The Role of Computational Epidemiology and Risk Analysis in the Fight against HIV/AIDS

Berhanu Tameru, David Nganwa, Asseged B. Dibaba, Vinaida Robnett,
Center for Computational Epidemiology, Bioinformatics and Risk Analysis (CCEBRA)
College of Veterinary Medicine, Nursing and Allied Health (CVMNAH)
Tuskegee University, Tuskegee, Alabama, USA

Gemechu B. Gerbi
Epidemiology and Surveillance Branch, Division of Viral Hepatitis, NCHHSTP
Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Tsegaye Habtemariam
Center for Computational Epidemiology, Bioinformatics and Risk Analysis (CCEBRA)
College of Veterinary Medicine, Nursing and Allied Health (CVMNAH)
Tuskegee University, Tuskegee, Alabama, USA
1 Introduction

The devastating effects of the HIV/AIDS pandemic are compounded by its complex patterns of transmission and spread. The rate of its transmission and the demographic spread of the disease both in time and space are influenced not only by direct factors such as age, gender, marital status, number of sexual partners, sexual preferences, frequency of extramarital liaisons, etc., but indirectly by factors such as psychosocial, socioeconomic, HIV/AIDS risky behaviors and ultimately by the complex interactions and interplay between all these factors. It is these aforementioned interactions and their evolution over time that provides a major difficulty in trying to predict the spread and social impact of the disease (Tameru et al., 2012; Gerbi et al., 2012; Diaz et al., 1994). Teasing out these various factors using the epidemiologic problem oriented approach methodology will greatly facilitate in developing knowledge bases that are an integral part in incorporating these factors in the different models developed.

In general a model can be considered as a pattern, plan, representation (especially in miniature), or description designed to show the main object or workings of an object, system, or concept. In Scientific modeling, modeling is considered as being the process of generating abstract, conceptual, graphical and or mathematical models which are implemented in computer models to study the behavior of a complex system by computer simulations. Science offers a growing collection of methods, techniques and theories about all kinds of specialized scientific modeling. Modeling is an essential and inseparable part of all scientific activity, and many scientific disciplines have their own ideas about specific types of modeling. Several researchers have developed computational tools and mathematical approaches to study the effects of various mitigation methods on HIV/AIDS at different population levels, in vivo or in vitro, macro to micro levels, just to name a few (Ho et al., 1995; Perlson & Nelson, 1999; Habtemariam et al., 2001; Tam et al., 2008; Tameru et al., 2012). The approaches range from compartmental models represented by sets of differential equations (Anderson & May, 1992; Rvachev & Longini, 1985; Sattenspiel & Simon, 1988); to highly complex individual-based models which represent daily activities and interconnections of individuals via transmission networks (Eubank et al., 2004). Compartmental models can be easily solved, but they cannot model adaptive behaviors of individuals and complex interactions of different groups of populations during disease outbreaks. While individual-based models like Agent Based Modeling can capture the spread of diseases with high-fidelity, modeling large populations often requires utilization of supercomputers and makes it impractical for quick what-if analyses of interventions or treatments under varying different conditions.

This volume which is an update review chapter provides a synthesis of our collective knowledge regarding the application of computational epidemiologic modeling and risk analysis in addressing HIV/AIDS pandemic in both macro and micro settings undergirded by sound epidemiology. Numerous case examples with different approaches for modeling the AIDS pandemic and the HIV pathogenesis in HIV infected individuals are given. The role and importance of computational epidemiologic modeling are discussed. A new avenue of research, involving the application of risk assessment in solving public health problems in general and its suitability in modeling HIV/AIDS dynamics in particular is explored.

