance indicator, it can have negative effects and unintended consequences on the processes, systems or entities to which it is applied.

The purpose of the measure will be defined by the intended use of the measure or the audience. In turn this will impact the level of information used in the measure as well as the way the measure is represented. As with any data or information there are an infinite number of approaches to follow to represent the indicator. The representation of the measure should be in line with the purpose of the measure to make it useful for its purpose, and not just for the sake of measuring or representing of data.

Eddy describes three main purposes: description of an effect of some intervention, measurement of an improvement, and comparison of quality of care being delivered by different entities [21].

Of these, the easiest to perform is the description of the effect of an intervention. Only one measure needs to be taken after a certain time to describe the effect of the intervention. To measure an improvement is slightly more difficult because the original process has to be compared to the improved process, while keeping all other possible influences to the outcome the same. This requires multiple measurements over time. The most difficult measure to perform is the comparison of care delivered by different entities. This is due to the variation in the operating environment, and the fact that the outcome of any treatment is only partially a result of the treatment itself. The outcome of treatment is dependent on a selection of factors not limited to clinical ones, but demographics such as socio-economic factors, medical history and family background all play a role in the eventual outcome of treatment. All this variation needs to be taken into account and normalised, only then can two entities be compared in the delivery of care.

ii. The entity being measured

The entity being measured will have a strong influence on the choice of health indicators used to measure the performance. Different collections of indicators will be used to measure a physician's performance compared to the performance of a hospital or region's health system. The type of entityto be measured is a strong indication of the level of detail that will be needed to measure the performance.

At a lower level of analysis, specific process measures could be more appropriate to use compared to high level outcome measures used for higher level securitisation of a hospital system or districts health services.

The entity being measured can place limitations on the measure that would be used from a practical point of view. In a very rural setting, complicated measures that are time-consuming to gather would be less likely to be reliable or of good quality than the same measures from a highly developed academic hospital which feed information directly into a health information system.

iii. The dimension of quality or performance that is being measured.

The Institute of Medicine (IOM) defines quality of care as "safe, effective, patient-centred, timely, efficient and equitable" [22]. From this definition six dimensions of quality are identified as:

- Safety
- Efficiency
- Patient focus
- Timeliness
- Efficiency
- Equitability

There are various definitions of quality or performance of care described in literature. The focus of performance measurement lies in the dimension of quality or performance to be measured, not the definition of quality. Each of the defined dimensions has different attributes and the one indicator will not necessarily be a good measure for another dimension. It is thus important to choose the correct indicators to measure the performance of a specific dimension of performance.

Because of the wide scope of the term "quality", the dimensions of quality vary greatly. In the IOM's definition, for example, the dimensions of "safe" and "timely" will need very different indicators to measure performance accurately and validly.

To get a complete picture of the performance of an entity across the dimensions being measured, composite indicators are often used [21]. A relative weight is assigned to a dimension of quality or performance and the scores of each dimension are added up to give an index of quality or performance.

When analysing such an index, it is important to know the background of the index and the weightings of the dimensions that were measured, as well as the reasoning behind the choice.

iv. The type of measure.

The fourth consideration identified by Eddy is the type of measure. Measures can be broadly categorised as, non-population based or population based.

Non-population based measures are used for individual monitoring. Biological measures play the most significant role in non-population based measures and are mostly used to assess the state of health of an individual.

Within a given population three types of measures exist: process measures, biological measures and outcome measures [21]. Hutton suggests the addition of a structural measure over and above the three types of measures proposed by Eddy [23].

Process measures give an indication of the performance or state of the process involved in providing care. The processes include indicators such as percentage of patients on ARV's with a Cluster of differentiation 4 (CD4) count of less than 350, or the number of patients at a primary care level being tested for HIV and TB. Waiting times, queue lengths and resource utilisation are further examples of process measures. Usually, operational performance is measured using process measures, although clinical information can also be shared using process measures.

Biological measures are used to measure the state of one patient or group of patients. The CD4 count of a patient would be a typical biological measure, as would the blood pressure or cholesterol levels of an individual. By utilizing the biological indicators of a population, an idea of the health of the population can be formed. Disease prevalence and the average blood pressure of a population are both examples of such measures.

Outcome measures represent the degree of a given outcome. The percentage of people dying of TB with a CD4 count of 300 or less, the percentage of patients being cured from skin cancer or the mortality rate for premature births are outcome measures. An outcome measure aggregates the whole process leading up to the measure; all of the preceding processes of care for premature births will influence the measure of mortality of prematurely delivered babies. While an outcome measure is much more comprehensive by being a function of all of the other processes, the limitations of outcome measures are also due to this attribute. Because of the levels of aggregation, the results of a population-based process measure can represent a large variety of sub-populations.

An example of the variation would be the life expectancy of a South African that was born in 2009. The live expectancy is estimated by the WHO to be 54 years, yet there is a very large population that is older than 54 years of age [24]. The relatively low life expectancy is largely due to the high incidence of HIV and AIDS and associated premature deaths.

The *structural measure* suggested by Hutton refers to the infrastructure, resources, and supporting services in place to augment the functioning of a health care system. The number of clinics per district, beds per hospital, health care worker to population ratio or the availability of ambulatory services would be examples of structural measures.

v. **Intended audience.**

Lastly, the intended audience is a consideration that mainly affects the acceptable level of complexity that can be measured whilst still conveying useful information. For the measure to be useful it needs to be easily understood by being clear, at the right level of detail, and of appropriate clinical sophistication.

3.1.2 Conclusion

Although the example of life expectancy stated above is only for one country there are significant differences between the life expectancies of sub-populations within that country. This would be a clear example where the considerations mentioned should be used to guide indicators chosen to represent the information.

The questions should be asked:

i. What is the purpose of the measure?
ii. What is being measured?
iii. What dimension of quality is being measured?
iv. What type of measure is needed?
v. Who is the intended audience of the information?

The choice of measure to use is essential in measuring the performance of a health care system. The measures that are chosen will directly influence the rating or performance measured in the system, and will also influence potentially far-reaching decisions based on the systems performance. Therefore good measures should be chosen that represent the state of the system accurately, and are appropriate for the intended use of the measure.

3.2 COMPARISON OF APPROACHES TO MEASURE HEALTH SYSTEM PERFORMANCE

On service provider level, various measurement systems are available to analyse facilities such as a health centres or hospitals [25]. These measures are usually obtained from a quality point of view and are used to compare the measured facilities.

Where low or medium-level provider measurement systems are plentiful, high level system measurement frameworks are less common. The two most prominent organisations involved in the measurement and comparison of health care systems are the World Health Organisation (WHO) and the Organisation for Economic Co-operation and Development (OECD). The approach used to measure health system performance by each of these organisations are now discussed

World Health Organisation:

The World Health Report 2000 was released in June 2000 by the WHO. This report arguably initiated the global interest in health system performance, and the measurement thereof along with the large amount of discussion on the topic. The purpose of the report was to "help the Member States measure their own performance, understand the factors that contribute to it, improve it, and respond better to the needs and expectations of the people they serve" [2].

To understand the framework used by the WHO, the different levels of indicators and measures will be discussed. The ranking of the health systems was done according to the combined relative score for each of the three major components. The major components score is calculated from relative scores of the sub components of the major component. The relative scores of the sub components are calculated using the chosen health indicator.

For the purposes of ranking the health systems of the WHO member states three major components were analysed [26];

- Goal attainment
- Health expenditure per capita
- Efficiency and overall level of health performance.

$$\boldsymbol{Health\ System\ Index} = \sum\left[\left(W_g \times Goal\ Attainment\right), \left(W_{he} \times health\ expenditure\right), \left(W_{oe} \times overall\ efficiency\right]\right)$$

With:

$W_g = Relative\ weight\ for\ Goal\ Attainment$

$W_{he} = \ Relative\ weight\ for\ health\ expenditure$

$W_{oe} = Relative\ weight\ for\ overall\ efficiency$

$$\boldsymbol{Goal\ attainment} = \sum_{1}^{n} [(W_n^{hi} \times S_n^{hi}) \times C_n^{hi}]$$

With:

$W_n^{hi} = Relative\ \boldsymbol{weight}\ for\ Sub\ component\ n$

$S_n^{hi} = Relative\ \boldsymbol{score}\ of\ Sub\ component\ n$

$C_n^{hi} \triangleq 1 : If\ Sub\ component\ \boldsymbol{n\ is\ chosen}, 0 : Ohterwise$

$$\boldsymbol{Health\ expenditure} = \sum_{1}^{n} [(W_n^{hi} \times S_n^{hi}) \times C_n^{hi}]$$

With:

$W_n^{hi} = Relative\ \boldsymbol{weight}\ for\ Per\ capita\ expenditure\ n$

$S_n^{hi} = Relative\ \boldsymbol{score}\ of\ Per\ capita\ expenditure\ n$

$C_n^{hi} \triangleq 1 : If\ Per\ capita\ expenditure\ \boldsymbol{n\ is\ chosen}, 0 : Ohterwise$

$$\boldsymbol{Overall\ efficiency} = \sum_{1}^{n} [(W_n^{hi} \times S_n^{hi}) \times C_n^{hi}]$$

With:

$W_n^{hi} = Relative\ \boldsymbol{weight}\ for\ Sysem\ efficiency\ measure\ n$

$S_n^{hi} = Relative\ \boldsymbol{score}\ of\ Sysem\ efficiency\ measure\ n$

$C_n^{hi} \triangleq 1 : If\ Sysem\ efficiency\ measure\ \boldsymbol{n\ is\ chosen}, 0 : Ohterwise$

Goal attainment:

For the measure of Goal attainment two desired considerations have been identified; the first being goodness and the second fairness. The three main goals in goal attainment are: improving health, enhanced responsiveness to the expectations of the population and assuring fairness of financial contribution. Combining improving health, enhanced responsiveness to the expectations of the population, and assuring fairness of financial contribution with goodness and fairness, the following measures have been created:

Improving health:

- *Level of health*
- *Distribution of health*.

Enhanced responsiveness to the expectations of the population:

- *Level of responsiveness*
- *Distribution of responsiveness*.

Assuring fairness of financial contribution:

- *Fairness of financial contribution.*

The reasonable level of financial contribution is not considered in the measure of goal attainment. However, the distribution thereof is one of the major components used by the WHO in the ranking of health systems is per capita expenditure. *Level of health*: The health indicator used to measure the level of health is disability-adjusted life years (DALY). With the use of DALYs the average number of healthy living years is used to compare the levels of health. Japan ranked 1st in this indicator.

Distribution of health: The equity of child survival was chosen to represent the distribution of health within a population. Chile ranked 1st according to this indicator.

Level of responsiveness: Using around 2000 subjects from selected countries, the level of responsiveness of the health system was measured. The USA ranked 1st according to this indicator.

Distribution of responsiveness: The distribution between the "level of responsiveness" indicated by the 2000 subjects were used for this measure. The United Arab Emirates ranked 1st according to this measure.

Fairness of financial contribution: The fairness of financial contribution is calculated based on the proportion of permanent income per household. Colombia ranked 1st according to this measure [2] [26].

Health expenditure per capita:

The component of health expenditure per capita is relatively self-explanatory. The level of expenditure as a portion of income per household is measured and averaged across the population. The USA ranks as the country spending the most on health care by a great margin.

Efficiency and overall level of health performance.

Coyne and Hilsenrath define the efficiency calculated by the WHO as follows [26]:

$$HS_e = {(DALE_o - DALE_{wo})}\Big/{(DALE_m - DALE_{wo})}$$

The Health system efficiency (HS_e) is calculated using measures of DALY's. The overall performance of the system is account for by using DALY's.

$DALY_o$ is the observed aggregated DALY's for a given health system

$DALY_{wo}$ is the calculated DALY's that would have been experienced without a functioning health system in the given population.

$DALY_m$ is the maximum possible DALY's that could have been experienced given the level of expenditure of the health system in question.

The intensive use of DALY's to measure not only the efficiency of the health system but also the level of health received some critique and indicates the limitations of using a composite measure to assess and compare health systems and their performance.

Organisation for Economic Co-operative Development:

Included in the convention that was signed by the founding states of the OECD in Paris on the 14 of December 1960 was an indication of the importance of a healthy society for economic growth: "good health is essential for people to flourish as citizens, family, members, workers and consumers" [27].

Compared to the WHO's framework, the conceptual framework proposed by the OECD has more of a focus on the health care system performance than the public health [28]. According to the OECD the goals of a health system performance framework are: health improvement and outcomes; responsiveness and access; financial contribution and health expenditure. Each of these goals is measured on two areas, namely average level and distribution. The six components used to measure the performance of a health system are as follows:

- Health improvement and outcomes average level.
- Health improvement and outcomes distribution.
- Responsiveness and access average level.
- Responsiveness and access distribution.
- Financial contribution and health expenditure average level.
- Financial contribution and health expenditure distribution.

Compared to the use of DALY's by the WHO, the OECD publishes the following health indicators. The 10 main indicators found in the database are [28]:

- Health status
- Health care resources
- Health care utilization
- Expenditure of health
- Financing and remuneration
- Social protection
- Pharmaceutical market
- Non-medical determinants of health
- Demographic references economic references

Arah *et al*. reviews four national frameworks with the WHO and OECD. From this review the performance indicators used by each organisation are compared in the following table.

International body	Main goals of health system (framework)	Health system performance components	Dimensions of performance	Main data fields used
OECD	Health improvement and outcomes	Health improvement and outcomes-average	Health improvements/ outcomes	Health status
	Responsiveness and access	Health improvement and outcomes-distribution	Responsiveness	Health care resources
	Financial contribution and health expenditure	Responsiveness and access-average	Equity(outcomes, access & finance)	Health care utilization
		Responsiveness and access-distribution	Efficiency	Expenditure on health
		Financial contribution and health expenditure-average		Financing and remuneration
		Financial contribution and health expenditure-distribution		Social protection
				Pharmaceutical market
				Non-medical determinants
				demographic references
				Economic references
WHO	Improving health	Improving health-average		DALY's
	Enhancing responsiveness to expectations	Improving health-distribution		Child survival
	Fair financial contribution	Enhancing responsiveness to expectations-average		Survey
		Enhancing responsiveness to expectations-distribution		Survey
		Fair financial contribution-distribution		Per Capita Expenditure

Table 2: OECD and WHO measurement comparison

3.2.1 *Elements describing the health systems*

By comparing the data field used by the WHO and the OECD, a picture can be drawn of the elements, each organisation finds relevant to define the performance of the system. Based on the assessment of the two organisations 8 generic system components are identified

WHO	OECD	Generic system component
DALY's, Child survival	Health status	*Clinical*
Per Capita Expenditure	Health care resources, Expenditure on health, Financing and remuneration	*Financial*
Survey, DALY's, Child survival,	Health care utilization	*Operational*
	Social protection	*Social*
	Pharmaceutical market	*Pharmaceutical*
	Non-medical determinants	*Non-medical*
DALY's , Child survival	Demographic references	*Demographic*
	Economic references	*Economic*

Table 3: System component comparison

The data used by the WHO and the OECD can be grouped to cover 8 system components or areas influencing the health systems of the world. These components or elements describing health systems are a benchmark for elements, or classes of element used at the highest level of modelling. At lower levels, the scope of the model will determine if all of these elements will be used or not.

3.2.2 *Validation of 8 system components*

To validate the 8 system components or main elements a Delphi-panel review of system elements was done. The aim of the validation was to collect a number of elements identified through a formal process and match those elements to the 8 identified classes of elements. This comparison will then serve as an indication of how well the identified classes of elements represent the reality as perceived by expert opinion.

The panel of experts represented the following areas of health care: Private practice and Management, Academic Institutions, Government as well as Non-Government role-player. The participants were briefed on the purpose of the survey as well as the background of the research. Each of the participants was contacted by telephone and then the survey was sent to them in an email containing the necessary supporting documentation. All of the supporting documentation can be found in the Appendix.

The process of the Delphi panel review was conducted throughout three phases:

1) Identification of elements describing the health system
2) Choose the 10 most influential elements
3) Rank the 10 chosen elements

Step1

The initial list of 10 elements was identified by each of the participants. All of the responses were collated to form the following master list of elements:

- Available equipment and resources.
- Budget allocation.
- Demographics of population.
- Disease prevention.
- Distribution of Health care services.
- Drug supply.
- Effectiveness of Health care services.
- Epidemiology.
- External trends influencing the system.
- Funding mechanism: Out of pocket, Tax and Insurance.
- Health care education to general population.
- Health Care system Model: Free market in South Africa.
- Health measuring.
- Human Resource training.
- Human Resource utilization.
- Human Resources competence: Clinical.
- Human Resources competence: Management.
- NGO involvement.
- Patients.
- Philosophy of health.
- Quality Assurance of health care services.
- Regulation of Health care services.
- Salaries to health care professionals.
- Service delivery structure: Referral protocol.
- Support services.
- Transportation to and from health care services.

Step2

In the second step the complete master list of elements was circulated to the participants. They were requested to identify the 10 most important elements from the list.

From the responses the following ranking of elements was received:

Element	Percentage of responses choosing element as top 10
Human Resources competence: Management.	83.3%
Health Care system Model: Free market in South Africa.	66.7%
Quality Assurance of health care services.	66.7%
Regulation of Health care services.	66.7%
Distribution of Health care services.	50.0%
Epidemiology.	50.0%
Funding mechanism: Out of pocket, Tax and Insurance.	50.0%
Human Resources competence: Clinical.	50.0%
Patients.	50.0%
Available equipment and resources.	33.3%
Demographics of population.	33.3%
Disease prevention.	33.3%
Effectiveness of Health care services.	33.3%
Human Resource training.	33.3%
Human Resource utilization.	33.3%
NGO involvement.	33.3%
Budget allocation.	16.7%
Health care education to general population.	16.7%
Health measuring.	16.7%
Philosophy of health.	16.7%
Salaries to health care professionals.	16.7%
Service delivery structure: Referral protocol.	16.7%
Support services.	16.7%
Drug supply.	0.0%
External trends influencing the system.	0.0%
Transportation to and from health care services.	0.0%

Table 4: Ranking of elements

Step3

Using the results of the previous step the following elements were identified to be used in the final step:

- Burden of disease.
- Clinical competence.
- Demographics of population.
- Disease prevention.
- Distribution of health care services.
- Effectiveness of health care services.
- Funding mechanism.
- Health care system model (Free market).
- Managerial competence.
- Patients.
- Quality Assurance of health care services.
- Regulations governing health care services.
- This list of elements was distributed to the participants

This list of elements was circulated to the participants to rank the elements according to importance. From the feedback the following ranking of relative importance was received.

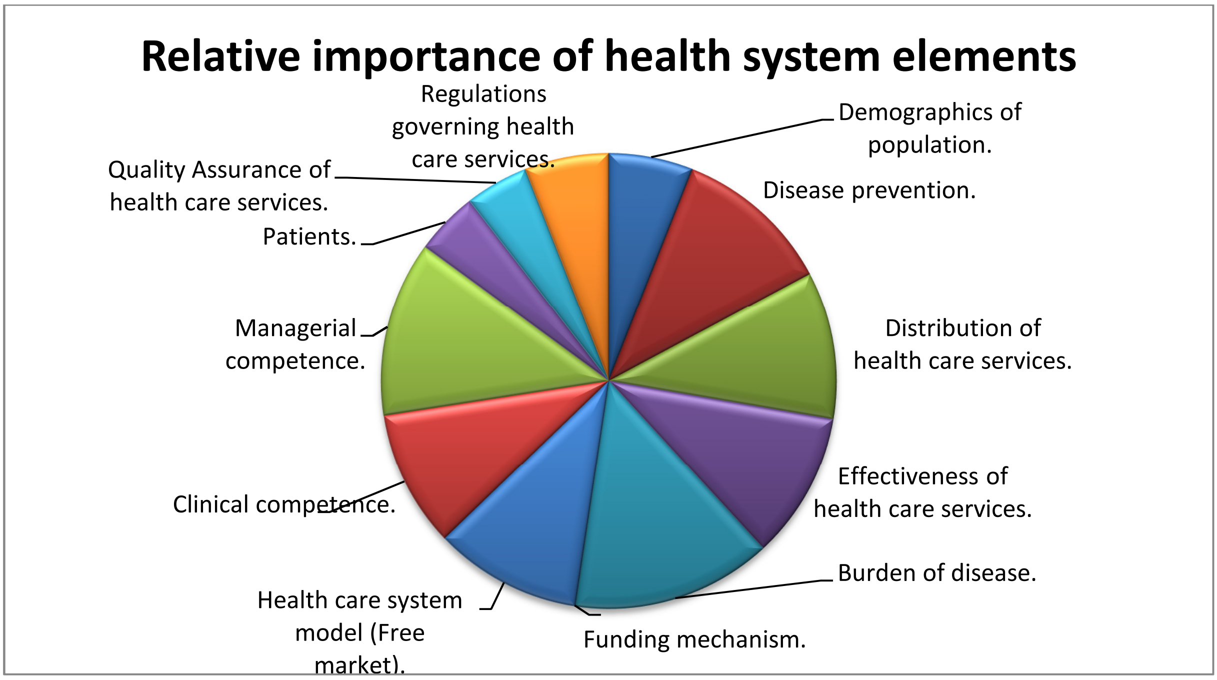


Figure 1: Relative importance of elements

By showing the most important elements in bold, and the lesser important elements not in bold, and how they fit into the classes of elements the following table is populated:

Class of element	Identified element
Clinical	**Disease prevention** **Burden of disease** **Clinical competence** Patients
Financial	**Health care system model (Free market)**
Operational	**Effectiveness of health care services** **Managerial competence** Quality Assurance of health care services
Social	**Distribution of health care services**
Pharmaceutical	-
Non-medical	Regulations governing health care services
Demographic	Demographics of population
Economic	Funding mechanism

Table 5: Element classification

It is clear that there is a good fit of the elements into the classes of elements. The Pharmaceutical classes of elements are represented to a lesser degree in the identified elements. This could be due to the pharmaceutical aspect of a health care system often being tied into the clinical and operational components of care and not necessary seen as a self-standing element.

Based on the outcomes of this review it is assumed that the elements of a health system can be represented by the following 8 classes; Clinical, Financial, Operational, Social, Pharmaceutical, Non-medical, Demographic and Economic.

Chapter 4

FRAMEWORK FOR DEVELOPING A SYSTEM DYNAMIC MODEL FOR A HEALTH SYSTEM

4.1 CHAPTER OVERVIEW

In this chapter a framework is suggested that can be used for modelling health care systems. The traditional frameworks used in system dynamic modelling are used as a starting point with specific considerations taken into account to make the framework relevant for health system modelling. Each of the steps in the framework is discussed with specific reference to previous work in this document. To measure the effectiveness of the framework a criteria will be used. The logic behind the criteria is also discussed and how the criteria will be implemented.

The modelling guide is the product of distilling the framework into a form that is clean and unambiguous to use. The use of this guide is illustrated in the next chapter.

4.1.1 *The traditional framework for creating a SD model:*

According to Randers most SD models are created in four stages, each including a number of steps. These stages with the steps are outlined below [17]:

1) Conceptualization
    a. Define the purpose of the model
    b. Define the model boundary and identify the key variables
    c. Describe the behaviour or draw the reference modes of the key variables
    d. Diagram the basic mechanisms, the feedback loops of the system
2) Formulation
    a. Convert feedback diagrams into level and rate equations
    b. Estimate and select parameter values
3) Testing
    a. Simulate the model and test the dynamic hypothesis
    b. Test the models assumptions
    c. Test model behaviour and sensitivity to permutations
4) Implementation
    a. Test the models response to different policies
    b. Translate the study insights to an accessible form

The suggested approach will be guided by the traditional framework for creating a SD model but will be aligned to the health care environment by customising it according to the information in the previous chapters.

4.1.2 *Framework Measurement criteria*

Taking the information provided in previous chapters on system modelling and health indicators, and performance measures this approach is developed to; use system dynamic modelling (guided by relevant health indicators) as a tool to identify and assess interventions aimed at improving system performance.

The purpose of this approach is to aid decision makers, in the health care environment, to make informed decisions using an SD model by providing practical guidelines.

To test the appropriateness of the suggested approach a criteria will be used as a reference for the measurement. Thereafter a test of the framework will be done in the form of a case study.

The approach will be considered successful if the following criteria are met:

1. Provide a formal structure to follow when the effect of changes in the system are assessed
2. The approach should address the practical challenges found in the health care environment
3. The approach should be generic enough to be applied at different levels of analysis

The weighting of the criteria was chosen by using a comparison matrix indicating the relative importance of each criterion.

- Formal Structure: W_{FS}
- Health Care Specific: W_{HCS}
- Generic across levels of analysis: W_{GA}

Relative Importance	W_{FS}	W_{HCS}	W_{GA}
W_{FS}	1	1.5	3
W_{HCS}		1	2
W_{GA}			1

Table 6: Weight matrix

By solving the matrix using simple matrix algebra, the following weights are assigned to each of the criteria:

Criteria	Weight
Formal Structure	0.454
Health Care Specific	0.363
Generic across levels of analysis	0.181

Table 7: Solved weight matrix

After the framework has been applied in Chapter 5 this weighting used to measure how well the framework is aligned to the set goals according to the weights.

4.1.3 Iterative nature of modelling

The iterative nature of the modelling approach is an essential part of the framework. By going through the different stages of developing the model repeatedly, a deeper understanding of the modelled environment is gained. With the increased understanding of the problem and associated model the initial model assumptions might be altered, these assumptions in turn have an impact on the rest of the model and how it was build. As a result, the rest of the model would also have to be reviewed according to the new insights gained.

The diagram below gives an illustration of how the phases of modelling fit together, and how the changes in understanding can influence the model.

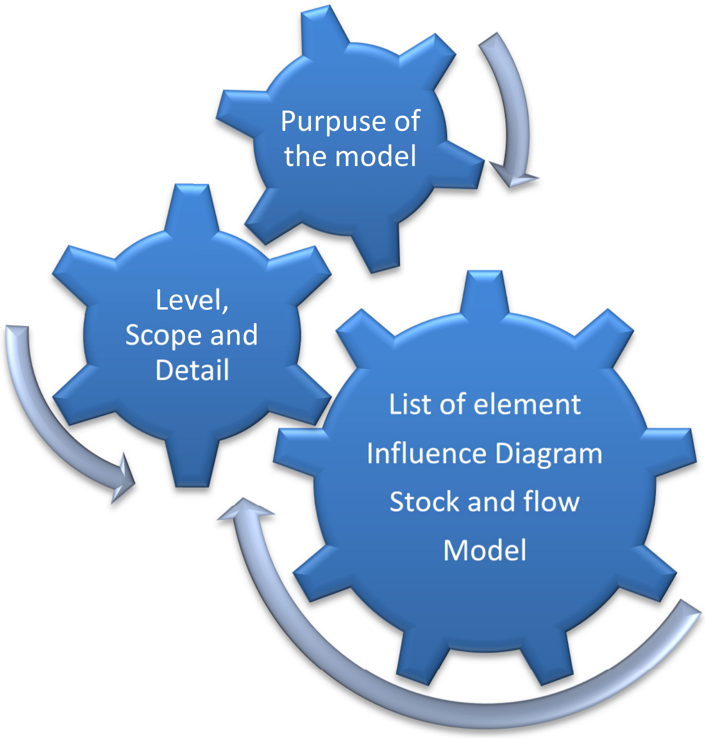


Figure 2: Phases of modelling

One of the mechanisms by which the process of modelling adds value to the outcome of the exercise is by creating a deeper understanding of the driving factors, and elements playing a role on the relevant environment.

This fact further augments the need to have an approach that, due to its practical nature, lend itself to aid in a fast turnover of model iterations, and as a result creates more understanding of the operating environment in a shorter period of time.

4.1.4 MODELLING ASSUMPTIONS

The suggested framework is based on some assumptions of the users modelling knowledge: It is assumed the modeller has:

a) A basic understanding of System Dynamics modelling
b) The mathematical competence to formulate equations and solve dimensional miss alignment.

In the following section the framework will be described. The aim of describing the steps in the framework is to motivate the inclusion of the each of the steps in the approach, and give insight to the value the step will add to the model, and exercise as a whole.

4.1.5 THE FRAMEWORK

The following framework is suggested to be used as a guide for the development of a health care performance model. The 12 steps described in the framework are of an iterative nature; none of these steps are intended to be cast in stone after they have been completed the first time. Any of the outputs of the steps should be open to revision as the model progresses.

After the framework is provided the steps described in the framework is condensed into a guide that can be used (once all of the steps are understood) as a practical and implementable tool to aid modelling.

1. **Defining the purpose of the model**

   "A system dynamics model is built to understand a system of forces that have created a "problem" and continue to sustain it. To have a meaningful model, there must be some underlying problem in a system that creates a need for additional knowledge and understanding of the system" [29]

The first step in the approach is to define the purpose of the model. The purpose of the model should be aligned to, and based on the problem that is being investigated.

Defining the purpose of the model could arguably be the most important step in the process of developing a model. The reason for this importance is because so many of the following steps are guided by, and based on the purpose of the model.

2. **Define the appropriate modelling level**

In the second step, the level at which the model will represent reality, or rather the health system is defined. The chosen level should be based on the purpose of the model. If the purpose of the model is to answer a question at a national level the modelling should reflect this high level of representation. A model that will be used to analyse the capacity of a facility, district or sub-district the level of the modelling will be much lower. The level of modelling does not reflect the level of detail of the model, but rather the level of representation. The level of modelling detail will be discussed in later steps.

3. **Scope and detail of the model**

Based on the purpose of the model, the scope and appropriate level of detail to be used can be defined.

Scope: The purpose of the model will drive the scope of the model along with the resources available, and the areas that can be influenced.

According to the review of the indicators used by the WHO and OEDC there are 8 classes of elements that can be used to describe health systems are:

- Clinical
- Financial
- Operational
- Social
- Pharmaceutical
- Non-Medical
- Demographic
- Economic

By using the 8 identified classes of elements the scope of the model can be clearly defined. Assuming the 8 classes represent all of the factors influencing any health system a choice of any number of these factors will define the scope of the model.

To further clarify the scope of the model Albin suggest the list of elements should be split into two categories, endogenous and exogenous [29]. Endogenous elements will influence the elements in the model, and will be influenced by the feedback loops and the behaviour of the model. Even though exogenous elements are modelled, their behaviour is not influenced by the behaviour of the model and sits outside the scope feedback influence of the model. Later when the list of elements is populated, it is essential to refer back to the scope of the model when the elements are grouped to inform the grouping, and ensure the scope is well defined.

Amount of detail to model: The amount of detail to model is guided by the purpose of the model, and based on the practical implications of resources and data availability.

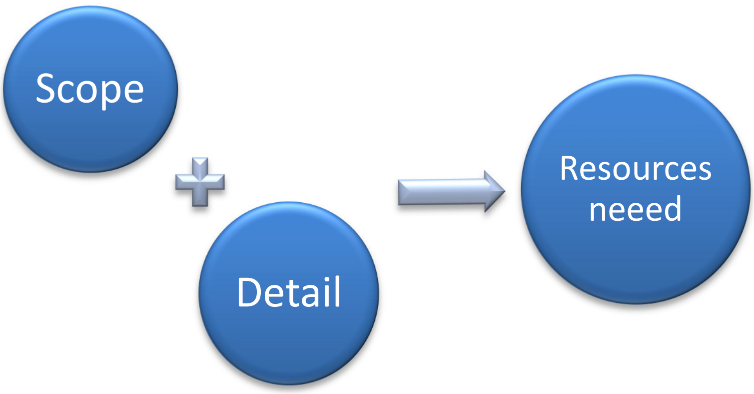


Figure 3: Scope, detail relationship

With a limited amount of resources a trade-off will have to be made between the scope of the model and the amount of detail that is modelled.

The relation between available Resources (Rs), resource need to model the Scope (Sc) and resource need to model the Detail (Dt) can be mathematically represented in non-dimensional terms as follows:

$$Rs \leq Sc + Dt$$

If only a limited amount of resources are available the scope will have to be decreased if the detail of the model is increased. If both the scope and detail of the model increases, the resources needed to build the model will have to be increased.

4. Defining the classes of elements

The elements of the system are the components of the system, but they should not to be limited to physical or tangible components. The elements of the system also include the non-tangible factors and components, the processes found in the system and the policies or rules governing the operation of the system.

The elements of the system will include all of the "topics" you will mention to someone in an attempt to describe the system in its entirety to someone that is not at all familiar with the system.

The elements describing a health system can fit into the 8 classes of elements previously defined:

- Clinical
- Financial
- Operational
- Social
- Pharmaceutical
- Non-medical
- Demographic
- Economic

Based on the scope of the model a number of these classes could be omitted thus they will, or will not be considered when the elements are chosen for the model. If during the modelling process there is a recurring need to include elements from classes not within the scope, the scope of the model should have to be reviewed and aligned with the purpose of the model.

5. Identification of elements

Traditionally the identification of the elements describing the system is a qualitative process relying heavily on the knowledge of the modeller, team or advisors. In an environment where experts in the field are easily accessible to the modeller, this approach works very well.

In an environment where there are relatively few people with modelling and operational knowledge of a health system there is a greater need for the modeller to be guided in the choice of elements.

Thus it is proposed that the mechanism used to guide element identification should be health indicators. This approach can be combined with brainstorming sessions, or a Delphi-panel review to provide further validation of the content that is modelled.

Using health indicators as the guiding mechanism to identify elements has the following significant advantages:

1) The choice of elements is not only dependent on the modelling group but based on pre-determined factors that describe the system's performance.
2) The elements that are chosen will be easier to model and validate using existing data that should be available for the indicator.
3) Reporting of the findings will be easier communicated due to familiarity to the people working in the environment with the indicators or measures.

Identification of the indicators that are going to guide the elements is dependent on the purpose of the model, along with the chosen scope and the level of detail used. For a high level model "high" level indicators must be used. These indicators are usually systematic indicators, and aggregated across the lower tiers of the system.

The lower level indicators will represent operational aspects of the system and will be appropriate to use for the modelling of a lower level system.

The factors Eddy identifies as considerations for a good measure can also be applied to the choice of measures to be used in the modelling. The factors include:

- Purpose
- Entity
- Dimension
- Type
- Audience

The indicators chosen to create the list of elements can be guided by the factors as follows:

Purpose: The purpose of the indicator should be aligned with the purpose of the model.
Entity that is measured: The indicator is in place to measure an entity. This entity can range from the throughput of a rural clinic to the performance to a hospital system in an urban area. By aligning the entity that is measured by the level of detail of the model the indicator will translate into relevant elements.
Dimension: The dimension of the indicator refers to the view or perspective of the indicator, indicators measuring quality can view quality from a patient, operational or community perspective. Again the dimension of the indicator used should be aligned with the purpose of the model, and the perspective used in the purpose definition.
Type: The types of measures defined by Eddy include Population, Process, Biological and Outcome measures. The scope of the model will guide the type of measures that can be used to identify elements. If Clinical class of elements are not in the scope of the model biological measures will most likely not be used.

Intended audience: The intended audience of a measure will have a close relation to the level of the model. For a high level model the intended audience might sit at the National Department of Health, in that case high level indicators (Population or Outcome) will guide the choice of elements. On an operational level process indicators will translate into elements providing valuable insight to the modeller.

6. Consider characteristics of elements

The list of elements will be used to create the influence diagram; the influence diagram will be translated to a stock and flow diagram that is used for the modelling. This flow of the information should be known because certain characteristics are needed from the elements to ensure consistency across the full process.

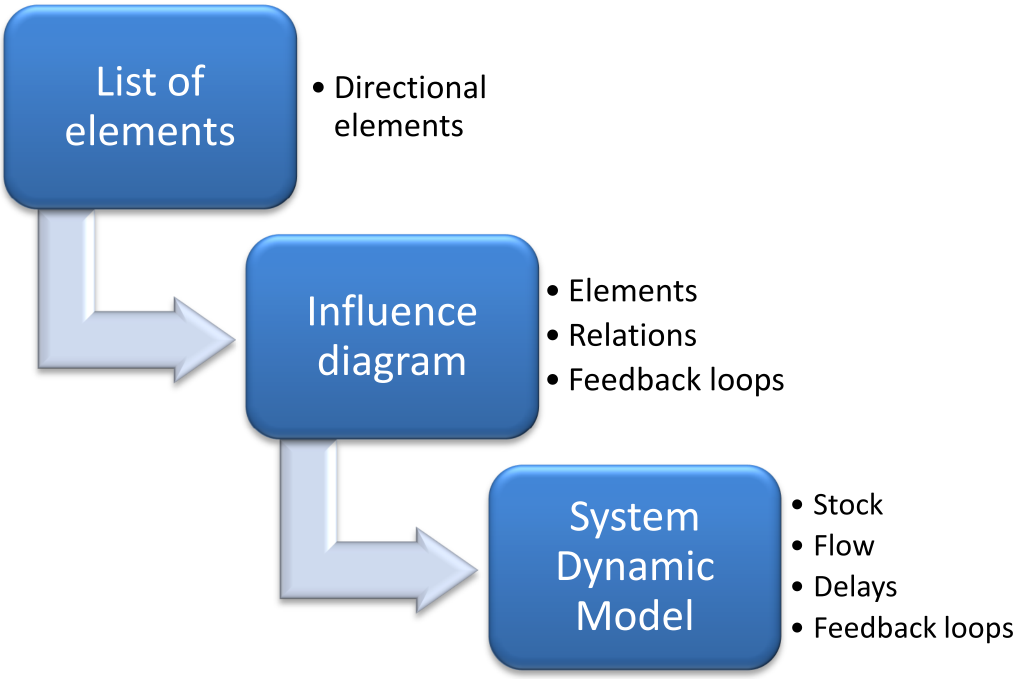


Figure 4: Process of modelling

Elements should be *unambiguous*. A clear working definition should be available for each element to eliminate confusion and miss communication. This factor is very important, especially in a team of cross-functional participants where a common frame of reference might not be available.

The elements should be *directional*. The reason to have this characteristic is to have mathematical consistency across the influences in the influence diagram. Directionality is the property of having a direction or quantity that can increase or decrease. Kirkwood highlighted the need to use nouns or phrases rather than verbs to describe elements [30]. When the influence diagram is drawn the links will represent the direction of change.

Elements do *not have to be dimensional*. An element could be dimensional, like capacity (measured in units/hour) but an element such as moral (measured on a scale from high to low with no dimension) could also be used as a non-dimensional element.

To illustrate the above mentioned considerations an example will be used.

Human Resource is a factor that is frequently used. If an element is named *Human Resource*, it is neither directional nor unambiguous. To have an unambiguous element a better name will be *Human Resource Moral*, *Human Resource Capacity* or *Staff Attitude towards patients*. The first two suggested names also adhere to the directionality criteria: *Human Resource Moral* can increase or decrease as can *Human Resource Capacity*. *Staff Attitude* towards patients is un-directional element; *Staff Attitude* cannot increase or decrease. The correct naming convention would be *Staff helpfulness* or *Staff Compassion*. Both of these terms can increase or decrease, and thus they are directional. *Increasing Staff Compassion* should be avoided as an element, thus increasing or decreasing will be indicated by the links and relations.

7. Endogenous and exogenous grouping of elements

Grouping of the elements into the two categories will give further clarification when the influence diagram and the model are created. The grouping of the elements is highly dependent on the scope of the model and should be referred to when done.

The steps; defining the scope and identifying the elements are highly iterative, even to such a degree that it might feel like one-step in the approach. As these steps are undertaken, the scope or more likely the level of detail of the model will change when new elements are added or removed from the list.

8. Create an influence diagram

The influence diagram gives a representation of the influences or causes the elements will have on each other. The purpose of the influence diagram is to start to gain a deeper understanding of how the elements interact with each other and how the system operates. Often causal loop diagrams are also used in the qualitative part of SD modelling but in this dissertation it was decided to only use the terminology of influence diagrams. The reason to not use casual loop is based on the fact that you could argue causal loop diagrams should only depict complete loops, all influences needs to be shown in this part of the framework thus influence diagrams are used.

Kirkwood developed a set of hints for the development of influence diagrams using Pugh and Kim's guidelines [31] [32]. Adapting the set of hints to the health care modelling environment yields the following steps:

1) Start by indicating influence of the elements on each other. If one element has a direct influence on another element connect the two with an arrow pointing in the direction of the cause. If indirect influences are known, the indirect paths should be

replicated as long as it is within the scope of the model, alternatively a direct influence should be considered, or a review of the scope of the model.

2) If influences are not clear between the two elements consider adding elements to ease the explanation of the relationship
3) The influence one element will have on another element should be indicated as positive or negative. When the change of the first element (up or down) causes the second element to change in the same direction (up or down) a positive relation exists. When the change causes the second element to change in the opposite direction a negative relation exists.

> *"Keep the diagram as simple as possible… The purpose of the diagram is not to describe every detail of the process, but to show those aspects of the feedback structure leading to the observed pattern of behaviour" Kirkwood*

9. Build the model representing the influence diagram

Translation of the influence diagram onto the model has two components. The first is the creation of the stock and flow diagram and the second is defining the equations governing the model.

The purpose of building the model is to create a mechanism to analyse the characteristics of the model in a quantitative manner.

Stock and flow diagram

The creation of the stock and flow diagram starts with the definition of the elements as stocks or flows. A stock is something with an associated quantity, or level at a given time, the water in a tank will be a stock element, an element of stock increases or decreases due to a flow. An element of flow changes/influences the element(s) of stock, the rate at which water is flowing into the tank (or out of) are elements of flow. As an illustration, if the flow of water into the tank is greater than the flow out of the tank, the "water" in the tank will increase.

According to Kirkwood most businesses has stocks of the types; materials, personnel, capital equipment, orders and money. [30] In the health care environment, the types of stocks can be adapted and defined as:

Materials: The stock or flow of goods used in the process of care, or people/patients in the process of being treated or in a specific state of health.

Personnel: The actual people represented in the system that are providing services. Clinical and non-clinical personnel will form part of this type of stock.

Capital equipment: Includes elements like hospital beds, clinical and diagnostic equipment, all the equipment needed to provide care to the patient.

Orders: This type of stock refers to things like drugs on order, an "order" of new personnel, material or capacity. Orders are the result of a decision to address something in the system that is "not right", by requesting or making an "order" to address the difference between the "not right" state and the "right" state.

Money: Representing the stock of money and how it changes according to the behaviour of the model.

The flows between the stocks of the model are the movement of entities over time. A flow is usually measured as entities over time or units/time. If time is frozen there will be no flow in the model, but still the stocks will have values.

Defining Equations

The mathematical equations are going to define operation of the model. Each equation will be a function of the inputs/elements influencing the element. When a stock is modelled, the level of the stock at any time will be the difference between the flows into and out of the stock over the time, the model has been running.

$$Stock\ level = \int_{Starting\ Time}^{Current\ Time} [\sum flow\ in\ (t)\ - \sum flow\ out\ (t)]\mathrm{dt}$$

The equations defining the flows in the model will be a function of the elements influencing that element. The equation defining a flow will have a time component associated with it because flows are represented as rates.

$$Flow\ rate = \int F(input\ 1, input\ 2, input\ N)dt$$

Dimensions of elements and equations:

When equations are being defined the dimensions used has to be consistent. Without consistent dimensions, elements cannot be mathematically compatible. Often the influence diagram does not make provision for consistent dimensions. In such a case, axillary variables are used to aid the translation from one dimension to another.

10.Model Validation

Validity: "adequacy with respect to a purpose" [33]

Validation of any modelling based activity is a highly important aspect of the modelling exercise. The validity of the model measures or indicates how strongly the model represents reality.

With the high reliance of SD on informal and subjective model building the validation of SD models has been a controversial topic [33]. In Arthur and Winch' review of the history of validation they argue that a lot of the controversy [34] [35] arose because of the narrower expectations and definitions of validity [36]. To clarify the expectation of validity from different points of view Barlas work is very helpful.

In an attempt to describe the formal aspects of model validation Barlas differentiates between "causal descriptive" (white-box) models and purely "correlated" (black-box) models [33].

The following table highlights some of the fundamental differences between causal descriptive and correlated models.

	Correlated Models (black-box)	**Causal descriptive model (white-box)**
Internal characteristics	No claim of causality	Internal causality represents reality
Focus/Importance to model	Aggregated output	The way systems operate in reality
Validity measured against	The degree to which the output represents reality	Internal explanation of reaction Characteristics of the output
Purpose of application	Forecasting	Behavioural analysis

Table 8: Model characteristics

If the direct outputs of one type of model, correlated or causal descriptive would be compared against the other, there could be a significant miss alignment of results. It is this misalignment or misunderstanding that could lead to controversy.

Barlas continues to show how the validation of causal descriptive models such as system dynamic models has strong ties to the philosophy of science. Where a system dynamics model is a depiction of a "theory" of how the system operates in reality, even though the outputs does not necessary represent reality. Herein lays the connection to the philosophy of science and how theories can be proven. The philosophy used to define a theory, or the model validity is then the point of reference that will influence the validity of the model.

Barlas and Carpenter summarises a comprehensive review of these philosophies [37] [38] [39] into two "camps"; the traditional reductionist/logistical positivistic camp and the opposing relativistic, holistic and pragmatist camp [40]. The traditional school of thought defines the model as either being correct or incorrect based on the empirical facts. From this point of view the validity of the model is based on accuracy of the output rather than the usefulness of the model.

The opposing camp does not define the model as correct or incorrect, but will define a valid model as one that describes the real situation or system. "Models are not true of false, but lie on a continuum of usefulness" [40].

System Dynamics approach is aligned to the relativistic/holistic philosophy. Model validation of SD models is a process of gradually building the confidence in the model to be "adequate with respect to its purpose".

Barlas summarises; the ultimate objective of SD model validation is to establish the validity of the structure of the model. The accuracy of the model behaviour's reproduction of real behaviour is also evaluated, but is meaningful only after we have sufficient confidence in the structure of the model [33].

With the focus of validity of SD models on the internal structure of the model rather than only the outcome or output of the model there is still a need for a formal structure of model validation.

The formal structure of model validation can be broken into three parts, where the first two tests the structure of the model and the third the output of the model.

1) Direct structure testing
2) Structure-oriented behaviour testing
3) Behaviour pattern testing

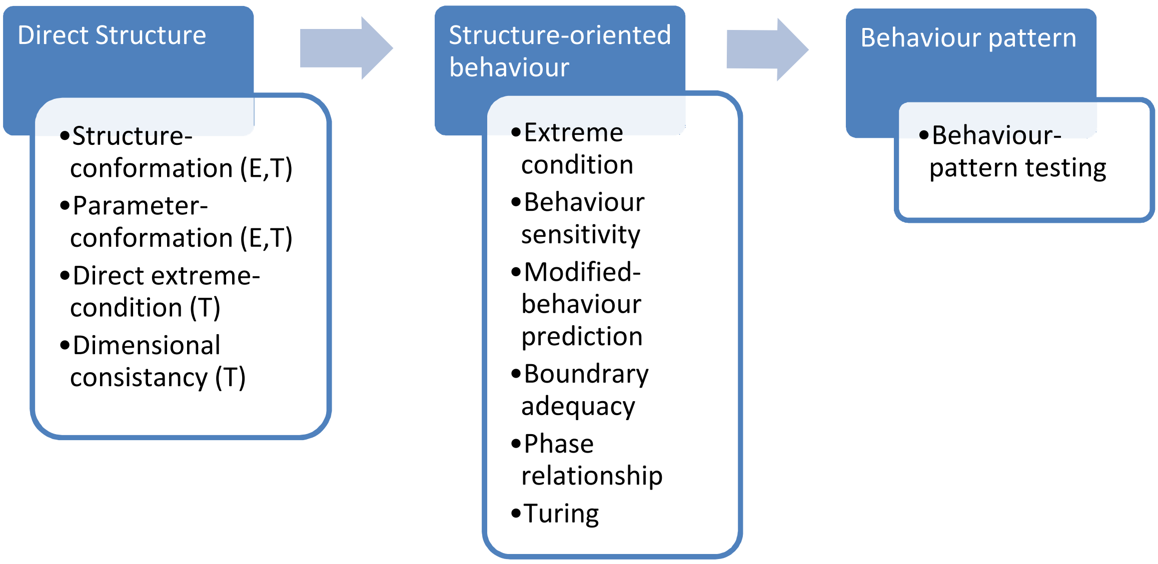


Figure 5: Process of validation

Direct Structure testing

Direct structure testing does not involve simulation, the testing is done in the static state of the model. Only the individual equations describing relationships are tested. This part of testing validates the direct structure of the model. Each relationship of the model is compared against the real system. This comparison is done with the available knowledge of the system (theoretical) or empirical data (empirical) from the system.

Structure conformation test - Empirical: The mathematical equation describing the relationship is compared to the relationships present in the real system. Theoretical: The mathematical equation is compared to the general knowledge about the system.

Review, walkthroughs and formal inspections are a number of ways to formalise these tests.

Parameter conformation test - Empirical: Comparison of parameters of the model to those found the real system for certain elements. Theoretical: Comparison of parameters of the model to theoretical values for the system in reality.

Direct extreme condition testing: By testing the model equations under extreme conditions, the validity of an equation can be measured. The resulting output due to the extreme condition input is assessed against the expected or logical reaction of the real system. Zero and infinite are usually the inputs used for the extreme tests [41].

Dimensional consistency test: The dimensional consistency test checks that both sides of the equation has the same dimensions, and are thus mathematically plausible equations.

When enough confidence has been gained in the model through the direct structure tests the following part of testing can be done.

Structure oriented behaviour tests

The structure oriented behaviour tests validates the structure of the model indirectly by assessing the behaviour patterns of the model [41] [42]. Sub areas or components of the model can be tested during this phase of testing. The behaviour of the model or sub models is compared to that of the system being modelled. The degree to which the model represents the reality gives an indication of how well the structure, sub-structure represents reality.

A number of different types of structure oriented behaviour tests are mentioned by Barlas. A short description of each type is given below:

Sensitivity testing: The sensitivity of the model or sub-model to various parameters is tested by manipulating those parameters and seeing how big an influence the change has on the output. The sensitivity of the model to the parameters is then compared to the sensitivity that can be expected in reality.

Extreme condition or Stress testing: Extreme values are assigned to inputs of the model. The behaviour of the model or sub model as a response to the extreme values are then compared to the expected reactions of similar inputs to the system that is modelled.

Modified behaviour: If data is available to show how the real system responded to an intervention the response of the model to the same intervention can be compared to the real data. When the model generates similar behaviour to the real system, the test is passed.

Phase relationship testing: Comparing the relationship between different variables' phases of the model, to the expected relationship gives an indication of the structural validity of the model. Barlas and Forrester indicated that when the relationship of the phases does not correspond to the expected behaviour of the system, the structure should be reviewed.

Turing test: The test developed by Alan Turing can be used to validate the structure of the model [43]. An external party is asked to differentiate between the outputs or behaviour of the model, compared to the expected behaviour of the system according to an expert or historical data for a fixed set of inputs. If the external party cannot differentiate between the two sources of the output with statistical significance the test is passed.

The model should be tested using a selection of the above mentioned types of tests depending on the purpose of the model, resources available to the modelling team and the nature of the model.

When the structure oriented behaviour tests are complete, and the modelling team is confident in the structure of the model the final part of validation is addressed.

Behaviour pattern testing

After the structure of the model has been validated the behaviour of the model as a whole can be tested. The behaviour pattern testing depends on the types of behaviour exhibited by the model. The model can exhibit either transient behaviour or long-term steady state behaviour. Different validation approached is appropriate to use for transient or steady state behaviour.