1.1 Computational Epidemiologic Models

The use of computational epidemiologic models, undergirded by sound epidemiologic and mathematical principles, has brought a substantial progress in the understanding of HIV and CD4 cell dynamics. In its early stages, the applications of these models were based on relatively simple mathematical templates that considered the body as a one compartment system. In spite of being very attractive, from the experi-
mental and/or mathematical standpoint of view, the underlying simplification means that a lot of factors that significantly affect the population dynamics both at macro (human) and micro (cellular) population levels are effectively avoided. This simplification also affects the kinetics linked to the infection, immunology, and chemotherapy dynamics throughout the host. As epidemiologic research involves a complex set of host, environment and agent factors as they interact to impact health in any given population (whether biotic or abiotic), generating large datasets which require the use of more advanced computational methods is a necessity. Computational epidemiologic methods are useful for studying such large and complex models (Habtemariam et al. 2001; Wai-Yuan & Hulin, 2008; Tameru et al., 2012). Another dimension of the challenges faced by the public health decision makers is epitomized by the emerging/re-emerging diseases, as they have to face and deal with a lot of uncertainty at the early stages of new disease outbreaks.

1.2 Risk Analysis

Epidemiologic problem-solving and decision-making often proceeds in the face of uncertainties and limited information. Emerging diseases are one such challenging problem to public health decision makers since they encounter a lot of uncertainty even at the early stages of disease outbreaks. Risk analysis is a process for decision making under uncertainty that consists of three fundamental tasks: risk management, risk assessment, and risk communication. Risk analysis can be considered as the process of examining the whole of a risk holistically by assessing the risk and its related relevant uncertainties for the purpose of its efficacious management, facilitated by effective communication skills about the risk. It is a systematic way of gathering, recording, and evaluating information that can lead to recommendations for a decision or action in response to an identified hazard or opportunity for gain. Risk analysis is not a science rather science based; it is not certain (usually expressed in probability distributions & confidence intervals), it is not a solution (from all possible options, the best one is picked under the pertinent available information at that point in time); and it is not static (needs to be updated whenever new information is available). In the context of public health, Quantitative Risk Assessment is the process that quantifies the likelihood of the introduction of infection, it’s establishment, or spread of a pathogen or disease in a given susceptible population. Standard epidemiological procedures are utilized to systematically evaluate health risks from the combined effects of multiple factors that lead to the identified hazard scenarios for each step in the process; mathematical and computing methods are utilized to assess the magnitude of the risk and mitigations effect. The ultimate goal is to support and facilitate public health officials in making informed decisions based on organized science based analyses.

1.3 The Epidemiologic Framework to Risk Analysis

Definitions of epidemiology and risk analysis as pertains to this book chapter are that: Epidemiology is defined as the study of the dynamics of health/ill health processes in a given population. It is often directed at problem solving and decision or policy making at the population level. On the other hand, risk analysis is defined as the practice of decision making based on scientific evidence (Risk Newsletter, 2000). Like epidemiology, risk analysis is often focused on population-based studies although both methodologies can be applied to any population such as cellular, molecular, genomes and others. This is because in epidemiology, the population under study can be groups of animals (e.g., herd health), humans (e.g., public health), plants (phytoepidemiology), cellular and molecular populations (molecular epidemiology), or populations of genes (genetic epidemiology). Epidemiology is a discipline that can be applied to the study of population dynamics from the molecular (microepidemiology) to higher levels (macroepi-
demiology) of population dynamics. This breadth of epidemiology provides risk analysis with the framework for its application in a vast array of population-based studies from genomics and biotechnology to even global health/international trade. The link between epidemiology and risk analysis is rational and intuitive. The two areas complement and supplement each other. In epidemiology, the basis for reasoning and explanation, the opportunity for dealing with choices, risks or benefits especially in the face of uncertainties, and the need to analyze and manage imperfect data are common occurrences. The same applies to risk analysis. The dilemma is that the paucity, incompleteness and uncertainty of available data further complicate quantitative models. Yet, epidemiologic problem solving and decision-making as well as risk analysis often must proceed in the face of uncertainties and limited knowledge. A risk assessment is never complete nor is it static. As more knowledge and information is gained over time, a risk analysis particularly the risk assessment component can be revised and updated as appropriate. To handle the types of challenges described above, computer modeling (Risk Analysis and Epidemiologic Modelling) provides a powerful alternative tool to traditional empirical (field or laboratory based) studies. Computational epidemiology provides a mechanism for approximating biological interactions, via bio-mathematical expressions that can be tested using computer models as the experimental medium. This new approach is the realm of computational science (Pool, 1992). Computational science integrates the two traditional areas of empirical and theoretical sciences. It also builds upon and extends the methods and tools available to research by exploiting computational resources. Computational epidemiology and risk analysis now provide alternative avenues where systems, which may be complex, too large, and not feasible because the information is scanty and uncertain; or the cost is too prohibitive, can be approximated and simulated realistically.