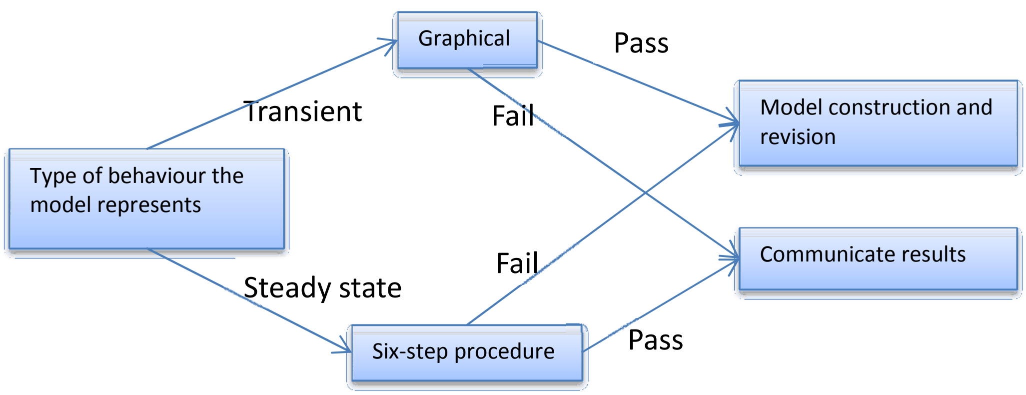


Figure 6: Behaviour pattern testing

Transient Behaviour: Because transient behaviour usually involves patters such as s-shaped growth curves or boom then bust patterns there are no statistical reference to measure the behaviour against [33]. Thus graphical measurements are used to perform pattern testing [41]. The use of graphical validation if further discussed by Carson and Flood, Forrester and Senge, and Barlas as a formal procedure [41] [42] [44].

Steady-state Behaviour: Based on technical and philosophical arguments the use of statistical significance testing is shown not to be useful for system dynamic model validation.

From a technical point of view one of the reasons statistical significance testing is not relevant to system dynamics models are due to the assumptions of data being i) serially independent ii) not cross-correlated iii) normally distributed. The first two of these assumptions are generally not valid due to the very nature of a system dynamic model [40] [45] [46] [47]. Further, the simplifications and technical difficulty to address these properties of system dynamic models are used as an argument against the use of significance testing for system dynamic model validation.

From a philosophical point of view, the “acceptance” or “rejection” of a hypothesis used in significance testing is not aligned to the holistic school of thought used by the system dynamics approach. By limiting the value or validity of the model to a binary accept or reject does not make provision for a continuum of use for the model. Further using the degree to which the model represents reality as a measure does not make provision to measure the model against the purpose of the model. The danger of focusing the validation of the model on the outputs of the model, or how the model’s outputs replicates the reality is that the model gets pushed to a black box model.

Even though statistical significance testing is assumed not to be a method of validation for the model behaviour, there are certain standard statistical measures that can be used to validate the behaviour of the model. If the validation of the model behaviour is appropriate for the modelling application Barlas’ six step approach can be used.

- Trend Comparison and removal
- Period comparison using autocorrelation function
- Comparing the averages
- Comparing the variation
- Testing for phase lag using the cross correlation function
- An overall summary measure using the discrepancy coefficient [42]

To conclude the validation process, the importance of using the original purpose of the model as a reference point for validation is emphasised.

> *“In the end the modeller or user must accept the model as valid if he/she thinks it is more beneficial to use the model than not to use the model.” [33]*

11. Implementation of the model

The model was developed according to a purpose. Now that the modeller(s) are confident with the validity of the model, it can be applied to serve its purpose.

If the purpose of the model was only to gain insight into the system, the implementation of the model could be by using it as a management flight simulator. The flight simulator will be used to show the implications of different scenarios or input parameters to the model, and thereby further the understanding of the system and its behaviour.

If the purpose of the model was to aid in the correction of undesirable system behaviour, the implementation of the model would be to try to identify how the undesirable behaviour could be mitigated or eliminated. Different policies, or operational approached can be tested to see if the behaviour would be altered.

12. Providing results and findings

Providing the results and findings of the model is the final step in the framework. The results of the analysis and implementation of the model should be provided in such a manner that relays the findings in a clear and concise manner. These findings should align to the original purpose of the model as well as resonate with the intended audience.

4.2 MODELLING GUIDE

Due to the comprehensive nature of the framework it does not lend itself to be used as a quick guide. The modelling guide was developed to ensure a practical and implementable tool is available to assist modelling in the health care environment while still providing the necessary guidance for successful modelling.

The guide is structured according to the logical flow of the framework, and includes all 12 steps discussed in the framework. Only the essential content of the framework is used in the guide, there is very little explanation of the steps in the guide. The guide focuses on the logical flow between the steps along with the information that is needed for each step

For each step three components are highlighted;

- The inputs used to complete the step
- The considerations to take before completing the step
- The output of the step

The steps are all aligned to the health care environment and include the specific suggestions discussed in the framework, such as the defined element classes and use of indicators as element identifies.

In the next chapter the guide is applied in a case study. To illustrate the usefulness of the guide it is directly applied to a real life problem.

System Dynamics Framework Guide

1. Purpose of the model

Based on: The problem that is being investigated

Considering: Purpose should be clear and unambiguous

Define the purpose of the model: ________________________________

2. Define the modelling level

Based on: The purpose of the model

Considering: Audience, data availability

Define the modelling level: _____________________________________

3. Define the modelling scope

Based on: The purpose of the model

Considering: Data availability, area of possible influence and resources available

Using: 8 Classes of elements:

Define the modelling scope:

Element class	In Scope
Clinical	
Operational	
Pharmaceutical	
Demographic	
Financial	
Social	
Non-Medical	
Economic	

4. Define the modelling level of detail

Based on: The purpose of the model

Considering: Data availability, model scope and resources available

Define the modelling level of detail:________________________________

5. Identify the elements of the model:

Based on: Health care Indicators

Considering:

a) Purpose of the model.

b) The factors identified by Eddy:

- Indicator Purpose
- Entity being measured
- Dimension of measurement
- Type of measure and
- Intended Audience

c) Characteristics of the elements:

a. Unambiguous
b. Directional and
c. Not dependent on dimensions

Identify the elements of the model:

Indicator used	Element	Aligned to the model purpose	Eddy's factors considered	Element has all of the characteristics mentioned in c)
Eg: Estimated Prevalence	Infected Population	Yes	Yes	Yes

6. Define elements as endogenous or exogenous

Based on: List of elements

Considering: Model scope and resources available

Define the elements as endogenous (End) or exogenous (Exo):

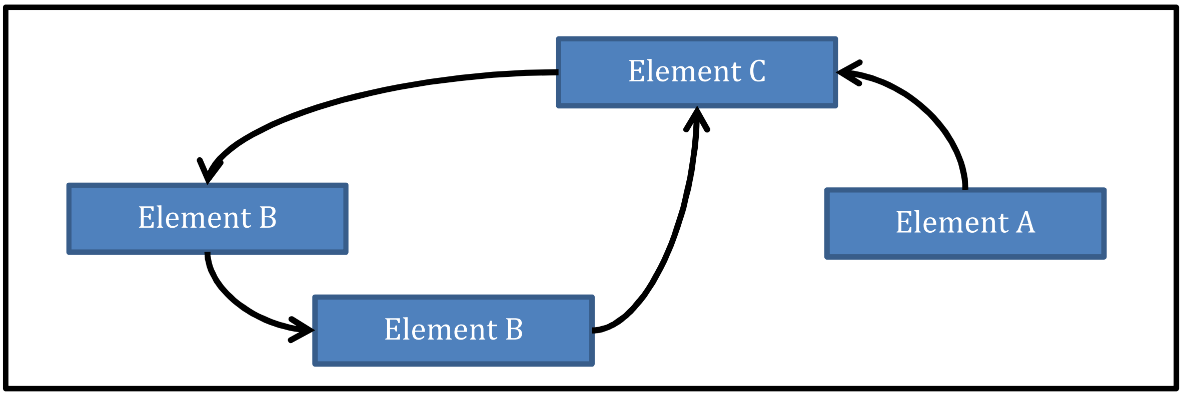
Element	End	Exo
E.g. Infected Population	X	

7. Create an Influence Diagram

Based on: List of elements

Considering: Direct influences elements and the relationship (positive or negative).

Create an Influence Diagram:

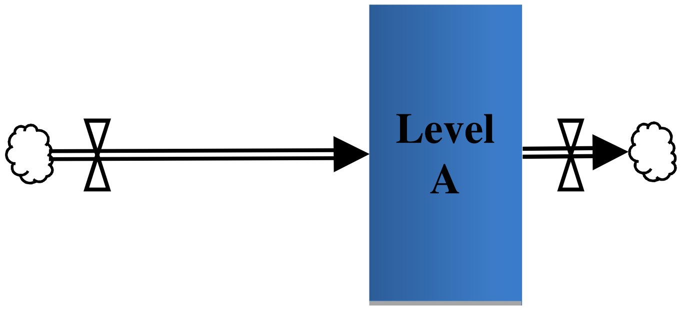


8. Create a stock and flow diagram

Based on: Influence diagram

Considering: List of elements, whether element is stock or flow.

Create a Stock and flow diagram:



9. Equations for Stock and flow diagram

Based on: Stock and flow diagram

Considering: Consistency of dimensions and auxiliary variables

Create equations for Stock and flow diagram:

(1) "Element A - Annual interest"= 0.06

Units: Dimensionless (Dmnl)

(2) "Element B - Balance"= INTEG (“Element C - Payment",20)

Units: Rands

(3) "Element C - Payment"="Element A - Annual intrest"*"Element B - Balance"

Units: Rands/year

(4) FINAL TIME = 100

Units: Month

The final time for the simulation

(5) INITIAL TIME = 0

Units: Month

The initial time for the simulation

(6) SAVEPER =

TIME STEP

Units: Month [0,?]

The frequency with which output is stored

(7) TIME STEP = 1

Units: Month [0,?]

The time step for the simulation

10. Model Validation:
Direct Structure Testing

Based on: Literature or Operational knowledge

Validation of: Individual equations describing relationships and model structure.

Consider the following tests

- **Structure conformation test**
- **Parameter conformation test**
- **Direct extreme condition testing**
- **Dimensional consistency test**

Model Validation: Structure-Oriented Behaviour Testing

Based on: Data, Literature, Operational knowledge or expert review

Validation of: Model or Sub-model structure

Consider the following tests

- **Sensitivity testing**
- **Extreme condition test**
- **Modified behaviour test**
- **Phase relationship testing**
- **Turing test**

11. Model Validation: Behaviour Pattern Testing

Based on: Data, Operational knowledge or expert review and Type of behaviour

Validation of: Model behaviour

Consider the following tests

- **Graphical**
- **Six-Step statistical tests**

12. Model Implementation

Based on: The purpose of the model

Considering:

Implement the model

Consider the following tests

- **Graphical**
- **Six-Step statistical tests**

4.2.1 *CHAPTER CONCLUSION*

In this chapter the information from previous chapters were combined to form a framework that can be used to assist in the modelling of health care system. This framework explains the process of creating a system dynamic model and taking specific considerations to the challenges found in health systems.

From the content of this chapter the Framework Guide was made to ensure the suggested framework is in an accessible form to facilitate use.

The intention of the framework is not to give a one-size-fits-all solution to health system problems, By giving direction on how the exercise of modelling should be approached, the framework aims to further illustrate when and where the application of system dynamic modelling is appropriate.

Chapter 5

CASE STUDY; PRIORITIZATION OF HEALTH CARE INTERVENTIONS

5.1 CHAPTER OVERVIEW

In the following chapter, the modelling guide described in the previous chapter is applied to a case study in the health care environment. The case study will serve a dual purpose a) it will be used to measure the success of the approach based on the set criteria, and b) it will be used to illustrate the use of SD models in the health care environment.

The criteria that will be used to define a successful framework are based on the following three capabilities:

1. Provide a formal structure to follow when the effect of changes in the system are assessed
2. The approach should address the practical challenges found in the health care environment
3. The approach should be generic enough to be applied at different levels of analysis

When the framework is measured a high score (/10) will indicate a good fit to the criteria, where a lower score will indicate a lesser fit to the criteria.

5.2 CASE STUDY DESCRIPTION: PRIORITIZATION OF INTERVENTIONS

Background: BroadReach HealthCare (BRHC) is a company providing assistance and support to various governments in the provision of health care. In South Africa BRHC is actively involved in supporting both the National Department of Health (NdoH) situated in Pretoria, and some provincial Departments of Health including; Gauteng, Mpumalanga, Kwa-Zulu Natal and the Eastern Cape.

The application of resources to strengthen the health care system can have far reaching positive impacts on many lifes, but when these resources are applied in areas of the health system that will not yield improvements in the system both the opportunity to improve is lost and the resources are wasted.

With limited resources available it is essential that the resources are used where there will be the greatest positive impact.

In the following section the suggested framework is used to guide the development of a model to assist in prioritising interventions.

From the framework a 12 step modelling guide was compiled [This guide can be found in the appendix]. To illustrate the process of developing the model the guide is filled in a step-by-step manner. The logic is also discussed at each appropriate step in the guide.

The framework was used to guide the modelling of the HIV continuum of care for a health district in Kwa-Zulu Natal. The purpose of the model was to assist in answering the question of where interventions should be focused to have the greatest positive impact on the number of people dying of HIV related causes. The modelling was done with the assistance of expertise from within BRHC and publically available information and current policies and guidelines.

The model was developed to represent the different stages for infected people (infected, status unknown, infected, status known, eligible for treatment, on treatment) and how the people flow between the stages. The mechanisms of infection that was modelled included sexual transmission as well as mother to child transmission.

Once a person is infected they do not know their status. The majority of these people are still in the disease's silent period. Only after the infected person has been tested and found positive the move from the state of "infected, status unknown" to "infected, status known is made. The silent period of HIV has a big influence on the rate of people going for testing; because the silent period is so long the symptoms of infection will only show long after infection and prompt the patient to seek medical attention, and get tested. The second factor influencing the rate of state change is voluntary testing. As voluntary testing is done before symptoms of disease appear there is a better prognosis.

When the infected person knows his/her status they can act accordingly by changing their sexual behaviour (in turn having an impact on further infections) and seeking anti-retroviral treatment (ART)

National policy governs the point at which a patient becomes eligible for ART. Recently the criteria have changed to allow a patient who is less advanced in the disease to be eligible.

The rate of which patients are initiated onto ART is governed by operational limitations and external factors. The aim is to have all eligible patients initiated onto ART as soon as possible.

The influences and interactions of the above environment were further expanded as the model was developed.

5.3 Intervention analysis case study

Problem definition: BRHC is experiencing challenges to choose areas of applying their resources and prioritise interventions because of the difficulty to identify the impact of interventions

This challenge can be broken up into two questions:

1) What area should be the focus of the intervention
2) How should interventions be prioritised

With a basic understanding of the background of the environment and the problem that is being used for the case study the steps suggested by the approach can be followed as illustrated in the modelling guide.

1. Purpose of the model

Based on: The problem that is being investigated

Considering: Purpose should be clear and unambiguous.

Define the purpose of the model: *The purpose of the model is to aid the decision of choosing and prioritising interventions in a HIV setting considering the infections of the UThungulu District in KZN*

2. Define the modelling level

Based on: The purpose of the model

Considering: Audience, data availability

Define the modelling level: *The model will be representing the reality at a district level.*

The reason for choosing district level as appropriate is due to the fact that interventions driven by BRHC are mostly at a district level and thus the usual operational level is district. This means the decision makers are familiar with the factors influencing operations at this level. Secondly, operational data from the District Health Information System is available at a district level.

3. Define the modelling scope

Based on: The purpose of the model

Considering: Data availability, area of possible influence and resources available

Using: 8 Classes of elements:

Define the modelling scope:

Element class	In Scope
Clinical	√
Operational	√
Pharmaceutical	
Demographic	√
Financial	
Social	
Non-Medical	√
Economic	

Table 9: Scope of elements

4. Define the modelling level of detail

Based on: The purpose of the model

Considering: Data availability, model scope and resources available

Define the modelling level of detail: *The model will not represent reality by a high level of detail, it will only represent the necessary level to illustrate the operational characteristics of the system to aid in the decision.*

The purpose of the model is not to give numerical predictions, but to aid in choosing and prioritising interventions, thus the indication of higher or lower results would suffice to fit the model to its purpose. Available resources will also limit the detail that can be effectively modelled.

5. Identify the elements of the model:

Based on: Health care Indicators

Considering:

a) Purpose of the model.

b) The factors identified by Eddy:

- Indicator Purpose,
- Entity being measured,
- Dimension of measurement,
- Type of measure and
- Intended Audience

c) Characteristics of the elements:

- Unambiguous
- Directional
- Not dependent on dimensions

Identify the elements of the model:

Indicator used	Element Name	Aligned to the model purpose	Eddy's factors considered	Element has all of the characteristics mentioned in c)
Number of persons tested HIV (excl ANC) - Total Number of persons tested HIV positive (excl. ANC) - Total	Infections	✓	✓	
Estimation	Infected population – Status unknown			
Number of persons tested HIV positive (excl. ANC) - Total	Infected population – Status Known	✓	✓	
HIV positive adult patient eligible for ART	Nr of People eligible for treatment	✓	✓	✓
Total started on ART new - during month	New people started on ART	✓	✓	✓
Total remaining on ART at end month (total in care)	People on ART	✓	✓	✓
Policy	Frequency of Treatment Management	✓		
Number of persons tested HIV (excl ANC) - Total	Testing for status	✓		
Clinical data	Disease infectiveness	✓	✓	✓
	Opportunities of infection			
Policy	Eligibility		✓	✓
OPD headcount - Total PHC headcount (total)	Capacity to treat people	✓		

Table 10: Indicator characteristics

6. Elements as endogenous or exogenous

Based on: List of elements

Considering: Model scope and resources available

Define the elements as endogenous (End) or exogenous (Exo):

Element	End	Exo
Infections	X	
Infected population – Status unknown	X	
Infected population – Status Known	X	
Nr of People eligible for treatment	X	
New people started on ART	X	
People on ART	X	
Frequency of Treatment Management		X
Testing for status		X
Disease infectiveness		X
Opportunities of infection		X
Eligibility		X
Capacity to treat people		X

Table 11: Endogenous vs. Exogenous

7. Create an Influence Diagram

Based on: List of elements

Considering: Direct influences elements and the relationship (positive or negative).

Create an Influence Diagram:

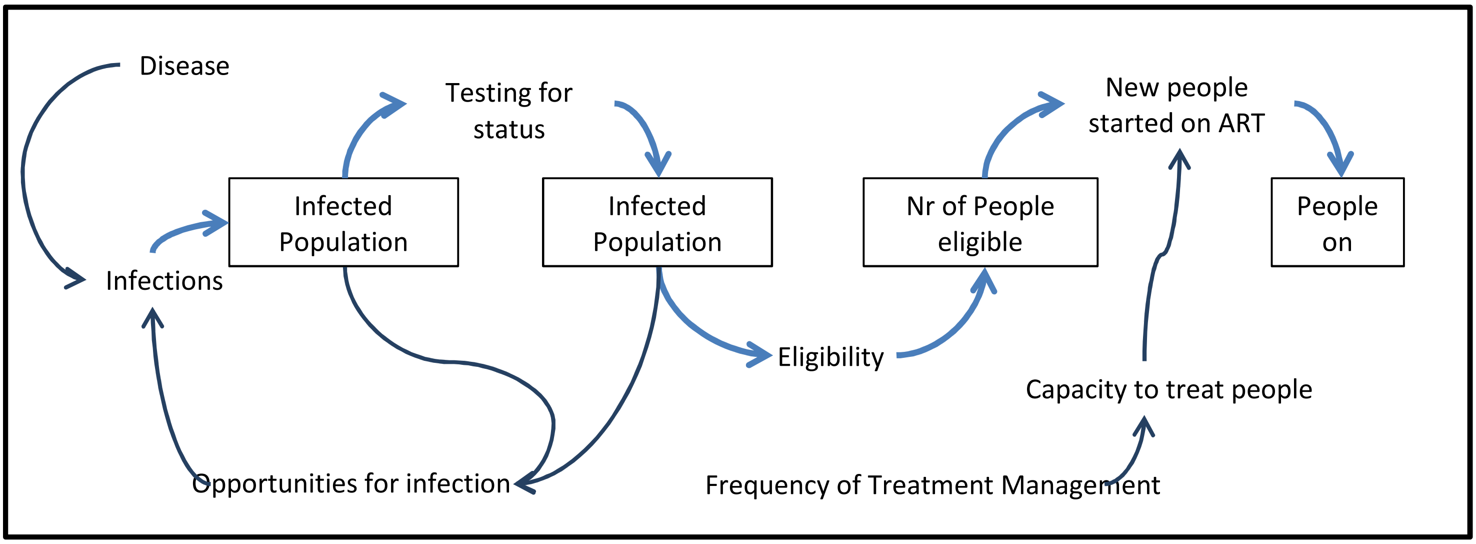


Figure 7: Case study; Influence Diagram

8. Stock and flow diagram

Considering: List of elements, whether element is stock or flow.

Create a Stock and flow diagram:

Based on: Influence diagram

[Figure on next page]

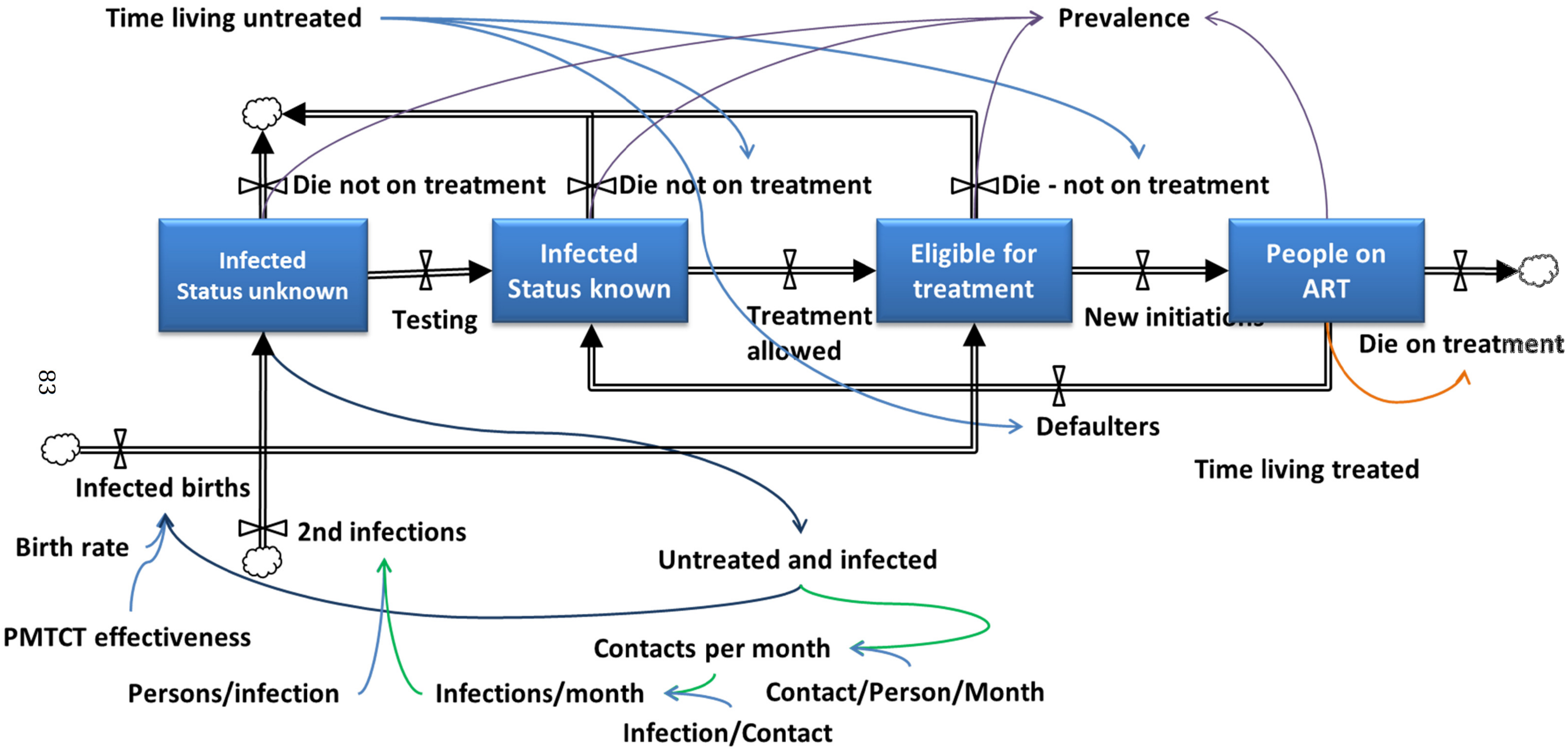


Figure 8: Case study; Stock and Flow Diagram

9. Equations for Stock and flow diagram

Based on: Stock and flow diagram

Considering: Consistency of dimensions and auxiliary variables

Create equations for Stock and flow diagram: All equations are described in the Model Validation section. For further reference each equation is numbered. This number corresponds to the model diagram and complete list of equations that can be found in the Appendix

10. Model Validation: Direct Structure Testing

Based on: Literature or Operational knowledge

Validation of: Individual equations describing relationships and model structure.

Consider the following tests

- **a) Structure conformation test**
- **b) Parameter conformation test**
- **c) Direct extreme condition testing**
- **d) Dimensional consistency test**

Model equations follow for each relationship. After the equations are defined the structural validation steps are indicated where appropriate:

The equations defining each relationship will be shown with the relevant validating for that equation. The stock and flow diagram will be used to illustrate what relationship is defined, further the equations are grouped logically according to patient flows and various data or information flows.

Flow of patients through the model:

The first section of the model that is discussed is the flow of patients between the states of an infected person in the model. The states include: Infected Status unknown, Infected Status Known, Eligible for treatment (Rx) and People on anti-retroviral treatment (ART). These four states represent the whole continuum for people living with HIV. Even though the complete representation of all the states of infected people is at a high level, it is necessary for the purpose of the model to represent each state. The states are modelled as levels. There are only four levels in the model, each indicating the number of people at that specific state of HIV infection.

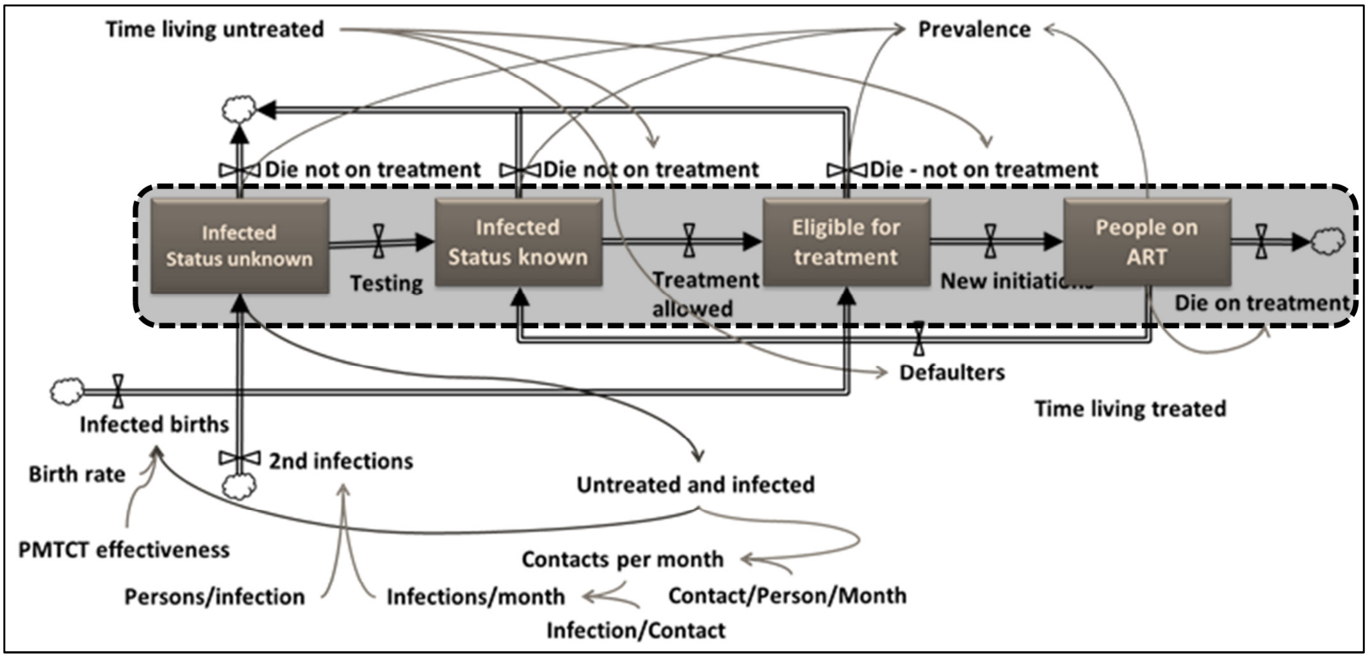


Figure 9: Flow of patients

The rates that define the flow of people between the different states are also modelled as people/month. The rate of testing represents how many infected people get tested for HIV and move to the state of "status known". Since all positive persons are not eligible for ART the flow of people from "status known" to eligible is governed by policy. The current policy allows for around 50% of positive people to be eligible. Once a person fits the eligibility criteria, there is a process of initiation that he/she has to go through to be initiated onto ART. The flow of patients through this process is also modelled. Once a patient is initiated onto ART he/she is represented in the "People on ART" level. The number of people that can be initiated each month is limited by the capacity of the system.

Due to the nature of ART, patients have to regularly return to health care facilities for check-ups and to collect their medication. The medication collection takes place every month, and the check-ups between monthly and 6-monthly according to national policy. This regular return of patients to facilities has an effect on the available capacity to initiate new patients. This capacity limiting characteristic is modelled with the new initiations rate.

The equations and validation thereof defining the above section of the mode are defined as follows with the level equations described first, the rate equations second and then the auxiliary variables:

Level equations:

(17) **Infected-Status Unknown** = INTEG ("2nd Infections"-Die not on Treatment-Testing, 100000)
Validation:
Units: person
Initial level based on estimate prevalence

(15) **Infected - Status Known** = INTEG (Defaulters + Testing - Die not on Rx with known status-Treatment allowed, 30000)
Validation:
Units: person
Defaulter flow into the category of infected "status known" and follow the generic flow from there.

(21) **Number (Nr) Eligible for Treatment** = INTEG (Infected births + Treatment allowed-Die not on Rx but Eligible-New Initiations + Infected births, 1100)
Validation:
Units: person

(22) **People on ART** = INTEG (New Initiations-Defaulters-Die on Treatment-Defaulters, 40000)
Validation:
Units: person

Rate equations:

(28) **Testing** = Testing Promotion*"Infected-Status Unknown"/Time before being tested
Validation:
Units: person/Month

(36) **Treatment allowed** = Rx criteria*"Infected - Status Known"/Time to test for RX
Validation:
Units: person/Month
Based on the fact that not all infected people are eligible for treatment, only a fraction of the infected patients can be treated, and there is also a factor of time associated with being initiated onto ART represented by the variable "Time to put on Rx"

(20) **New Initiations** = min (Treatment Capacity-(**People on ART/Time for revision of Rx**),0.4*Nr Eligible for Treatment/"1month")
Validation:
Units: person/Month

Number of new initiations is limited by the capacity of initiation, which is further limited by revisions of treatment.

(37) **Treatment Capacity** = 161361
Validation:
Units: person/Month
Total headcount as in the District health information system (DHIS).5 for HIV usage for 80 per cent utilization of staff*

Auxiliary variables:

There are various auxiliary variables used in the model. These variables are used in other equations to complete influences from the environment or represent constants. The relevant auxiliary variable for the flow of patients are described below:

(26) **Rx criteria** = 0.5
Validation:
Units: Dmnl
Based on literature: Estimating cost of eligibility change – WHO
Stover et al. suggested an estimate of between 30% and 60% of HIV positive patient having a CD4 count of lower than 350 [48] [49].
The widely accepted percentage of 50% in South African health care applications, and 50% falling between the suggested percentages of Stover et al. thus 50% is used.

(29) **Testing Promotion** = 1
Validation:
Units: Dmnl
Total tested / total estimate being Pos.
To be used as a variable.
The number of people being tested for HIV is based on the "Testing Promotion" variable. The "Testing Promotion" variable is used to account for the fact that not all infected people are being tested.

(30) **Time before being tested** = 15
Validation:
Units: Month
Quarter of infected untreated life, based on silent period
There is a large variance of the "silent" period in the disease [50]. *Thus the assumption is made that the silent period of the disease is a quarter of infected untreated life span.*

(34) **Time to test for Rx** = 2
Validation:
Units: Month
Based on the duration between being identified as eligible and CD4 count tests are conducted to initiate treatment. Two months is the current period observed in practice in public health.

(31) **Time for revision of Rx** = 1
Validation:
Units: Month

Infections

Infected births and sexual transmission (2nd infection) are the two sources of infection that is modelled. An infected mother infects her baby when the necessary preventative measure are unsuccessful. These preventative measures are called: Prevention of Mother to Child Transmission (PMTCT). A sexual transmission occurs when the HIV virus is transferred from an infected person to an uninfected person during a sexual contact.

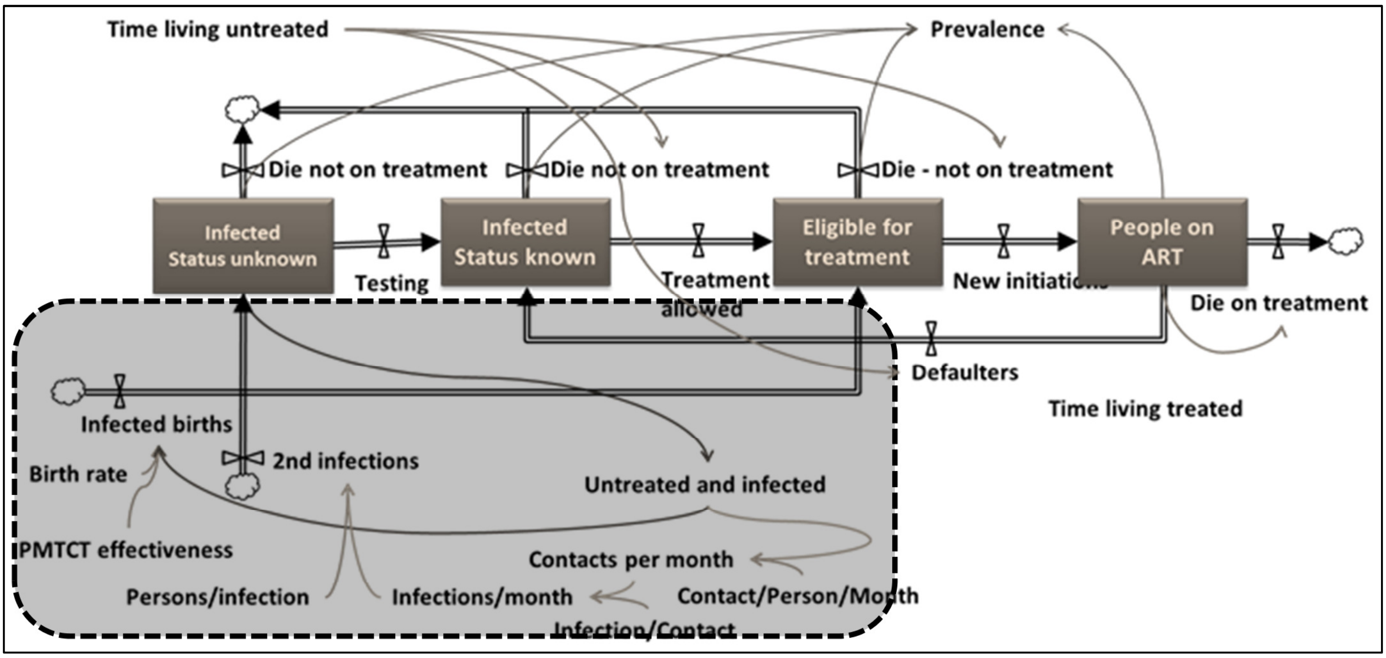


Figure 10: Infections

The number of infected births depends on the PMTCT effectiveness in the area that is modelled, as well as the size of the infected population. When a baby with an infected mother is born they are tested and if the baby is infected that baby is by default eligible for ART. Because of this protocol all infected births go directly into the state of "eligible for treatment"

Secndary infections is mainly a function of the infected population, the uninfected population and the sexual interaction between the tow. The 2nd infections are modelled by using the total number of infected people that is not on treatment and determining the number of potentially infective contacts there are each month. This number of contacts is then used to establish the number of new infections each month. In the model it is assumed a person that is on ART does not infect any further people. This assumption is based on the fact that people on ART has a higher CD4 count and thus they have a much lower infection rate of their sexual partner.

Rate equations:

(16) **Infected births** =Birth rate * Untreated and infected*(1-Pmtct effectiveness)
Validation:
Units: person/Month

Births are the only alternative way to get infected by HIV. All babies born to HIV positive mothers are tested and eligible for treatment.

(24) **PMTCT effectiveness** = 0.75
Validation:
Units: Dmnl
Based on internal BRHC assessment, and data from DHIS.

(03) **Birth rate** = 0.043
Validation:
Units: person/person/Month
Based on estimation of District Health Barometer [7]

(38) **Untreated and infected** = "Infected-Status Unknown" + "Infected - Status Known" + Nr Eligible for Treatment
Validation:
Units: person

(02) **2nd Infections** = "Infections (Inf) /month" * "Persons (Pers)/Inf"
Validation:
Units: person/Month

(14) **Inf/month** = "Contacts/month"*"Infection/contact"
Validation:
Units: infection/Month

(05) **Contacts/month** = SMOOTH(Untreated and infected*"Contact/Pers/Month",2)
Validation:
Units: contact/Month

(18) **Infection/contact** = 0.009
Validation:
Units: infection/contact
According to the study by Leynaert et al. the probability of transmission of HIV during a heterosexual contact is 0.9%. This percentage is excluding the CD4 count, the stage of the infected person or the type of sexual contact. [51]

(04) **Contact/Pers/Month** = 0.5
Validation:
Units: contact/person/Month
Estimation of the number of sexual contact one infected person has with an uninfected person

Auxiliary variables:

(23) **Pers/Inf** = 1
Validation:
Only one person is affected per infection, in some other diseases, such as airborne diseases, multiple people can be infected form one infection.

Deaths

Deaths are modelled for people in each state. These deaths are split into two groups; the first is deaths of people that are not on treatment, and the second for those that is on treatment. The reason for this differentiation is because of the increase in life expectancy when an infected person is on ART.

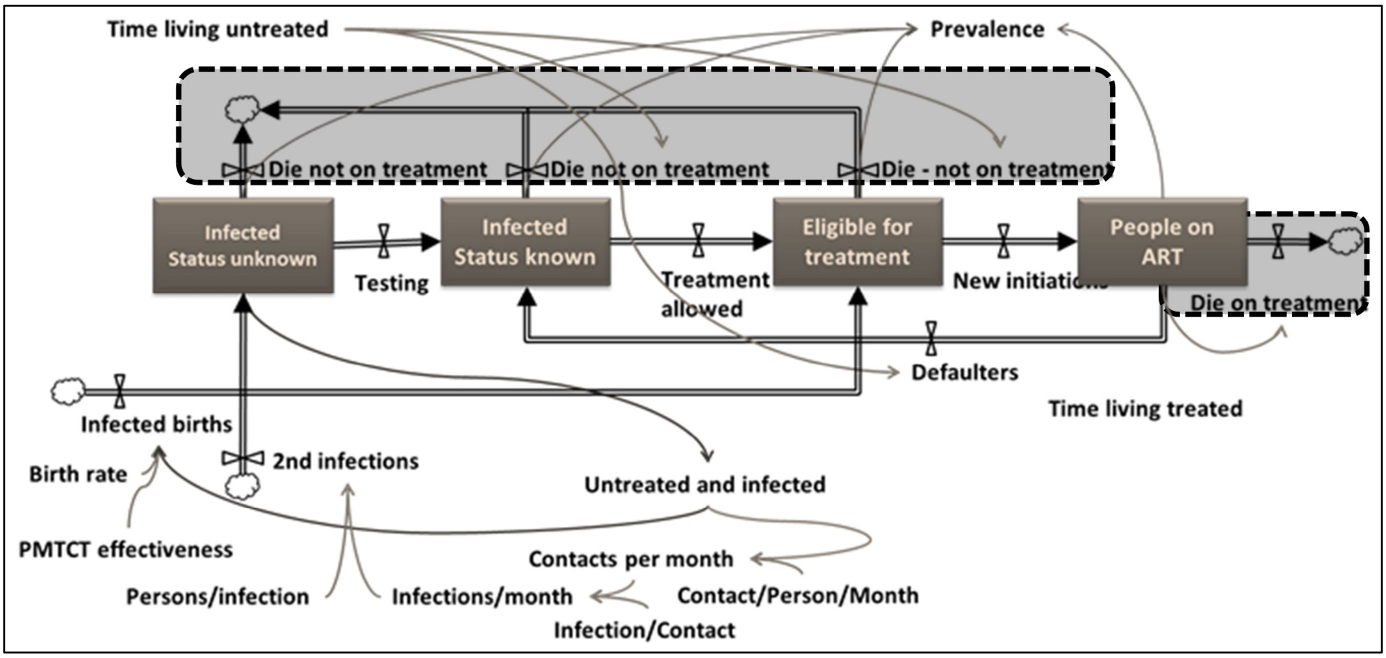


Figure 11: Deaths

The rate at which people from a certain state die is a function of the life expectancy and the number of people in that state.

Rate equations:

(10) **Die not on Treatment** = "Infected-Status Unknown"/Time living untreated
Validation:
Units: person/Month

(09) **Die not on Rx with known status** = "Infected - Status Known"/Time living untreated
Validation:
Units: person/Month

(08) **Die not on Rx but Eligible** = Nr Eligible for Treatment/Time living untreated
Validation:
Units: person/Month

(11) **Die on Treatment** = SMOOTH(People on ART/Duration living on Rx, 3)
Validation:
Units: person/Month
Death rate smoothed over a period of 3 month to compensate for fluctuations in human physiology.

Auxiliary variables:

(12) **Duration living on Rx** = 120

Validation:

Units: Month

Lifespan of 10 years used based on discussion with clinical HIV specialist.

(32) **Time living untreated** = 60

Validation:

Units: Month

5 Years used as untreated standard life duration if not treated.

Defaulters

During the course of treatment some people default on their treatment. When a patient defaults he/she does no longer adhere to the prescribed regime of taking their medication. Defaulting of ART can be due to various reasons that include; lack of patient education, lack of access to medication or side effects and stigma of medication.

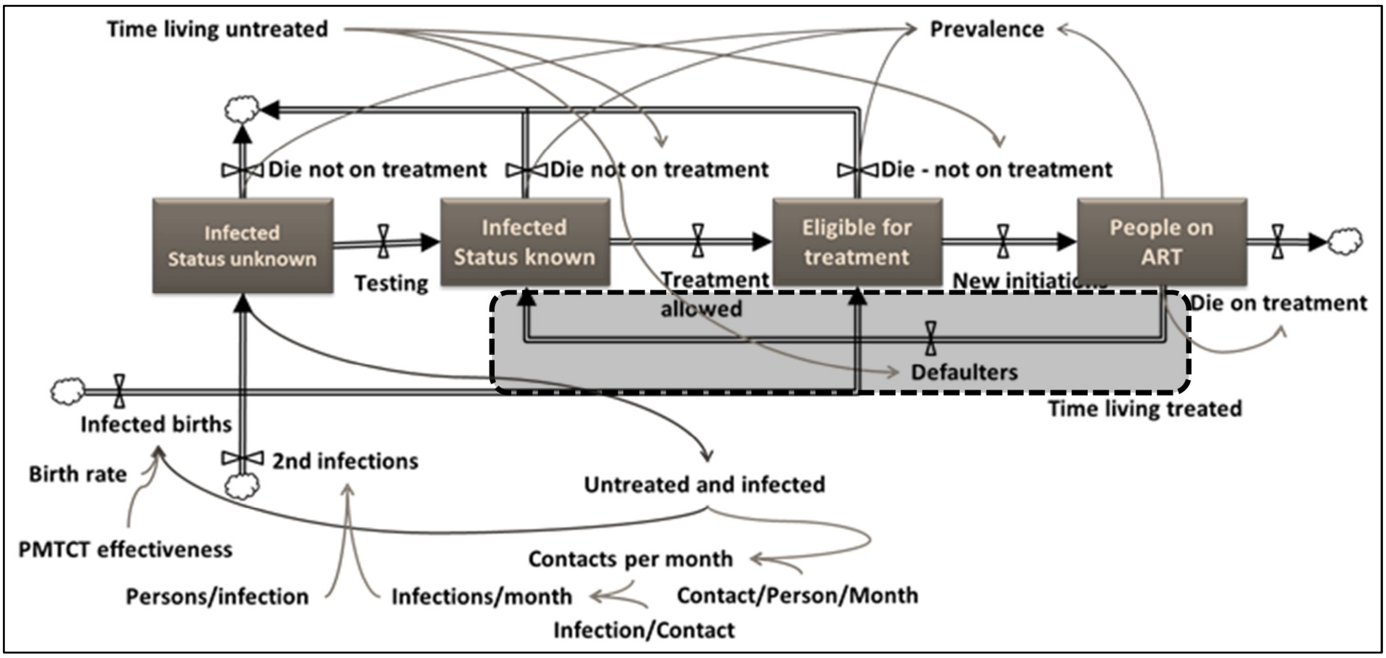


Figure 12: Defaults

In the model a defaulted patient returns to the state of "infected status know" and will have to move through the rest of the process before they can be re initiated onto ART.

Rate equations:

(07) **Defaulters** = Default Rate*People on ART/Time living untreated

Validation:

Units: person/Month

Auxiliary variables:

(06) **Default Rate** = 0.05

Validation:

Units: Dmnl

Estimation based on discussion with field workers

Indicators

To assess the model some variables was built into the model to act as indicators. These variables were used to measure the prevalence of HIV and the total number of deaths for the duration of the modelling run.

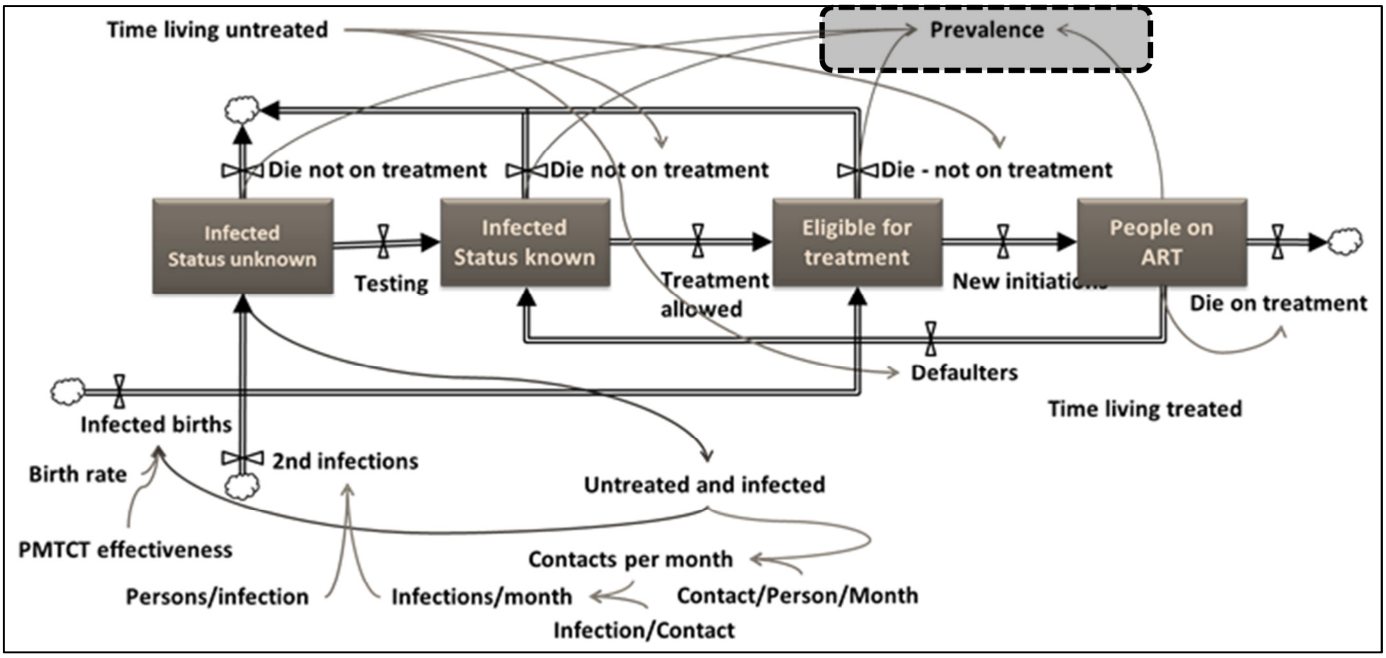


Figure 13: Indicators

Level equations:

(25) **Prevalence** = ("Infected - Status Known" + "Infected-Status Unknown" + Nr Eligible for Treatment + People on ART)/950000

Validation:

Units: Dmnl

Estimated prevalence rate for uThungulu using DHIS average data between 2011/01 and 2012/04: "Number of persons tested HIV positive (exl ANC)"/ "Number of persons tested HIV (excl ANC) – Total" ; = 0.171

(35) **Total Die** = INTEG (Die not on Rx but Eligible +Die not on Rx with known status + Die not on Treatment + Die on Treatment, 0)

Validation:

Units: person

Modelling parameters

The model was executed for a period of 20 years (240 months) with 1 month increments. No formal warming up period was used because the starting states of the levels wee calibrated to represent the current measures as found in the DHIS.

Model Validation: Structure-Oriented Behaviour Testing

Based on: Data, Literature, Operational knowledge or expert review

Validation of: Model or Sub-model structure

Consider the following tests

- **Sensitivity testing**
- **Extreme condition test**
- **Modified behaviour test**
- **Phase relationship testing**
- **Turing test**

Sensitivity testing:

Five variables have been identified as potential areas where interventions can be made. The sensitivity testing of the model was based on the five identified variables. These variables are:

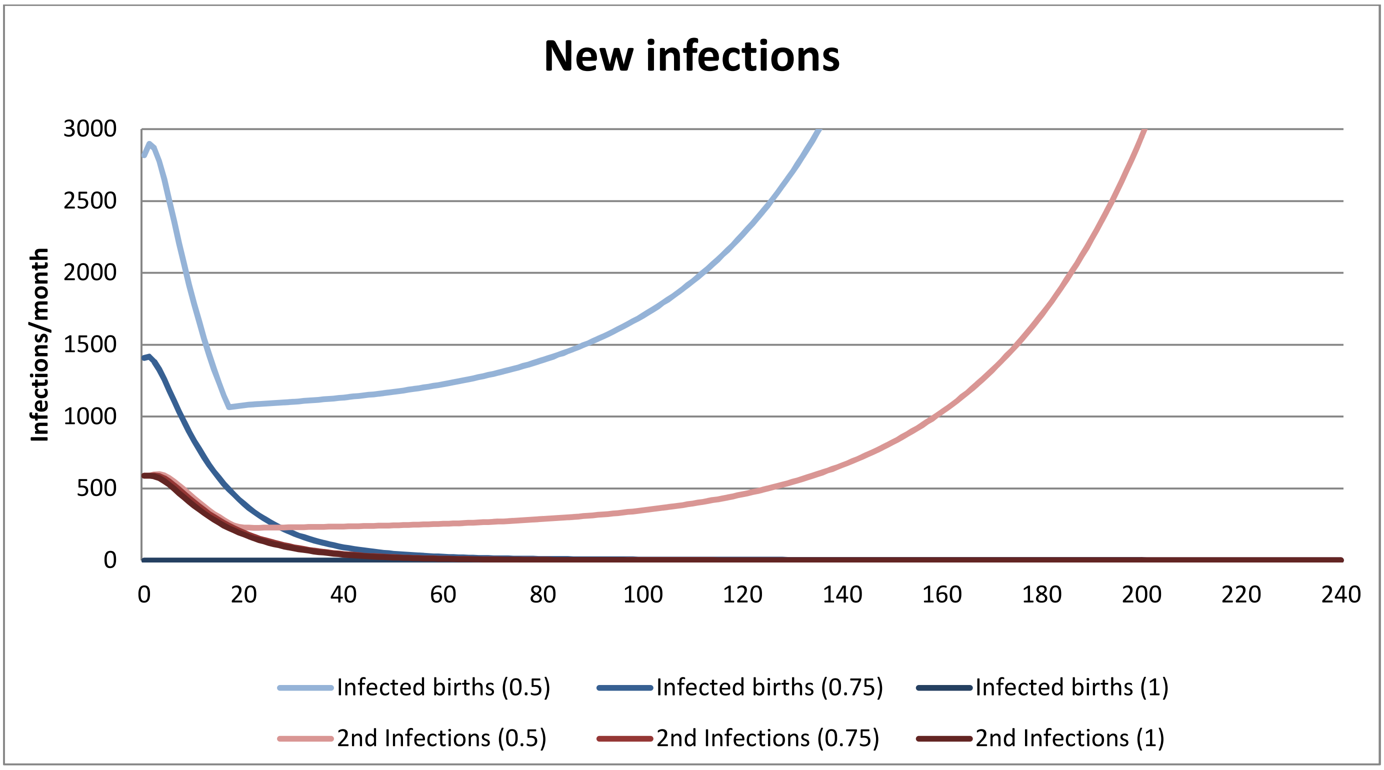
1) PMTCT
2) Infectiveness of disease
3) Treatment criteria
4) Capacity
5) Time for revision of treatment

The sensitivity of the model was tested for three values of each variable and the response of the model is show in relevant graphs.