2 History of Computational Epidemiology and Risk Analysis Models

Computational epidemiologic models and simulations are emerging as vital research tools in epidemiology, biology, and various other fields in advancing the bench (wet) lab and public health policy research agenda. Scientists are recognizing the huge potential of these tools in solving some of today’s biggest health problems. Computational epidemiology is a multidisciplinary field utilizing techniques from computer science, mathematics, geographic information sciences and public health to develop tools and models to aid epidemiologists and other scientists in their studies of the temporal-spatial spread of diseases. Research in computational epidemiology is now considered to be an exponentially expanding arena of scientific exploration. In particular the HIV/AIDS pandemic, where more than 25 million people have so far died from the disease since 1981, making it amongst the most serious threats to global health that we face today. Epidemic models of infectious diseases date back to Daniel Bernoulli’s mathematical analysis of smallpox in 1760 and have been developed extensively since the early 1900s. Mathematical epidemiologic modeling, with the help of computational tools, has provided new insights on such important issues as drug resistance, rate of spread of infections, epidemic trends, and effects of interventions such as treatment and vaccination.

The term computational epidemiology was first coined by Professor Tsegaye Habtemariam (a founding member of the Center for Computational Epidemiology, Bioinformatics and Risk Analysis (CCEBRA) at Tuskegee University) (Habtemariam et al., 1988) to better understand the complex biomedical systems in diseases like the HIV/AIDS pandemic. Computational epidemiology enables infectious and non-infectious diseases and risk agents of plants, animals and humans to be examined and in-
vestigated without jeopardizing lives or creating hazards. This relatively young strand of computational science is being used to understand a range of problems from soybean and wheat rust to HIV/AIDS, swine influenza, foot and mouth disease, rift valley fever and bioterrorism to name a few. In light of this, computational epidemiology has the potential to influence global issues that both directly and indirectly affect human, animal, plant health and the environment, representing a milestone in modern science (Tameru et al., 2012). The CCEBRA is a unique facility that has been involved since the early 1980s in groundbreaking work in this niche of science, that of Computational Science. Epidemiologic research involves the study of a complex set of host, environment and causative agent factors, with the most advanced of these efforts focusing on micro (cellular/molecular) and macro (host) population levels.

Computational epidemiologic models of HIV/AIDS provide important insights in population dynamics through studies at the molecular and cellular levels as well as at the human population level (Ho et al., 1995; Habtemariam et al., 2001; and Tameru et al., 2008). As of today, wet lab science has not yet achieved a level of success to produce a viable vaccine or effective medication for preventing or treating HIV/AIDS. The alternative research approach of computational modeling (the so called the third dimension of science), has also been a priority for researchers in other disciplines.

While risk analysis especially the risk assessment component has existed in various forms for many years, the process used by US Environmental Protection Agency (EPA) and others was formalized in the pivotal 1983 National Research Council (NRC) report known as the “Red Book” (National Research Council, 1983). The Red Book codified the well-known four steps of risk assessment (hazard identification, exposure assessment, dose-response assessment, and risk characterization) and it emphasized the necessity of a conceptual distinction between risk assessment, risk management and risk communication. Over the intervening quarter-century, risk assessment has evolved substantially, driven in part by additional NRC reports, EPA, World Trade Organization, and other agency guidelines, and publications in the peer-reviewed literature. There are two major types of risk assessments namely Qualitative Risk Assessments and Quantitative Risk Assessments (Probabilistic Risk Assessment (PRA)). The PRA is used to estimate risk by computing probability distributions to determine what can go wrong, how likely is it to happen, and what are its consequences. Health (human, animal and plant) PRA provides insights into how the risk propagates from the source to the end point (Scenario tree); how likely is each scenario to happen; what will be the consequence (e.g., the number of people potentially infected or killed); the efficacies and weaknesses of different mitigations to reduce the risk and associated consequences. However, there is a need in expanding the use of PRA in human health as most of the human health risk assessments are of environmental and food safety in nature.