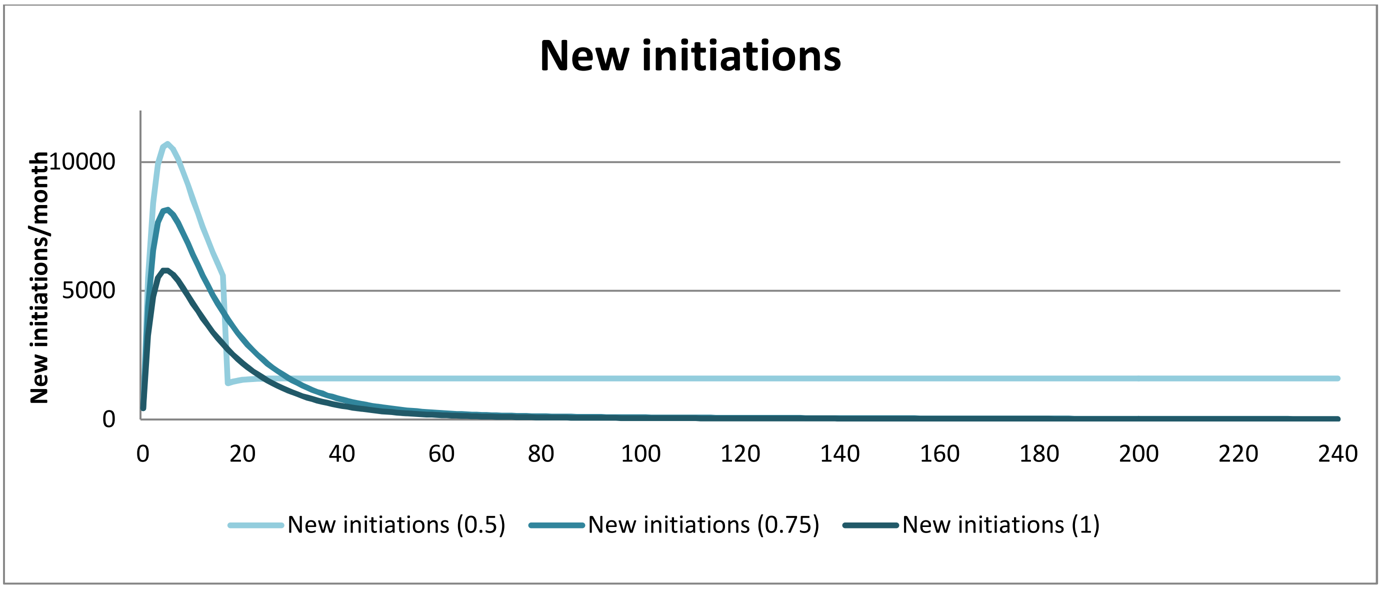
Variable: PMTCT effectiveness.
Values: 0.5, 0.75, 1

The sensitivity of the model to the PMTCT effectiveness is tested. PMTCT effectiveness is chosen as a variable, because it can be manipulated through various program implementations and interventions.

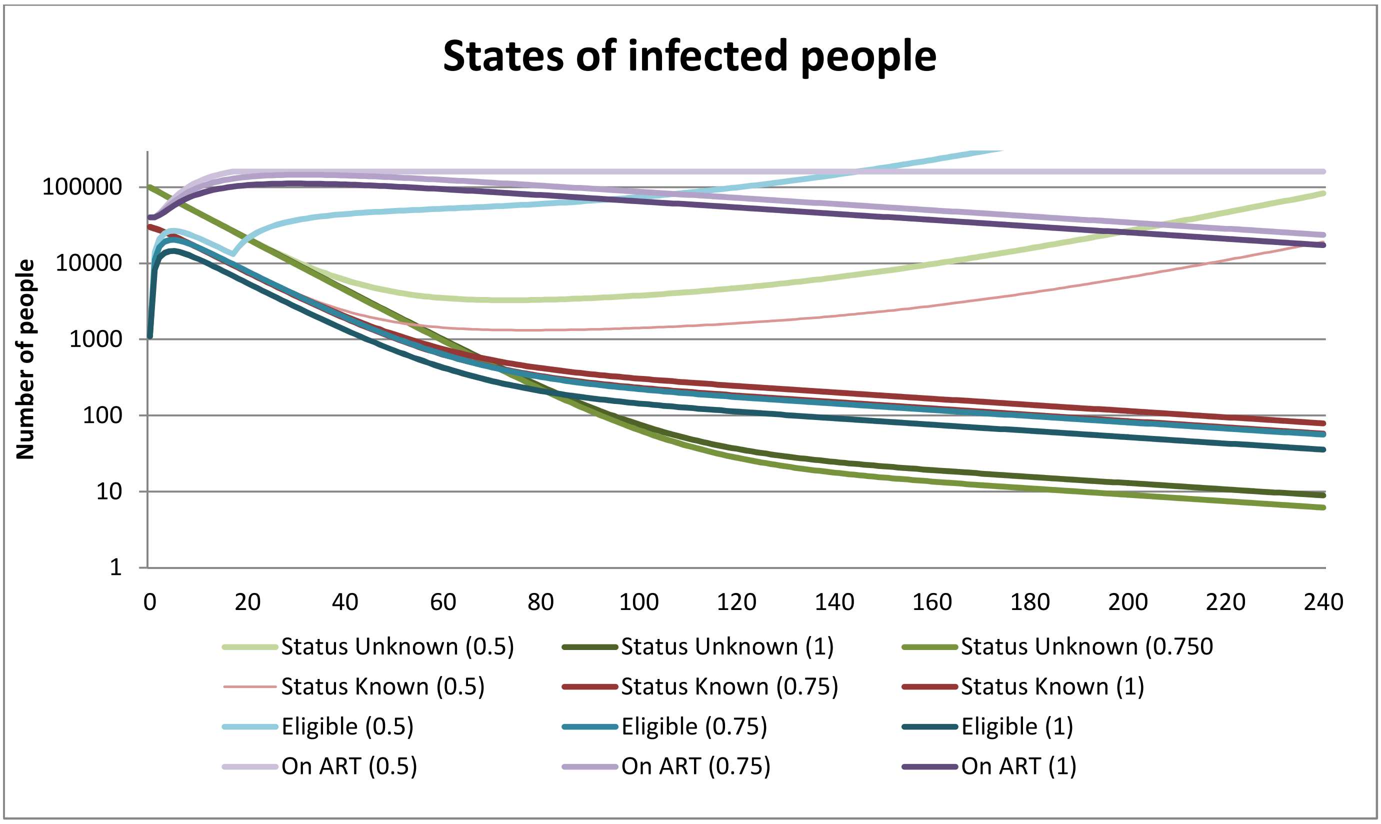
The effect on the model can be seen in the following graphs for PMTCT effectiveness values of 0.5, 0.75 and 1.



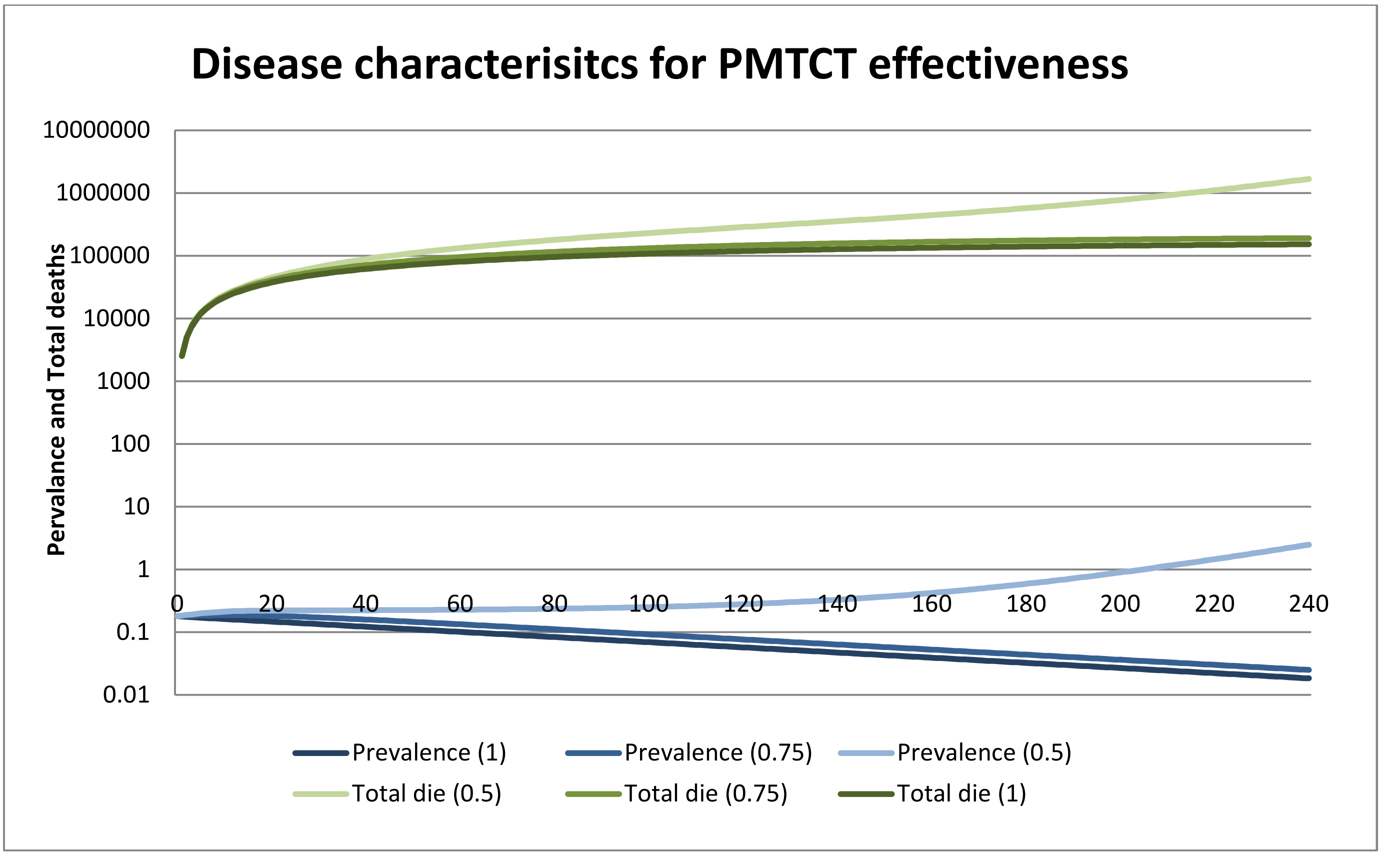
Graph 1: New infections based on PMTCT



Graph 2: New initiations based on PMTCT



Graph 3: States of infected people based on PMTCT



Graph 4: Disease characteristics for PMTCT

Sensitivity observations

New infections: Where PMTCT effectiveness is low the infections (both new births and 2nd infections) climbs drastically. It is thus suggested that there is a critical PMTCT effectiveness needed for stabile behaviour. At a PMTCT effectiveness as high as 100% there is Only a slight drop in 2nd infections.

New initiations: is sensitive to changes in PMTCT infectiveness. Where the PMTCT effectiveness is very low the number on new initiations jump higher initially and then stabilises much higher. This stabilisation of higher number is due to model behaviour, where new initiations are limited by capacity.

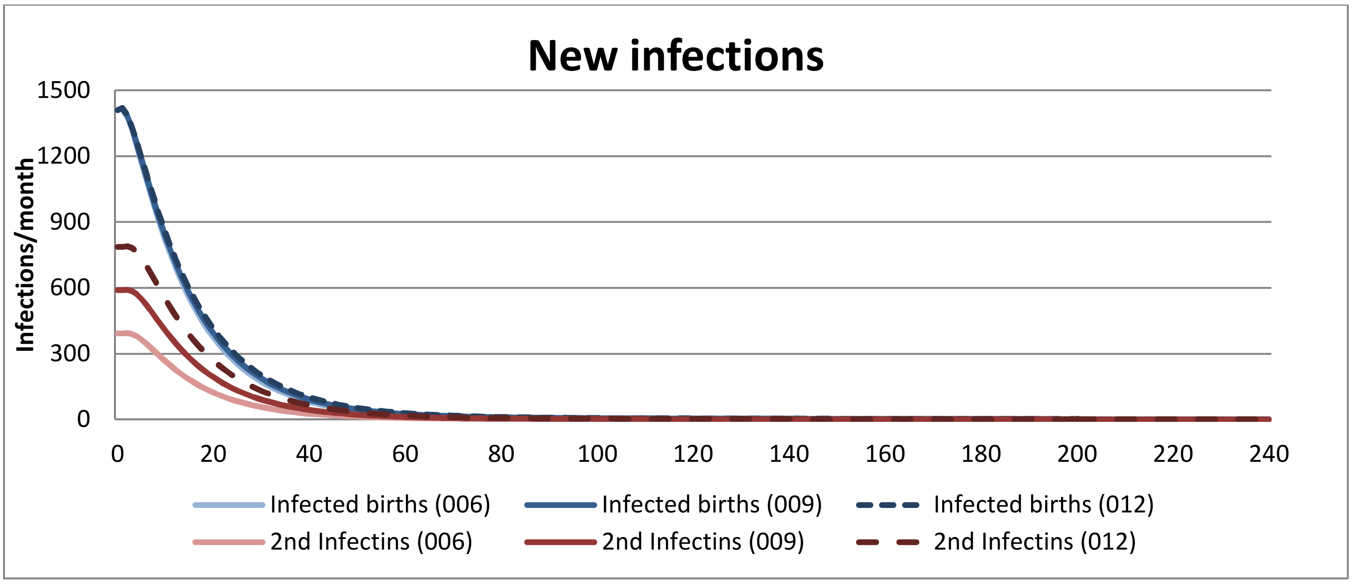
States of people: With low PMTCT effectiveness there is a general rapid increase in numbers of people in all states. With high effectiveness there is no major change is behaviour.

Disease characteristics: The effect of low PMTCT effectiveness has a big influence on both the number of people who die, as well as the prevalence of the disease. The number of people who die increases drastically. When the PMTCT effectiveness is improved, there is only a slight reduction in total number of deaths.

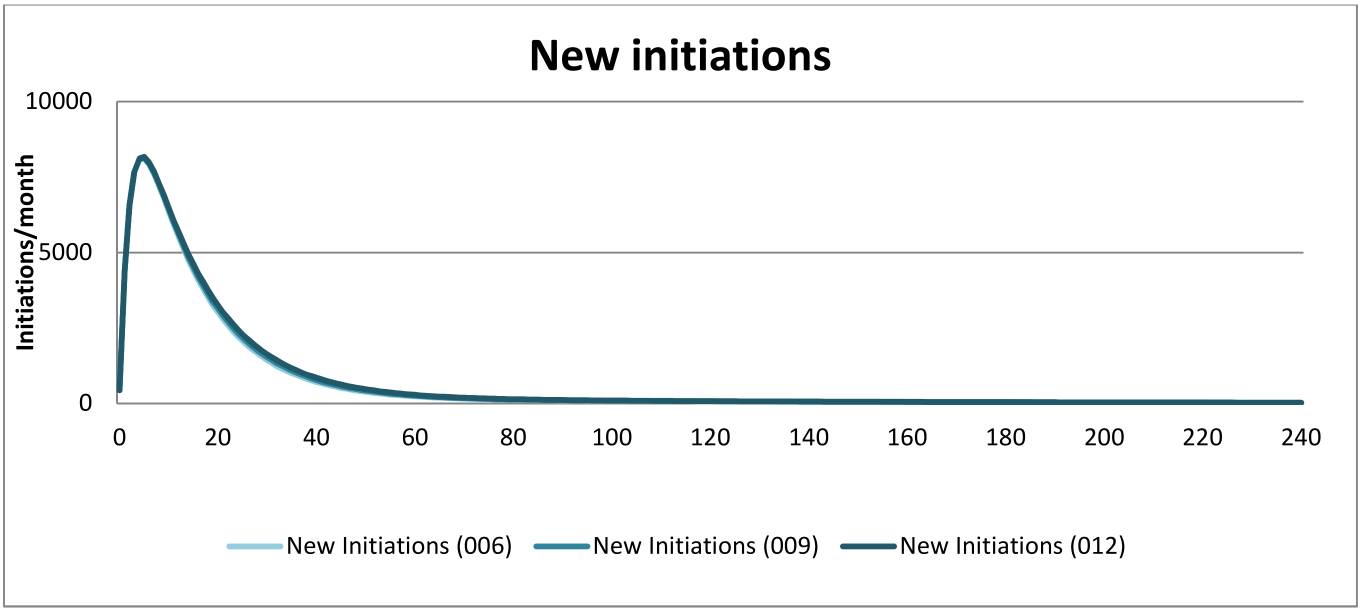
Variable: HIV infectiveness for a sexual contact.
Values: 0.6%, 0.9% and 0.12%

The sensitivity of the model to the HIV infectiveness is tested. HIV infectiveness is chosen as a variable because it can be improved through safe sexual practices, male circumcision and community education.

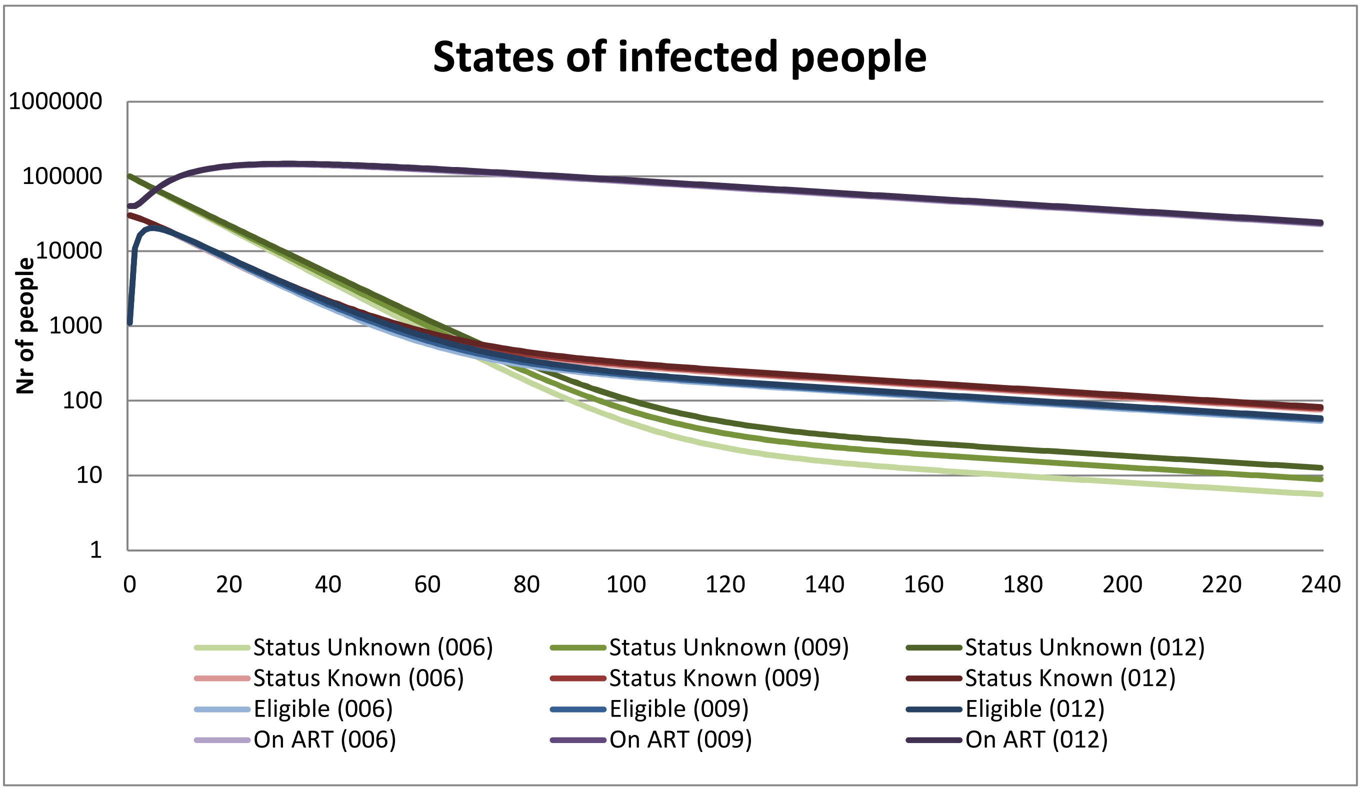
HIV infectiveness is modelled for the values of 0.6%, 0.9% and 0.12%.



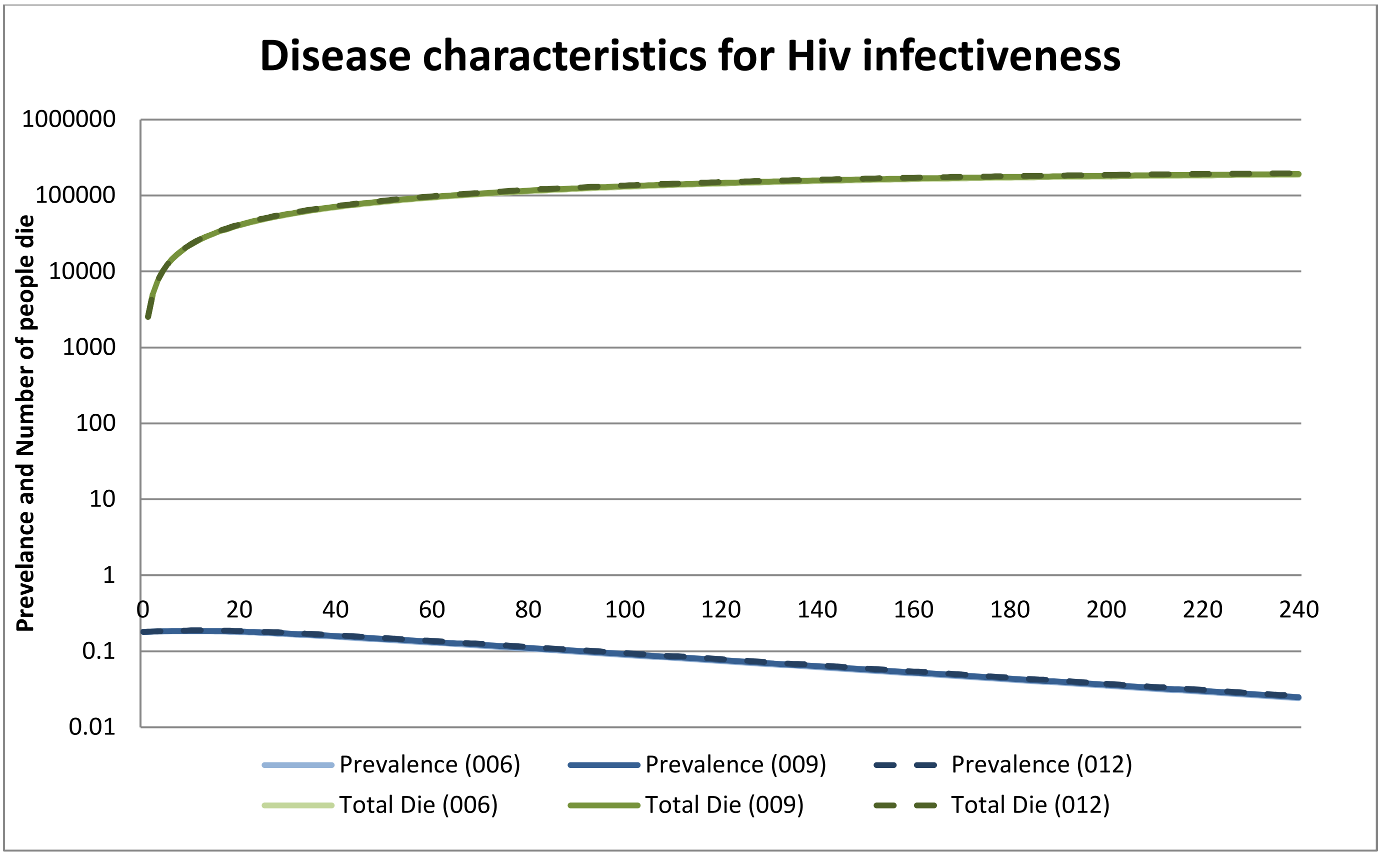
Graph 5: New infections based on infectiveness



Graph 6: New initiations based on infectiveness



Graph 7: States of infected people based on invectiveness



Graph 8: Disease characteristics based on infectiveness

Sensitivity observations

New infections: Infections through birth is almost not affected at all by the change in infectiveness of sexual contacts. The number of 2nd infections has a linear relationship to the change in infectiveness.