3 The Steps Utilized in Developing Computational Epidemiologic Models

Systems dynamic modelling (SDM), a tool widely used in epidemiological and mathematical modelling, allows researchers and scientists to study and develop a holistic way to assess not only the behavior of the system, but the relationships and interactions between different entities within the system so that scientists can predict what will happen if these systems behaviors persist over time into the future.

Systems dynamic modelling is a concept based on systems thinking whereby dynamic interactions between the elements of the system is considered in order to study the behavior of the system as a whole. This methodology, introduced in the mid-1950s by Forrester and first described at length in his book “Industrial Dynamics” (Forrester, 1961) with some additional principles presented in his later works (Forrest-
er, 1985), involves development of causal diagrams and computer simulation models that are unique to each problem setting. A central principle of SDM is that the complex behaviors of organizational and social systems are the result of ongoing accumulations of people, material or financial assets, information, or even biological or psychological states. Both balancing and reinforcing feedback mechanisms and the concepts of accumulation and feedback have been discussed in various forms for centuries. However, SDM uniquely enables the practical application of these concepts in the form of computerized models so that alternative policies and scenarios can be tested in a systematic way that answers the questions of “what if” and “why”.

The systems analysis approach to model development consists of seven major steps which are all interlinked and we will follow these seven steps in sequence to illustrate how to develop computational epidemiologic (risk assessment) models.

- **Step 1**: Develop the Epidemiologic Problem Oriented Approach (EPOA) methodology, (to collect, synthesize and organize the epidemiologic data).
- **Step 2**: Create a conceptual systems model diagram based on the EPOA. This involves defining the subsystems in the model and conceptualizing their relationships and interactions with each other (see Figure 2).
- **Step 3**: (i) Dynamic model development - the systems dynamics and the underlying structure for the mathematical formulations will be described by means of differential (difference) and partial differential equations (ii) risk assessment model development: (a) identifying the hazard (i.e., HIV/AIDS); (b) developing a scenario tree which outlines a series of mitigations and all the failures which could occur, culminating in the occurrence of the identified hazard; (c) gathering and documenting the evidence.
- **Step 4**: Develop a Mathematical Model
- **Step 5**: Develop a Computer Simulation Model (A simulation model)
- **Step 6**: Test, Validate, Perform Sensitivity Analysis, and update the model (update the knowledge base). Compare the models’ response/behavior/performance to reality, other models, or published works
- **Step 7**: Implement the Model.

As an example a systems dynamics model for HIV/AIDS identifies links between the sub-systems, state variables, rate variables, parameters and constants in a system. The developed models can be used to provide important insights in population dynamics at the macro (human) population level and at the micro (cellular) population level and helps in giving insight for alternative HIV/AIDS control and prevention strategies.

### 3.1 Knowledge Base: The Epidemiologic Problem Oriented Approach (EPOA) Methodology

In the epidemiologic modeling of HIV/AIDS and other diseases, the Epidemiologic Problem Oriented Approach (EPOA) methodology facilitates the development of systematic and structured knowledge bases, which are crucial for development of computational epidemiologic models (Nganwa *et al.*, 2010). A detailed analytic understanding of the epidemiology of a population under study and a decomposition of all relevant determinants of health and disease provides the essential framework for the development of computational epidemiologic and risk assessment models and enables the laying out of the comprehen-
sive and fundamental structures for the models. The method of decomposition of any epidemiologic or risk assessment task relies heavily on EPOA. As in any problem solving and decision making exercise, the EPOA essentially consists of a problem identification/definition/characterization component, followed by a problem management/solution/mitigation component. We use the classical epidemiologic triad (epidemiologic triplet) consisting of host, agent and environment interactions, and examination of agent transmission pathways and spread of disease both in time and space as the first key step to computational epidemiology or risk assessment (problem identification triad). Rational intervention strategies (mitigations) that minimize the risk of transmission and introduction of a disease or pest are then integrated into such an epidemiologic framework. The second set of triad, composed of prevention/control, treatment or therapeutics to eliminate a risk agent and health maintenance/promotion is the decision making step (problem/management/solution/mitigation triad).