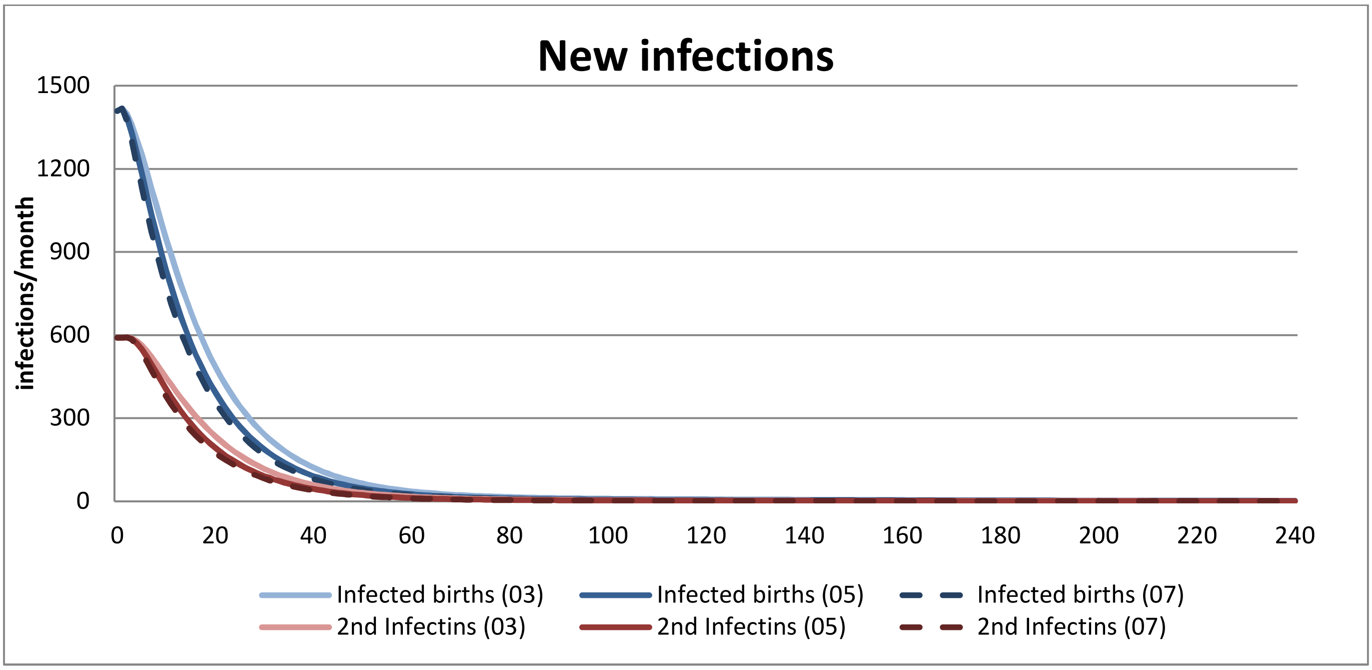
New initiations: Almost no effect can be seen on new initiations per month.

States of people: There is only a small change is the number of people in the different states as a result of manipulated infectiveness

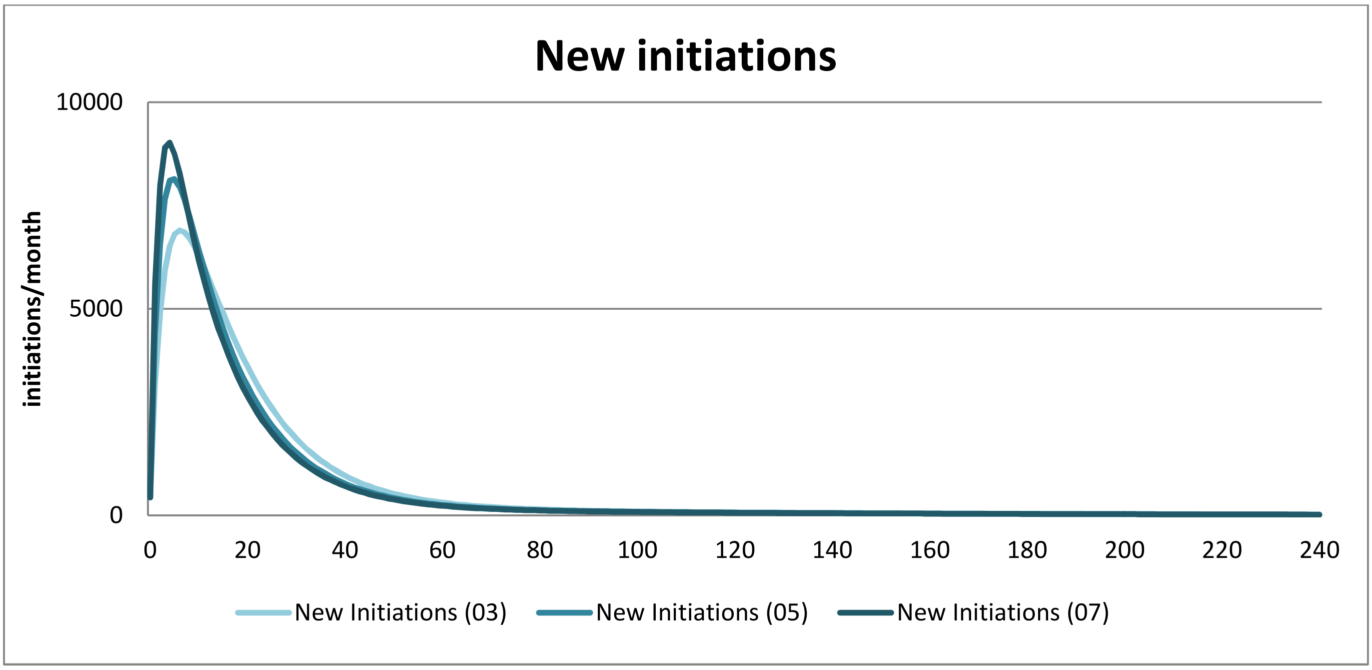
Disease characteristics: The disease characteristics are almost not affected at all.

Variable: Treatment criteria.
Values: 30%, 50%, 70%

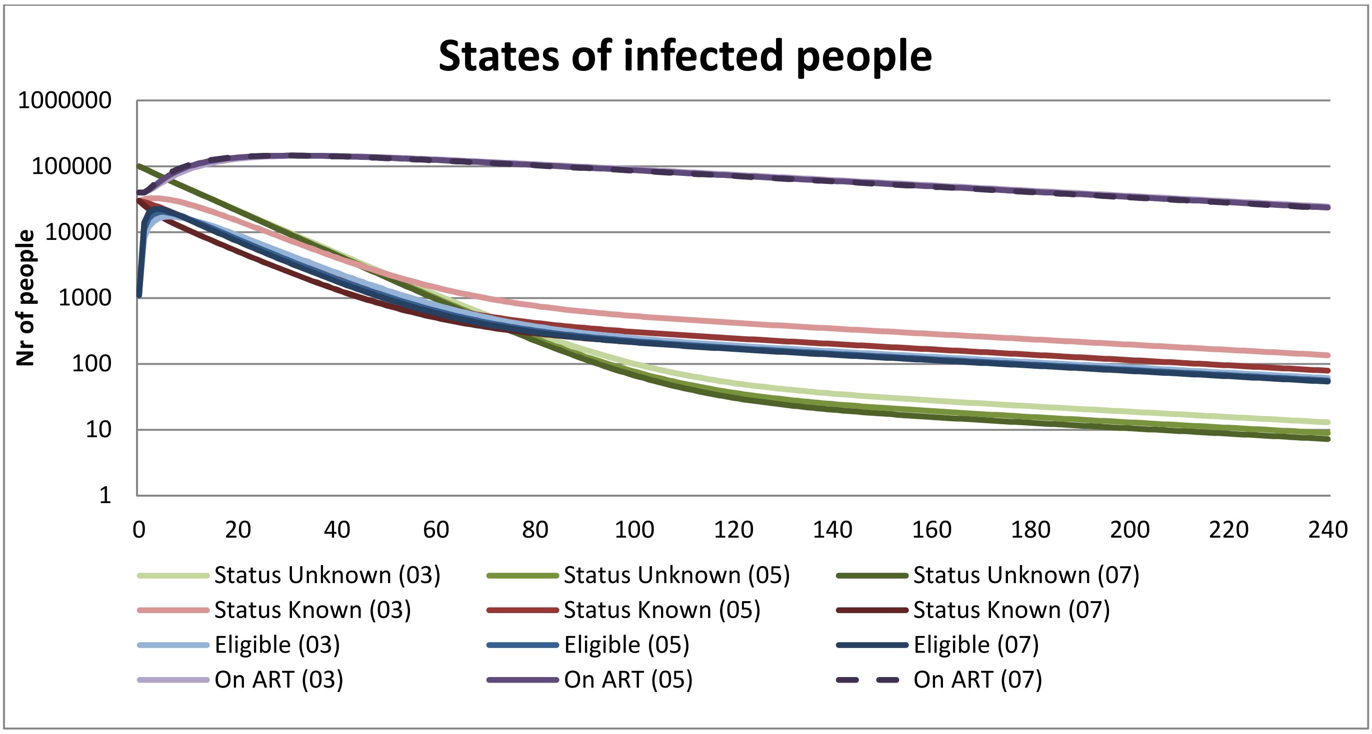
The sensitivity of the model to changes in treatment criteria is tested. Treatment criteria is based on national policy, thus with a change in policy the criteria might change. The effect such a change will have is represented below.



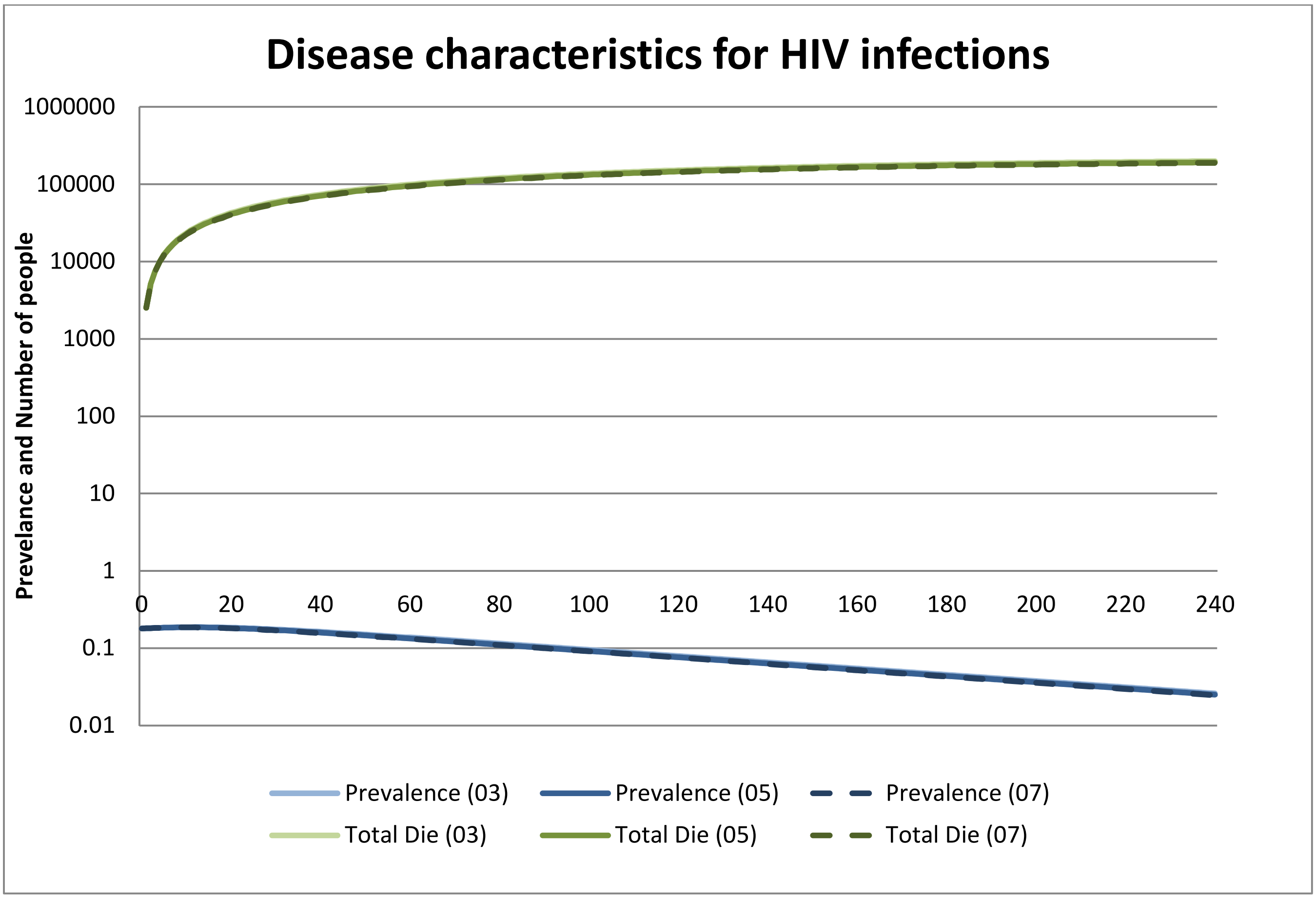
Graph 9: New infections based on treatment criteria



Graph 10: New initiations based on treatment criteria



Graph 11: States of people based on treatment criteria



Graph 12: Disease characteristics based on treatment criteria

Sensitivity observations

New infections: Infections through birth is almost not affected by the change in infectiveness of sexual contact. The number of 2nd infections has a linear relationship to the change in infectiveness.

New initiations: Almost no effect can be seen on new initiations after initial variation.

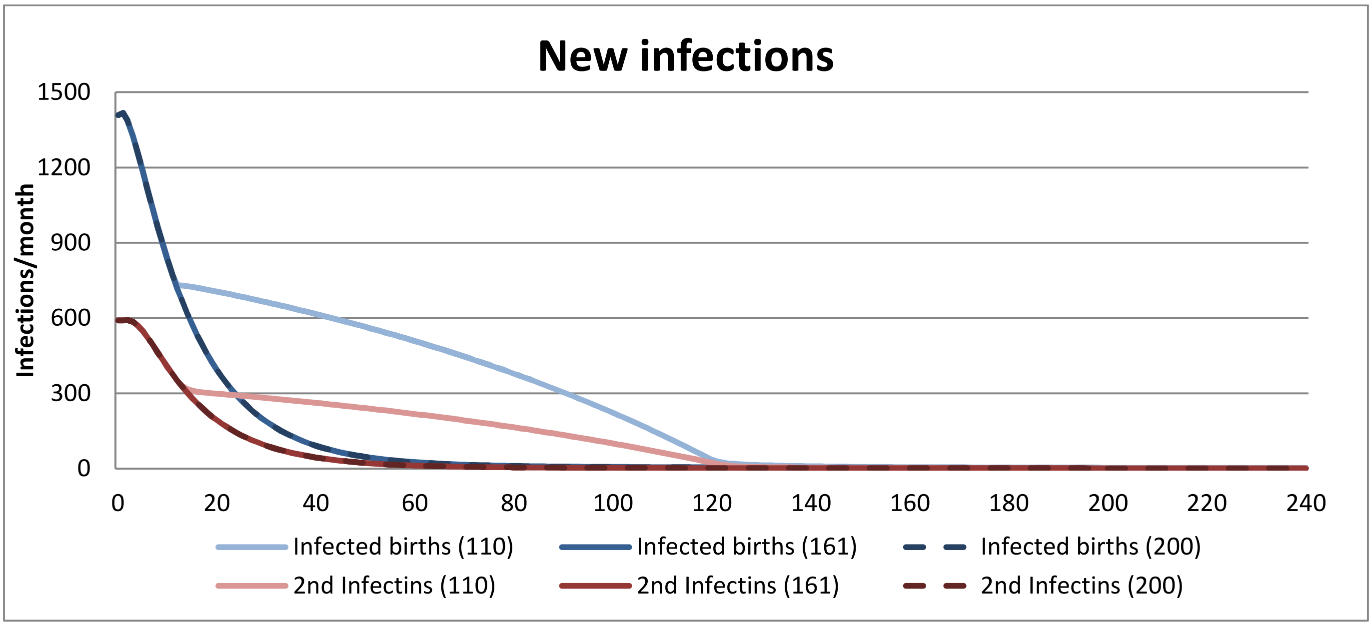
States of people: There is only a small change in the number of people in the different states as a result of manipulated infectiveness. The highest variation is in the number of people with state known.

Disease characteristics: The disease characteristics are almost not affected at all.

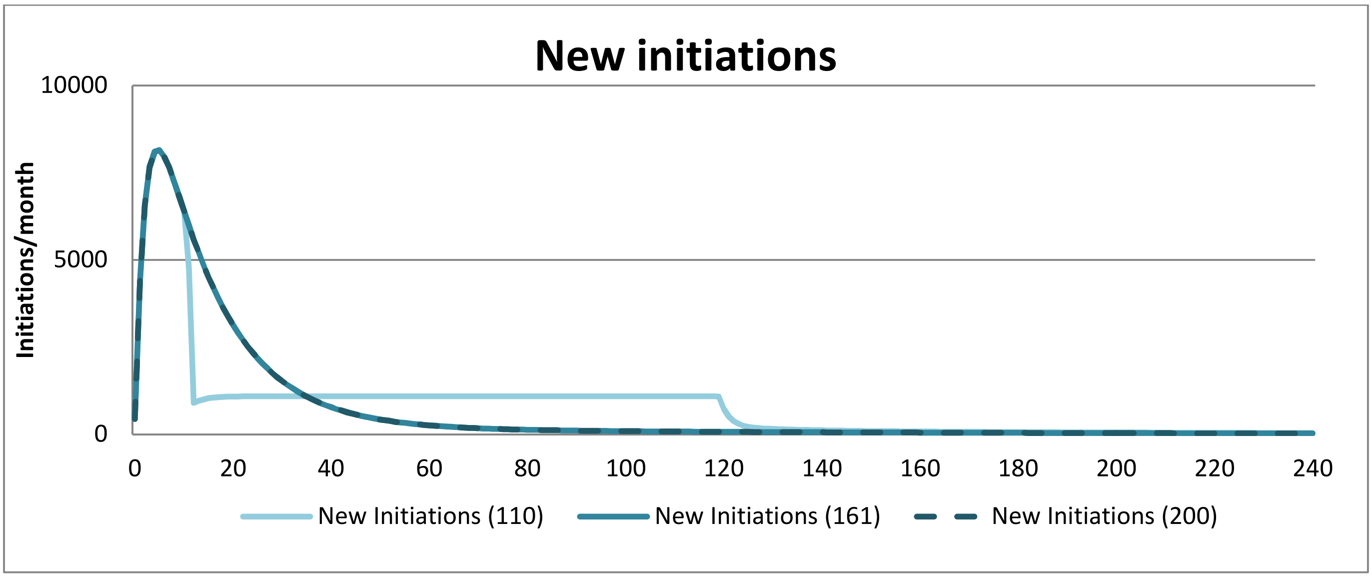
Variable: Capacity of initiation.
Values: 100000 people, 161000 people, 200000 people

The sensitivity of the model to changes in the available capacity to initiate patients is tested. With a fluid working environment the capacity can suddenly reduce, or with continuous improvement intervention capacity can be increased. The reaction of the model to these tests can give an indication of where interventions should be focused.

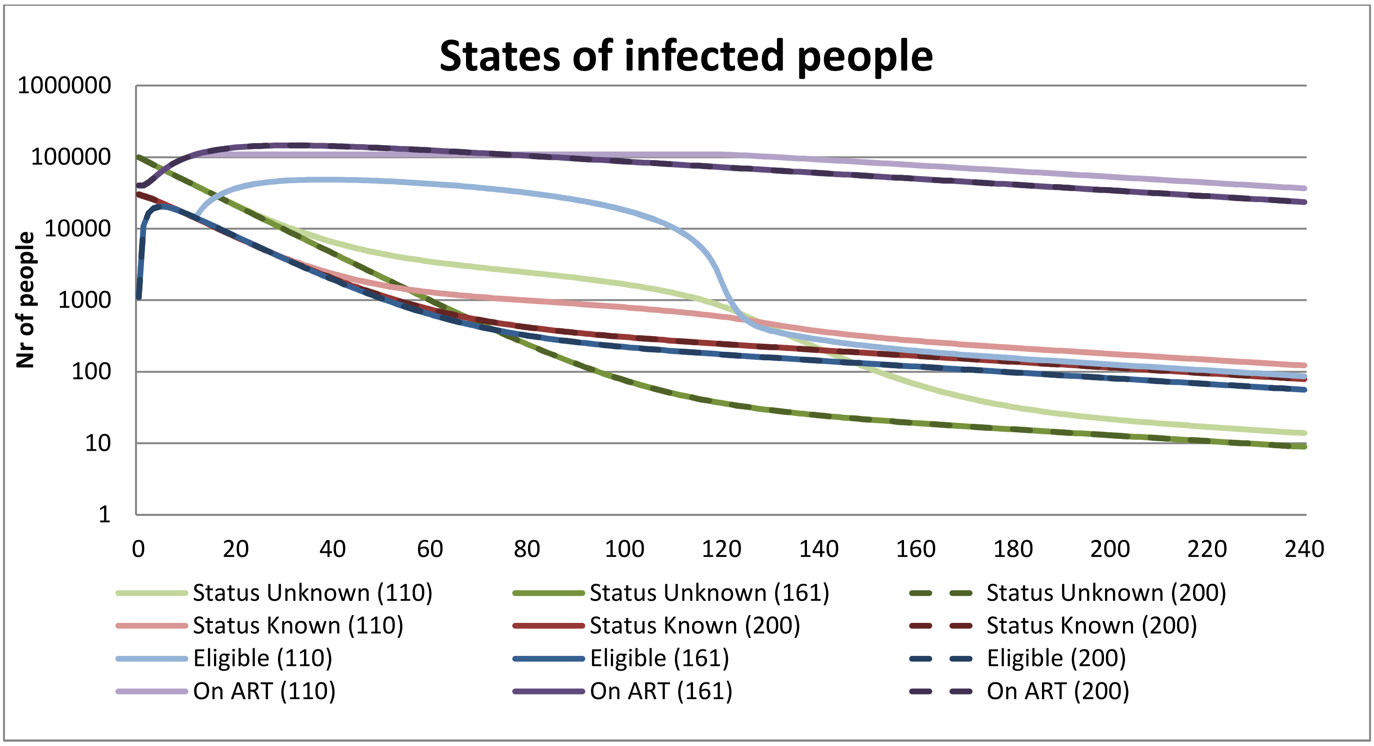
The model is run for three scenarios where the capacity is modelled between 100 000- and 200 000 people.



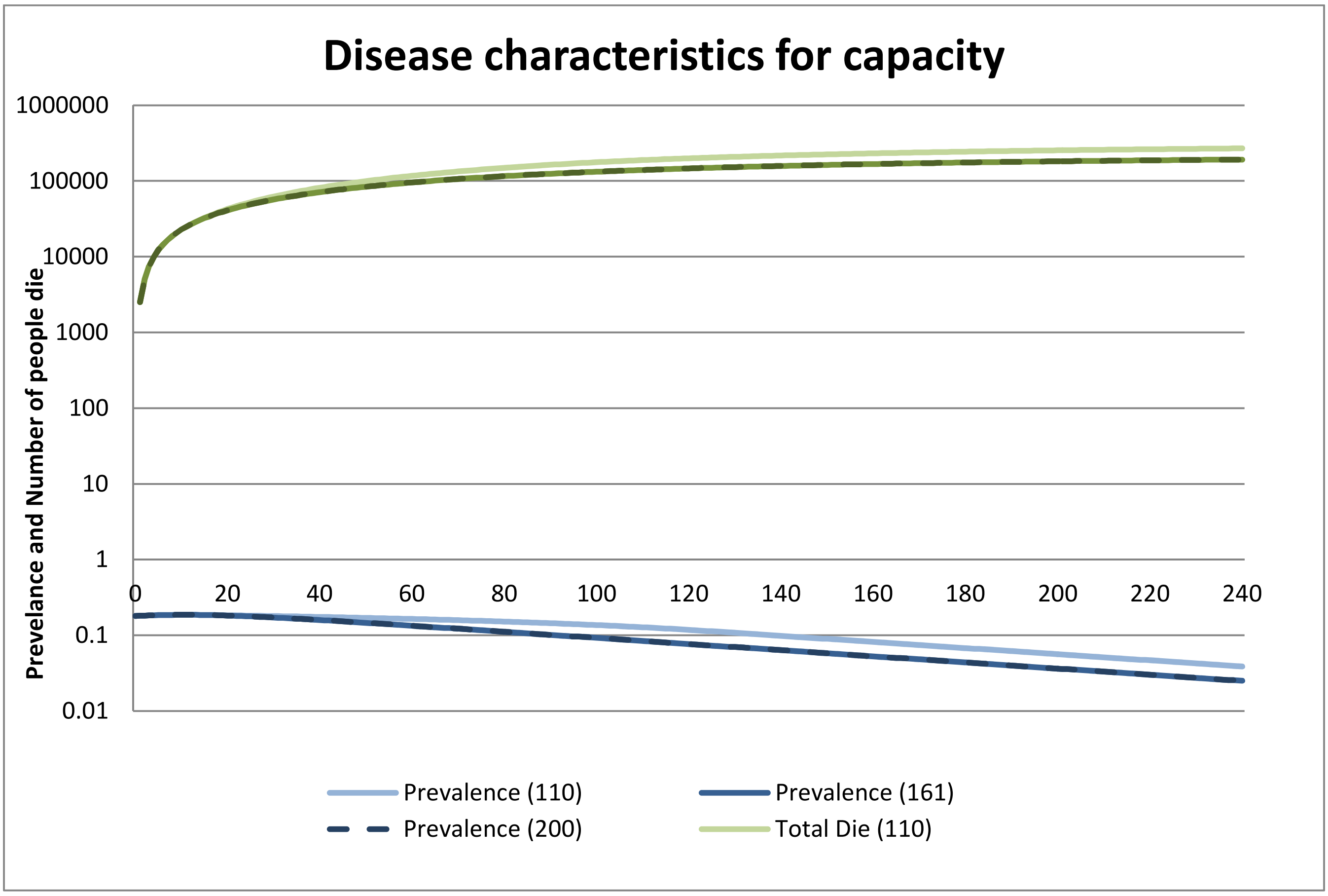
Graph 13: New infections based on capacity



Graph 14: New initiations based on capacity



Graph 15: States of infected people based on capacity



Graph 16: Disease characteristics based on capacity

Sensitivity observations

New infections: When the capacity is decreased initially there is no change in behaviour up to the point where the reduced capacity is reached, where after there is a significant push-back and as a result higher infections for both births and 2nd infections. With increased capacity there is no change in behaviour.

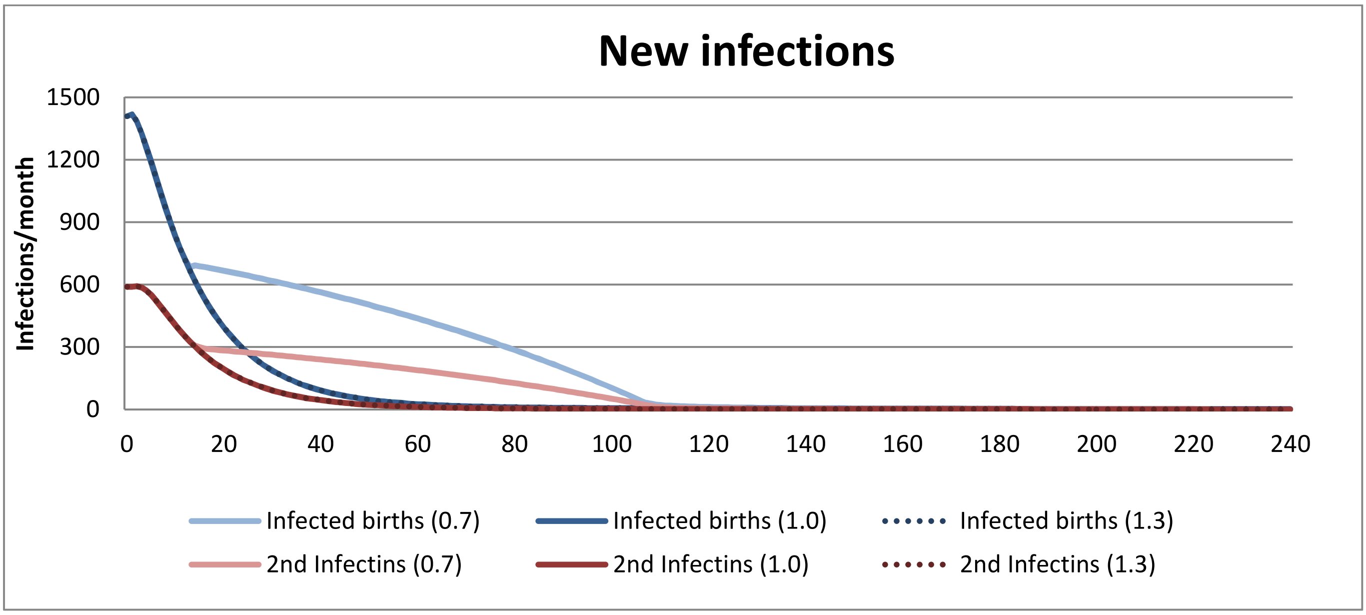
New initiations: When the capacity is decreased initially there is no change in behaviour up to the point where it is reached, then there is a sharp drop up to the point where only new initiations can be done due to deaths or defaults from the group on treatment.

States of people: With decreased capacity there are significant changes in the behaviour. As the lower capacity is reached there is a large increase in people eligible for treatment until the need falls under capacity. There is also an increase of both status known and unknown. There is no change of behaviour if the capacity is increased.

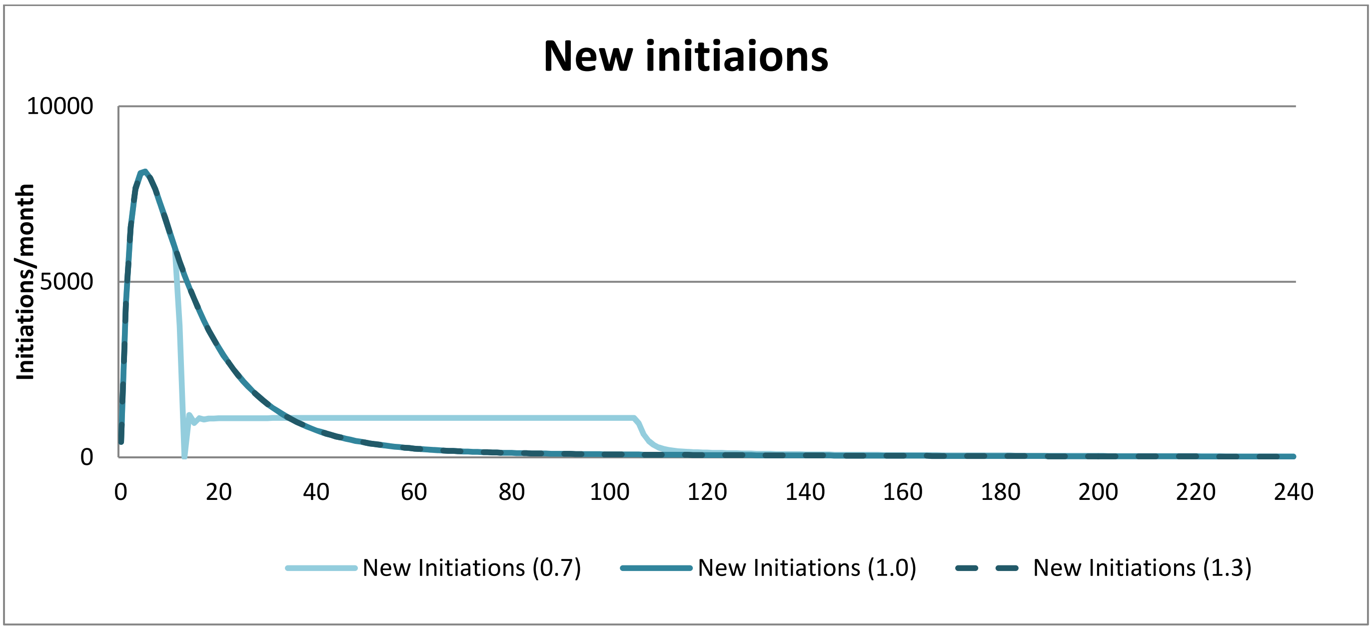
Disease characteristics: Capacity decreasing has an impact on the total deaths and higher prevalence rate, where increased capacity does not have an impact.

Variable: Time between treatment revision.
Values: 0.7 months, 1 month, 1.3 months

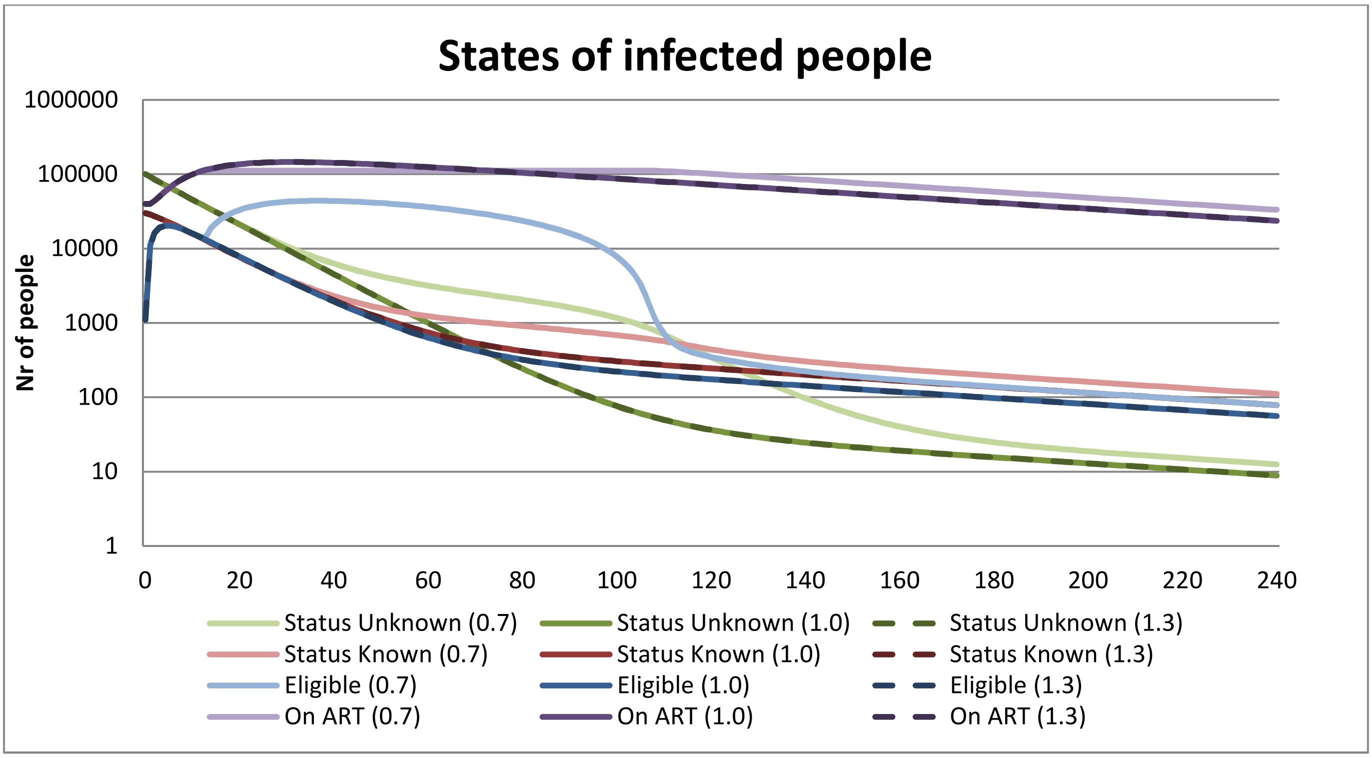
The sensitivity of the model to a variation of time for treatment is represented in the following graphs. The time between treatment revisions is based on national policy and could change. Thus it is important to know what effect such a change will have on the characteristics of the disease.



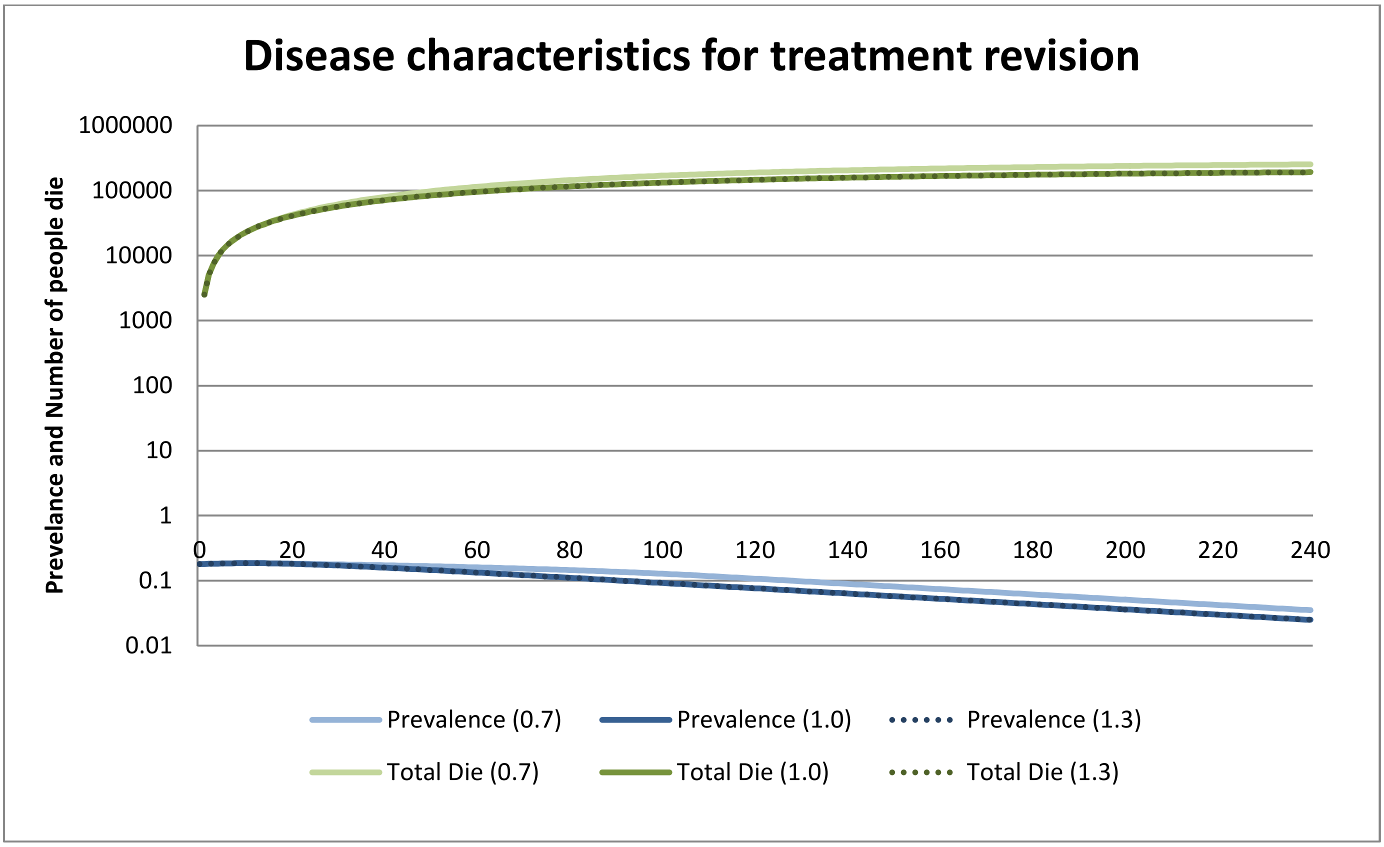
Graph 17: New infections based on treatment revision



Graph 18: New initiations based on treatment revision



Graph 19: States of people based on treatment revision



Graph 20: Disease characteristics based on treatment revision

Sensitivity observations

New infections: When the treatment revision is increased initially there is no change in behaviour up to the point where capacity is reached, then there is a significant push-back and as a result higher infections for both births and 2nd infections. With increased capacity there is no change in behaviour.

New initiations: When the treatment revision is increased initially there is no change in behaviour up to the point where capacity is reached, then there is a sharp drop of new initiations up to the point where only the only new initiations are done due to deaths or defaults from the group on treatment.

States of people: With increased treatment revision there is significant changes in the behaviour. As the capacity is reach there is a large increase in people eligible for treatment until the need falls under capacity. There is also an increase of both status known and unknown. There is no change of behaviour when the treatment revision is decreased.

Disease characteristics: Treatment revision increasing has an impact on the total deaths and higher prevalence rate. In the case where decreased treatment revision does not have an impact. virtual capacity is created.

Model Validation: Behaviour Pattern Testing

Based on: Data, Operational knowledge or expert review and Type of behaviour

Validation of: Model behaviour

Consider the following tests

- **Graphical**
- **Six-step procedure**

Due to the fact that the model represents transient behaviour graphical validation of the behaviour is appropriated.

Pattern testing was done with the use of Vensim's SyntheSim mode.

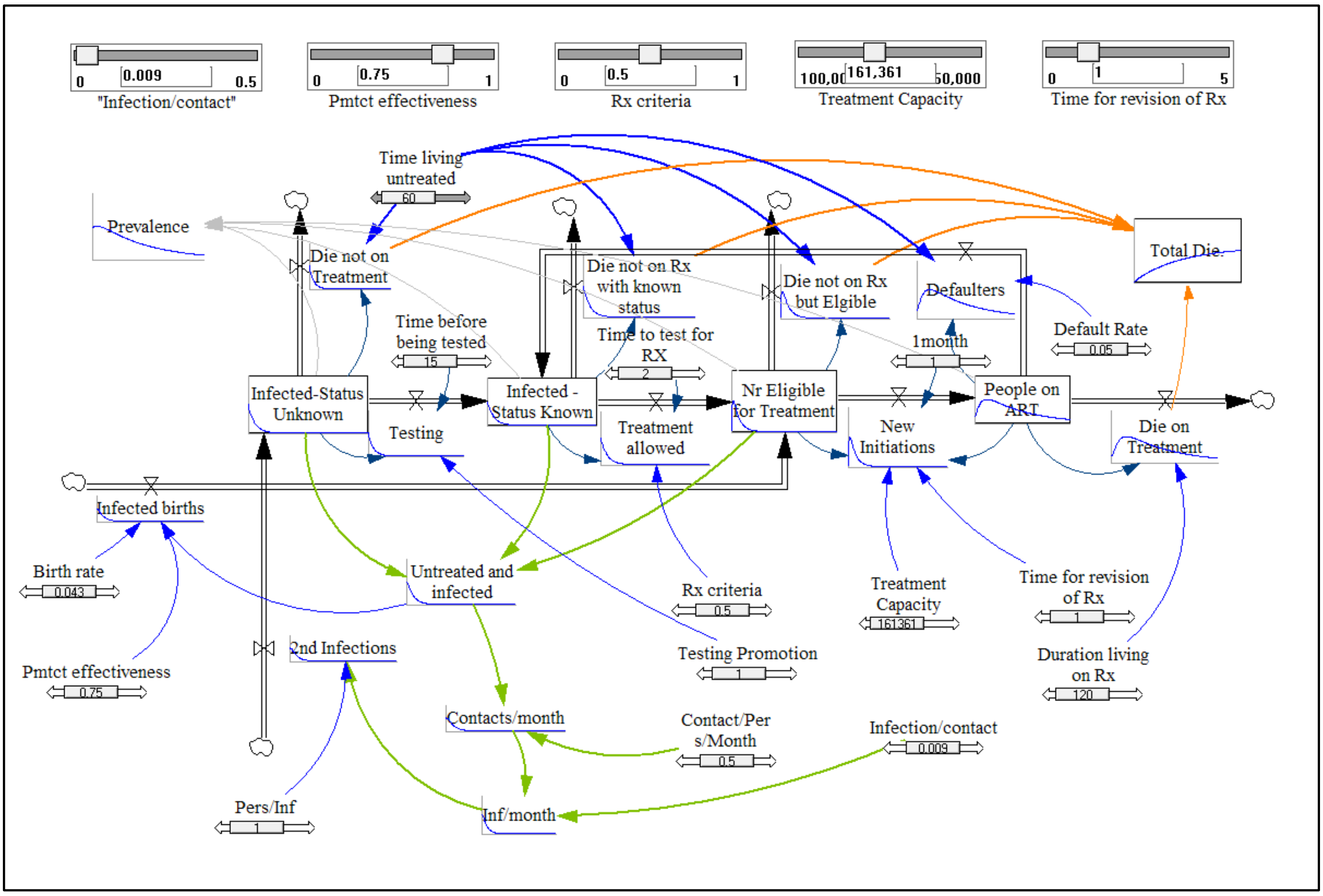


Figure 14: Pattern testing

Based on the graphical outputs of the model as displayed in the sensitivity analysis and the live interaction with the model the model's behaviour can be seen as a useful representation of reality.

Model Implementation

Based on: The purpose of the model

The purpose of the model is to aid the decision of choosing and prioritising interventions in a HIV setting considering the infections of the UThungulu District in KZN

Considering: The findings of the model

Findings from the model:

One of the most insightful finding is to see how the model can be split into two sections: demand for care and supply of care.

Capacity to supply care: With the current parameters the model does not reach its assumed capacity of 161 000 patients per month. The maximum capacity that it reaches is 145 800 patients in month 30.

Need for care: Any intervention that improves (get initiated onto ART faster) the utilisation of the available capacity will improve the outcomes; reduce the total number of deaths. The need for care can also increase when negative factors transpire. A reduction in the effectiveness of PMTCT or increased infectiveness of HIV through sexual contact will also increase the use of the available capacity, but with a negative effect on total deaths.

It is important to understand the negative impact illustrated by the model if patients are not initiated onto ART if they are eligible. If the numbers of eligible patients are increased up to a point where the capacity is reached, the improvement in outcomes will stop. Only when the capacity is increased will the outcomes (total number of people die) further improve.

The model's sensitivity to the five identified parameters can be illustrated in the following graph. The effect of a change (both improving and deteriorating) in parameter is compared to the base simulation for the variable Total Deaths.

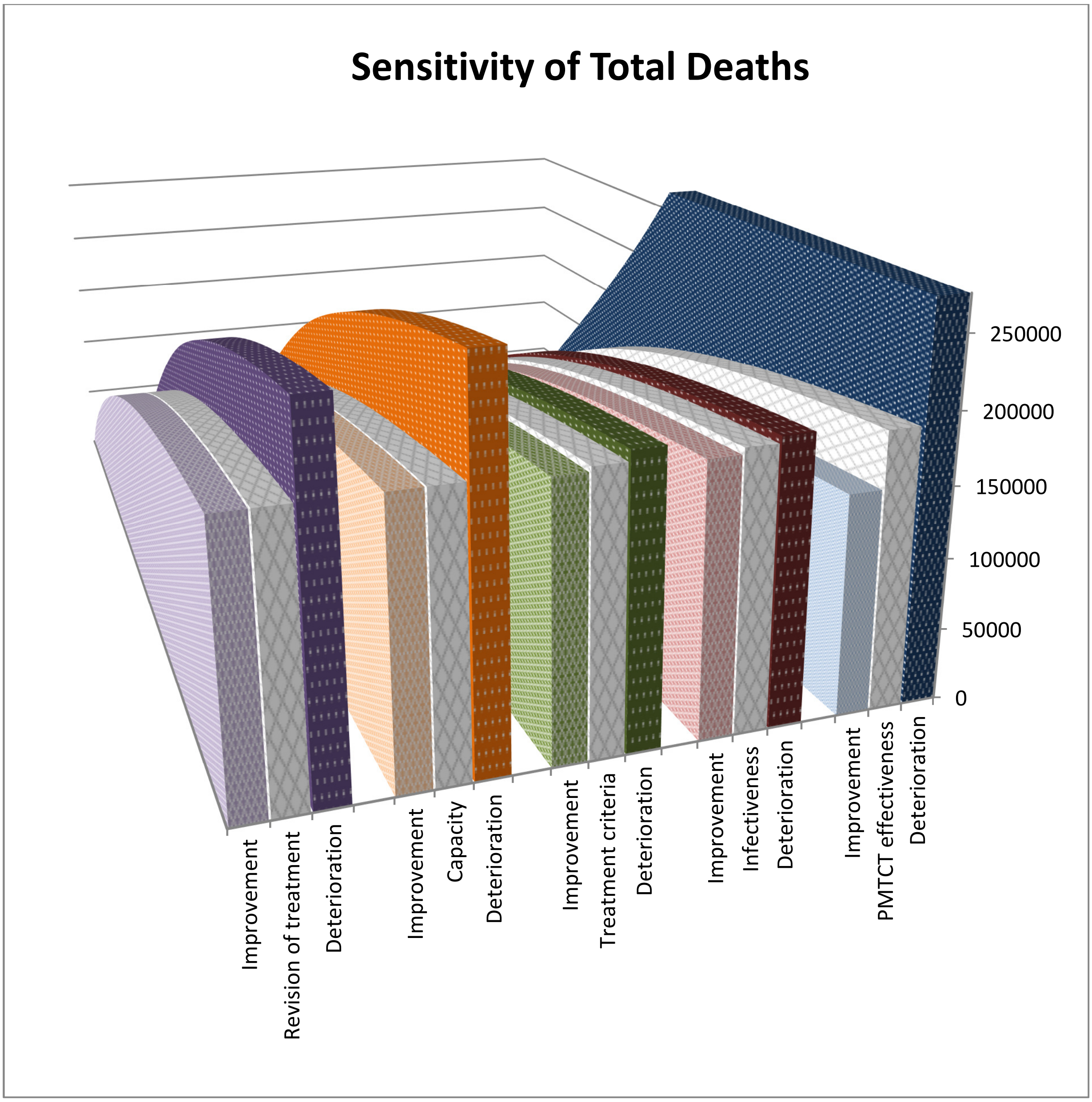


Figure 15: Sensitivity analysis

Recommendation:

It is recommended that interventions be focused on interventions in the uThungulu District that will increase the number of eligible patients. When the capacity of patients on ART is reached the focus should move to increasing the capacity to provide care rather than the need for care.

Special consideration should be taken to ensure that the interventions implemented does not affect PMTCT effectiveness in a negative manner as the system is very sensitive to a reduction in PMTCT effectiveness. When capacity is needed increasing the time between revisions of treatment will increase the available capacity to a large degree, while not requiring any investment.

5.3.1 Framework analysis

To test the framework it will be measured according to the developed criteria

The framework will be considered successful if the following criteria are met:

1. Provide a formal structure to follow when the effect of changes in the system are assessed
2. The approach should address the practical challenges found in the health care environment
3. The approach should be generic enough to be applied at different levels of analysis

With weighting as follows:

Criteria	Weight
Formal Structure	.454
Health Care Specific	.363
Generic across levels of analysis	.181

Table 12: Weight of criteria

Score of the framework for the criteria:

Formal Structure: 8/10 – The framework has a clear layout and gives an easy to follow guide with well-defined steps, especially if the modelling tool is used.

Health Care Specific: 7/10 – All of the content is based on the health care environment. The guidelines provided as part of the framework is sourced from health care relevant resources and addresses the specific challenges found in the environment.

Generic across levels of analysis: 6/10 – The framework scores lower as a generic across levels due to the fact that it so specific, this attribute makes the framework reliant on the ability of the modeller to interpreted the level of modelling and do the analysis appropriately.

Total score for the framework is thus

$$0.454 \times 8 + 0.363 \times 7 + 0.181 \times 6 = 7.26$$

Based on the above analysis of the framework it can be said to have succeeded in the criteria set prior to development.

5.4 Conclusion

The suggested framework from Chapter 4 was applied in a case study to illustrate the use of the modelling exercise through answering the questions posed, and to test how well the framework fits the criteria set to determine its success.

By following the steps indicated in the modelling guide two main insights was gained. First the so-called bottleneck of the system is on the demand for care, and not at the supply side, even though the demand for care, when at a maximum, is close to the capacity for supply. Secondly the sensitivity of the system to some variables especially the importance of keeping the PMTCT effectiveness above 0.75

After the framework was applied in the case study it was measured against the criteria set previously. Based on observation the framework scored 7.3/10 this gives the indication that the framework is successful according to the criteria.

Chapter 6

CONCLUSION AND RECOMMENDATIONS

The application of System Dynamics to health care systems is a growing field of research, but no literature was discovered that gives a clear framework for developing models specific to the health care environment in low resource settings. This dissertation should therefore make a significant contribution to the field of health system modelling.

The importance of a well-functioning health care system is not always realised in our day-to-day activities, but when the need arises for health care interventions it can literary make the difference between life and death. Research advances in the clinical, as well as operational aspects of the health system are common, yet it can be stated that health system performance is still poor. There is thus a need to address the challenges of health care system performance at a system level.

The aim of this dissertation was to develop a framework that will improve the performance of the care system through better understanding of its behaviour by using a formal modelling approach. The framework should be specific to the health care environment, and should be relevant at all levels of health system modelling. To achieve this task the framework for developing the system dynamics model to represent the health system was developed based on content specific to health care systems.

To ensure the correct modelling approaches was studied health care system had to be classified. First a health care system was classified as a complex system, then as an engineering system. With the classification of the system complete techniques that is used to model engineering system was reviewed.

Four different modelling techniques (Monte Carlo-, Discrete Event simulation, System Dynamics- and Agent Based modelling) were reviewed based on the level of application, the requirements to build a model and finally how they were applied in health care previously. Based on the information of the review the characteristics of the modelling approaches were compared to the requirements needed to answer the research question. System Dynamics modelling was chosen as an appropriate modelling approach to model health care systems.

The framework was developed to make use of health indicators as a mechanism to identify the elements that will be used as the components of the model. Health indicators are commonly used as the main source of information to measure the performance of health systems. The framework goes further to provide defined areas applicable to health system to define the scope of the model. From the framework, a guide was written to transform the framework into a practical and accessible tool that can be widely used. To test the appropriateness of the framework and modelling guide they were applied to a case study.

The framework was applied in the case study to answer the question of where resources should be applied to strengthen the HIV program in the uThungulu District in South Africa. The model was developed to illustrate the different states of the infected population, the mechanisms of infection as well as the life expectancy of each group. The calibration of the model was done according to data directly from the District Health Information System. The process of model validation was iteratively applied to ensure the model's validity is sufficient to fit the purpose of the model. The model indicated that the system is very sensitive to the effectiveness of the PMTCT program, but also a ceiling effectiveness is reached at around 80% effectiveness. Further the capacity to initiate patients was highlighted as an area of future concern. Currently the system is operating close to its capacity, and when the capacity is reached there could be significant effects on the rest of the system.

By applying the suggested framework to the case study in uThungulu the framework could be measured against the criteria of performance set to establish if the framework is successful. According to the criteria, the framework performed sufficient to be found useful.

In Chapter 1 the research question was stated:

> *What would be a reasonable framework to aid decision makers to increase understanding of the health system through a modelling approach, while being sensitive to the challenges specific to the health care environment, but be generic enough to be useful in various scenarios?*

Based on the research conducted in the dissertation and the results of the case study the research question can be answered as follows:

> *A modelling framework based on the System Dynamics approach, that is sensitive to the challenges found in health system modelling, and provides clear guidance through the process of model development and gives adequate insight into the understanding of the behaviour of health care systems when applied.*

Even though the research question is answered, there is still a need for improvement in health system performance to sufficiently address the burden of disease experienced in South Africa, and the rest of the world. Health and health care is such a complex environment that the improvement needed in the health care system will not only be found in one area of expertise, but through the collaboration of varies fields of study.

The framework to guide development of health system models is a starting point for the further application of models to help gain insight into the behaviour of health systems, identification of factors affecting health care performance and understanding of how these factors can be addressed. By using this framework, the efficiency of health systems analysis and problem solving can be increased. More areas of expertise will have access to a framework and guide that can be used to solve health system related problems.

For the framework to be used as a day to day tool to aid decision making it should be absorbed into both the public and the private health care sectors. The decision makers should understand the value of a formal approach to aid decision making and what role modelling can play in that approach. For successful implementation, the initiative should be lead from the top to lower levels of authority to ensure buy-in.

Along with the implementation considerations, there are areas for further research and improvement opportunities that can use the work presented in this dissertation. The use of indicators as a guide to identify elements in a System Dynamics model can be applied to various other fields of study where formal system measures are in place. The value of going through a modelling exercise can be clarified between the conceptualisation and the actual modelling of the system. Further the optimisation of the way resources should be applied to modelling to deliver the greatest gain in insight into the system that is modelled. The usefulness of the framework can further be studied when applied in different contexts. The process of modelling can also be translated into a format relevant for workshop-type interactions where only a limited time of contact is possible between contributors.

In conclusion, the research conducted for this dissertation focused on addressing an existing problem found in the health care environment, and finding a suitable solution that can practically be applied to the health care environment. To the best of the author’s knowledge it is the first time that a framework specific for modelling health care systems using System Dynamics has been described in a formal step-by-step manner and applied to a real life problem in the South African health care system.

REFERENCES

[1] NHS Confederation, “The potential of system dynamics,” Leading edge, 2005.

[2] WHO, “World Health Report,” WHO, Geniva, 2000.

[3] South African Department of Health, “National Strategic Plan for HIV and AIDS, STIs and TB, 2012-2016,” DOH, Pretoria, 2012.

[4] Minai AA, Braha D, Bar-Yam Y, “Complex Engineered System: A new Paradigm,” in *Complex Engineered Systems*, Berlin, Springer Berlin, 2006, pp. 1 - 21.

[5] CL, Magee. OL, de Weck, “Complex System Classicication,” Cambridge, 2004.

[6] CL, Magee. OL, de Weck, “An attempt at Complex System Classification,” Cambridge, 2002.

[7] HST, “Health Systems Trust - Indicators,” HST, 01 02 2010. [Online]. Available: http://www.hst.org.za/content/health-indicators. [Accessed 1 2 2012].

[8] L. v. Bertalanffy, General Systems Theory, New York: George Brazeller Inc, 1968.

[9] H. M. Paynter, “Analysis and Desigh of Engineering Systems,” MIT Cource 2.751, 1960.

[10] V. Hubka, W.E. Eder, Theory of Technical Systems, Berlin: Springer-Verlag, 1988.

[11] Sally Brailsford, Nicola Hilton , “A comparison of Discrete Event Simulation and System Dynamics for modelling Health care Systems”.

[12] N. Mustafee, K. Katsaliaki, S. Taylor, “Profiling Literature in Healthcare Simulation,” *Simulation,* vol. 86, no. 8, pp. 543-558, 2010.

[13] G. W. Morgenthaler, “The theory and application of simulation ond operations research,” *Progress in Operations Research,* 1961.

[14] J. Sammet, “History and Fundamentals,” in *Programming Languages*, Prentice-Hall, 1969, p. 650.

[15] R. E. Nance, “A History of Discrete Event Simulation Programming Languages,” Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, 1993.

[16] J. Forrester, "Counterintuitive behaviour of social systems," *Technology review,* vol. 73, no. 3, pp. 52 - 68, 1971.

[17] J. Randers, "Elements of the System Dynamics Method," pp. 117 - 139, 1980.

[18] D. Lane, "You just don't understand me: modes of failure and success in the discourse between system dynamics and discrete event simulation," 2000.

[19] N. Gilbert, Agent Based Models, London: Sage, 2007.

[20] B. Dangerfield, C. Roberts, "Foreword to special issue on health and health care dynamics," *Systems Dynamic Review,* vol. 15, no. 3, pp. 197-199, 1999.

[21] D. Eddy, "Performance measures: problems and solutions.," vol. 17, no. 4, 1998.

[22] Institute of Medicine , "Crossing the quality chasm: A new health system for the 21st century," National Academies Press, Washington, 2001.

[23] G. Hutton, "Indicators for Monitoring Health Sector Reform and the Sector-Wide Approach," in *Moving from Policy to Practice*, Amsterdam, 2000.

[24] WHO, "Global Health Observation," 05 11 2011. [Online]. Available: http://gamapserver.who.int/gho/interactive_charts/mbd/life_expectancy/atlas.html. [Accessed 05 11 2011].

[25] J. Loeb and A. Buck, "From the Joint Commission on Accreditation fo Healthcare are organisations - Framework for selection of performance measurement systems: attributes for conformance.," JAMA, 2006.

[26] Coyne, JS. Hilsenrath, P., "On the other hand The World Health Report," vol. 92, no. 1, 2002.

[27] Internal C0-ordination Group of Health, "Health update," OECD, 2007.

[28] O. Arah, G. Westert, J. Hurst and N. Klazinga, "A conceptual framework for the OECD health care quality indicators project," *International Journal for Quality in Health Care,* pp. 5-13, 2006.

[29] S. Albin, "Building a System Dynamics Model Part 1: Conceptualization," MIT, Massachusetts, 1997.

[30] C. W. Kirkwood, "System Dynamics Resource Page," 2011. [Online]. Available: http://www.public.asu.edu/~kirkwood/sysdyn/SDRes.htm. [Accessed 2011].

[31] D. Kim, "Toolbox: Guide for Drawing Causal Loop Diagrams," *The Systems Thinker,*

vol. 3, no. 1, pp. 5-6, 1992.

[32] G. Richardson and A. Pugh, "Introduction to System Dynamics Modelling with Dynamo," 1981.

[33] Y. Barlas, "Formal aspects of model validity and validation in system dynamcs," vol. 12, no. 3, pp. 183 - 210, 1996.

[34] W. Nordhous, "World dynamics: measurement without data," vol. 83, pp. 1156 - 1183.

[35] A. Zellner, "Comment on: Information Sources for Modeling the National Economy, JW Forrester," vol. 75, no. 371, pp. 567 - 569.

[36] D. Arthur and G. Winch, "Extending model validity concepts and measurements in system dynamics," Plymouth, 1998.

[37] T. Naylor and J. Finger, "Verification of Computer Simulation Models," vol. 14, pp. 92 - 101, 1968.

[38] I. Mitroff, "Fundamental Issues in the Simulation of Human Behavior," vol. 15, pp. 635 - 649, 1969.

[39] J. Bell and M. Bell, "Elements of the System Dynamics method," in *System Dynamics and Scientific method*, Portland, Productivity Press.

[40] Y. Barlas and S. Carpenter, "Philosophical roots of model validation: Two paradiggms," vol. 8, no. 1, pp. 148 - 166.

[41] J. Forrester and P. Senge, "Tests for building confidence in System Dynamics Models," in *Systems Dynamics*, Amsterdam, North-Holland, 1980.

[42] Y. Barlas, "Tests of Model Behaviour That Can Detect Structurel Flaws: Demonstration with simulation experiments," in *International System Dynamics Conference*, 1989.

[43] A. Turing, "Computing Machinery and Intelligence," *Mind,* no. 236, pp. 433 - 460.

[44] E. Carson and R. Flood, "Model Validation: Philosophy, Methodology and Examples," *Trans Inst MC,* vol. 12, no. 4.

[45] N. Mass and P. Senge, "Alternative Tests for Selecting Model Variables," in *Elements of the System Dynamics Method*, Portland, Productivity Press, 1980.

[46] P. Senge, "Statistical Estimation of Feedback Models," vol. 28, pp. 177 - 184, 1977.

[47] J. Sterman, "Appropiate Summary Statistics for Evaluating the Historical fit of System Dynamic Models," *Dynamica,* vol. 10, no. 2, pp. 51 - 66.

[48] J. Stover, L. Bollinger and C. Avila, “Estimating the effects of cost changing guidelines for ART Eligibility,” World Health Organisation, Geneva, 2009.

[49] J. Stover, L. Bollinger and C. Avila, “Estimating the Impact and Cost of the WHO 2010 Recommendations for Antiretroviral Therapy,” Futures Institute, UNAIDS, Glastonbury, Geneve, 2010.

[50] WHO, “Fact Sheet 1 HIV/AIDS,” World Health Organisation, [Online]. Available: http://www.who.int/hiv/abouthiv/en/fact_sheet_hiv.htm. [Accessed 6 2012].

[51] B. Leynaert, A. M. Downs and I. d. Vincenzi, “Heterosexual Transmission of HIV,” vol. 148, no. 1, 1998.

Appendix:

A: Supporting documentation for Delphi-panel review

B: Vensim model and code

A. LETTER OF REQUEST TO PARTICIPATE.

The following letter was sent to experts in the field of health care management as a request to participate in the survey to identify the main elements defining health care systems. The letter to participate was preceded by a telephone conversation describing the background and purpose of the study.

Along with the letter a document was also provided describing the background of the study in more detail.

Request to participate

Your participation in identifying the most influential elements defining the South African Health Care System (sahcs) would be greatly appreciated.

Identification of the elements defining the sahcs

The process of identifying the elements defining the sahcs would be done in three steps. The first step would identify all of the most influential elements according to each participant.

The list of elements would typically be the topics you would discuss when describing the sahcs to someone that is totally unfamiliar to it.

In the second step the list of identified elements would be reviewed by the participants and a short-list of elements would be created. In the final step the shortlist will be reviewed by the participants and a final list of elements would be identified to use in the modelling.

Step 1

To complete Step 1 all that is needed is to respond to this request or to millar.nienaber@up.ac.za with your list of 10 elements. A contact number and time of convience could also be indicated and you would be contacted to indicate your 10 most influential elements.

All of the responses will then be combined and a global list of elements would be published.

The global list of elements would be published online and the following steps would be taken anonymously by identification of elements online.

The link to the global list would be send to all of the participants responding to the request to participate.

The intention of the panel is to be as thorough as possible without wasting your time. Thus a limited window would be identified for each step to compete the panel as effective as possible.

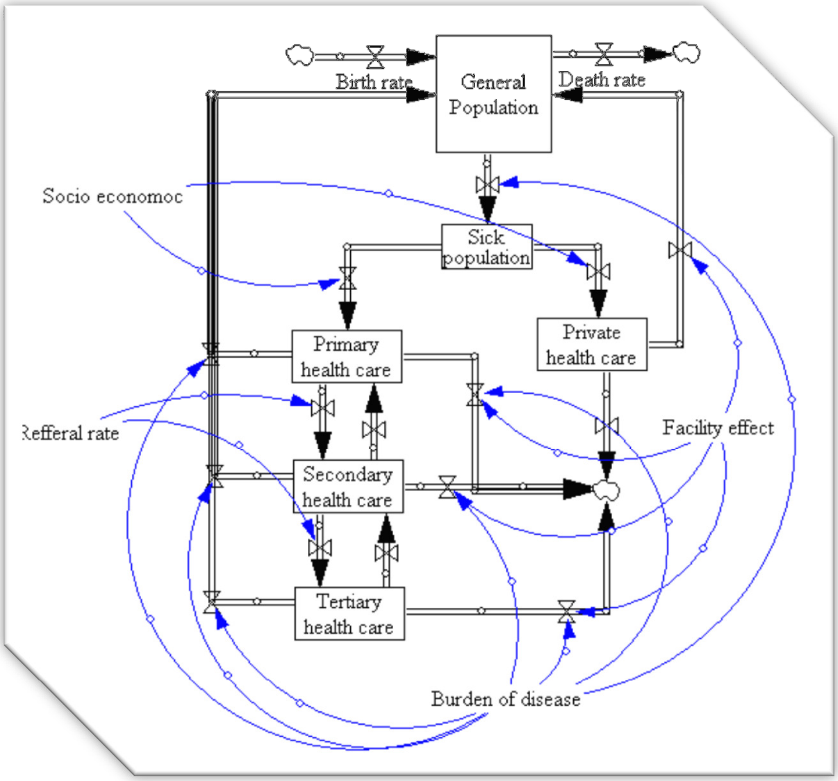
The global list of elements will be compiled from the responses receive by the 5th of September.

Thank you very much for your time, it is greatly appreciated.

Regards

Millar Nienaber

Health care system analysis



This document gives a brief overview of using System Dynamics as a modelling methodology to model and analyse the South African health care system. The use as a Delphi panel to identify the relevant elements en the health care system is also discussed.

Project aim:

The aim of this research project is to create a model of the South African health care system.

This model should be of such nature that it would be expandable, and be used as a reference to

build on in the future.

The purpose of this model is to gain insight into the dynamic nature of the South African health care

system. With the validated model various policy level changes could be simulated and analysed.

Health care system modelling

Simulating and modelling within health care is mainly done by the following methods:

a) Monte Carlo simulation.

b) Discrete event simulation.

c) Agent based modelling.

d) System dynamics modelling.

Through literature and a comparative analysis System Dynamics modelling has been identified as the

appropriate modelling approach to model a health system at a high level.

The inherent complexity of a health care system, and thus the sahcs, is captured in the sd modelling

approach by the attributes.

a) Feedback loops: As one element in the system interacts with other elements there can be a

direct interaction back to the initial element, or through a series of other elements that in

turn impacts the initial element.

An example of scenario this would be a health care facility that underperforms. To counteract this performance an intervention could improve the facility's performance. Now

due to the increased performance clients or patients is drained from other facilities to that

facility, this puts the resources under strain again and the performance is as a result reduced

again.

a) Time delays: There can be a long delay between the time of the implementation of an

intervention and the result thereof. This is especially apparent in health care. The slow

responses of a population to the epidemiology is an example hereof.

b) Non-linear relationships: With non-linear relationships between elements the complexity of

the system increases.

The impact a change of one element has on another is seldom linear. The first R100 spent

per annum per client will have a very different impact than the second R100 spent per

annum per client.

Steps forward

According to Stephanie Albin the modelling of the health care system there are four major steps:

a) Conceptualization

b) Formulation

c) Testing

d) Implementation

In the conceptualization step identification of the elements within the system are needed. By using

System Dynamics as a modelling methodology not all the elements within the system needs to be

modelled. Only the “biggest” and most influential elements need to be used in the model.

The elements must be necessary to the model to represent the reality that is to be modelled. The

elements should be aggregated to reduce the complexity of the model but still represent the nature

of the system that is modelled. Elements should be directional. The element identified should be

able to grow larger or smaller. Using elements like budget or MMR is easier to work with than

population health status.

Identification of the elements

The elements of a system are used to create a stock and flow diagram that is the basis of the

working System Dynamic model.

Identification of the elements can be done by qualitative or quantitative methods. Because of the

high level system modelling applied, quantitative methods are valid for identifying elements.

Brainstorming or panel discussions of field experts are the most common methods used.

Due to the unique structure of the South African health care system I wish to use a Delphi panel of

experts from a wide spectrum of stakeholders in the health care system.

The structure of the Delphi panel will incorporate suggested elements and then rank them according

to group preference. After the identification of the system elements the influences of the elements in the system will on each other will need to be identified.

Use of the research

By identifying the elements that impact the South African health care system with a scientific sound

procedure a validated picture of the given elements driving the health care system can be

generated. With the generation of this picture and the dynamic modelling of the given system

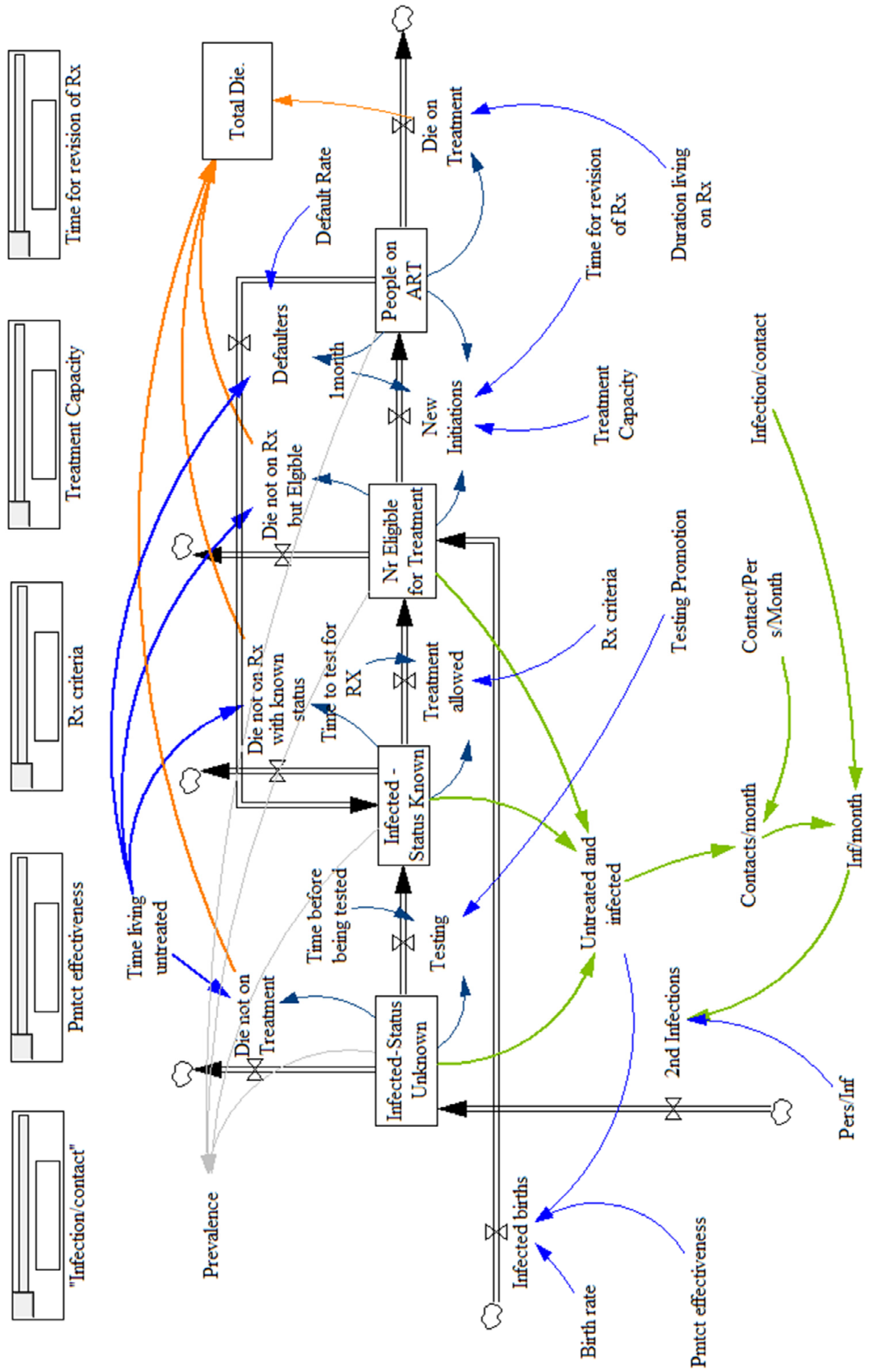
valuable insight will be gained on the dynamic nature of the system over time.

As a result the reaction of the South African health care system to interventions and policy changes

would also be better understood.

B. VENSIM INFORMATION

Vensim model:



Vensim code:

(01) "1month"=1

Units: Month

(02) "2nd Infections"= "Inf/month"*"Pers/Inf"

Units: person/Month

(03) Birth rate=0.043

Units: person/person/Month

Estimated prev rate for Uthungulu using DHIS is Total test pos = 2589 / Total tested for HIV = 15149 = 0.171

(04) "Contact/Pers/Month"=0.5

Units: contact/person/Month

(05) "Contacts/month"= SMOOTH(Untreated and infected*"Contact/Pers/Month",2)

Units: contact/Month

(06) Default Rate= 0.05

Units: Dmnl

(07) Defaulters= Default Rate*People on ART/Time living untreated

Units: person/Month

(08) Die not on Rx but Eligible= Nr Eligible for Treatment/Time living untreated

Units: person/Month

(09) Die not on Rx with known status= "Infected - Status Known"/Time living untreated

Units: person/Month

(10) Die not on Treatment="Infected-Status Unknown"/Time living untreated

Units: person/Month

(11) Die on Treatment=SMOOTH(People on ART/Duration living on Rx, 3)

Units: person/Month

(12) Duration living on Rx=120

Units: Month

(13) FINAL TIME = 240

Units: Month

The final time for the simulation.

(14) "Inf/month"= "Contacts/month"*"Infection/contact"

Units: infection/Month

(15) "Infected - Status Known"= INTEG (Defaulters+Testing-Die not on Rx with known status-Treatment allowed, 30000)

Units: person

(16) Infected births=Birth rate*Untreated and infected*(1-Pmtct effectiveness)

Units: person/Month

(17) "Infected-Status Unknown"= INTEG ("2nd Infections"-Die not on Treatment-Testing,100000)

Units: person

(18) "Infection/contact"= 0.009

Units: infection/contact

See ref of infectiveness by Leynaert *et al.*

(19) INITIAL TIME = 0

Units: Month

The initial time for the simulation.

(20) New Initiations=

max(0,min(Treatment Capacity-(People on ART/Time for revision of Rx),0.4*

Nr Eligible for Treatment/"1month"))

Units: person/Month

(21) Nr Eligible for Treatment= INTEG (Infected births+Treatment allowed-Die not on Rx but Elgible-New Initiations+Infected births,1100)

Units: person

(22) People on ART= INTEG (New Initiations-Defaulters-Die on Treatment-Defaulters,40000)

Units: person

(23) "Pers/Inf"=1

Units: person/infection

(24) Pmtct effectiveness=0.75

Units: Dmnl

(25) Prevalence= ("Infected - Status Known"+"Infected-Status Unknown"+Nr Eligible for Treatment+People on ART)/950000

Units: Dmnl

(26) Rx criteria=0.5

Units: Dmnl

Estimating cost of eligibility change - WHO

(27) SAVEPER = TIME STEP

Units: Month [0,?]

The frequency with which output is stored.

(28) Testing=Testing Promotion*"Infected-Status Unknown"/Time before being tested

Units: person/Month

(29) Testing Promotion= 1

Units: Dmnl

Total tested / total est being Pos

(30) Time before being tested= 15

Units: Month

Quarter of infected untreated life

(31) Time for revision of Rx=1

Units: Month

(32) Time living untreated=60

Units: Month

5 Years used as untreated std

(33) TIME STEP = 1

Units: Month [0,?]

The time step for the simulation.

(34) Time to test for RX=2

Units: Month

(35) "Total Die."= INTEG (Die not on Rx but Elgible + Die not on Rx with known status + Die not on Treatment + Die on Treatment, 0)

Units: person

(36) Treatment allowed= Rx criteria*"Infected - Status Known"/Time to test for RX

Units: person/Month

(37) Treatment Capacity=161361

Units: person/Month

Total headcount *.5 for HIV at 80 pers utilization

(38) Untreated and infected="Infected-Status Unknown"+"Infected - Status Known"+Nr Eligible for Treatment

Units: person