The two triads are interlinked by diagnostic linkage procedures used in identifying and characterizing the risk agent when possible. The individual pillars of each triad are interlinked and intertwined. Each pillar of the triads is decomposed into its respective variables and parameters.

The first triad the Problem Identification/Characterization Triad: The agent pillar; identifies the agent and its characteristics like infective dose and route(s) of infection, survival under different conditions, pathogenicity, life cycle, transmission pathways etc. The host pillar; identifies and characterizes the characteristics of all possible hosts whether they are definitive, intermediate, reservoir or paratenic. Host characteristics are identified in detail including intrinsic and extrinsic factors. The environment pillar; characterizes the physical (abiotic), biological (biotic) and socio-economic environments for both the host and agent and how they interplay. Psychosocial factors and determinants are considered. In the second triad Problem Management/Solution/Mitigation Triad: The therapeutics/treatment pillar; considers if the disease is treatable curatively, palliatively or for secondary problems. Options, ease of availability and accessibility to treatment are taken into consideration; Prevention/control pillar; considers if prevention is primary, secondary or tertiary; and the Health Maintenance/Health Promotion pillar; considers in general the health maintenance of the population mainly after a disease or condition has already occurred and is geared towards lowering the prevalence and incidence rates to their lowest level coupled with prevention, control strategies and eventual eradication of the disease being the ultimate goal.

Although the generic term risk analysis is composed of: a) risk assessment, b) risk management, and, c) risk communication, our emphasis in this chapter is on risk assessment. These components viewed through the EPOA methodology are part of the classical problem solving steps of: a) problem identification and characterization (risk assessment), and, b) problem management (risk management and risk communication). It is noteworthy to emphasize that both risk management and risk communication rely on sound risk assessments, which may be qualitative or quantitative in nature.

Risk mitigation in this chapter is broadly defined to include all activities and resources required to: a) prevent introduction of risk agents, b) eliminate the risk agent if possible, and/or c) manage the risk event by taking steps to minimize or reduce the risk of spread once introduced into a disease free population (for macro on population level modeling) or susceptible individual (for micro modeling). We contend that effective approaches to risk management rely upon: a) sound science-based risk assessment which in turn depends on a detailed understanding and decomposition of the epidemiologic factors and the transmission pathways for the risk agent under study, and, b) education and information sharing (nationally and internationally).
Data collection is crucial and central for development of the knowledge base, a lot of information and data have to be gathered depending on disease under study in this case HIV/AIDS. Sources and types of organizations that can provide or direct you to the key information include: academic and research institutions, ministries of health, other government agencies, hospitals and clinics, nongovernmental and community-based organizations, international organizations and partners involved in HIV/AIDS work, such as USAID, CDC, UNAIDS, WHO, national and/or regional associations of People Living With HIV/AIDS (PLWHA), private companies, media and the internet. Comprehensive program reviews produced by governments, donor agencies, and others, National HIV/AIDS program updates, including reports from behavioral and biological surveys, and from sentinel surveillance, Web-accessible libraries, meeting and conference reports, books, journals, and medical databases generated by research endeavors. The EPOA organizes this data collected into a well-structured format that is easily retrievable in the process of identifying variables and estimating parameters used in model development.

### 3.2 Conceptual Systems Dynamics Model Diagram

Once the knowledge base is developed based on the EPOA methodology a conceptual systems dynamics model diagram needs to be developed. The systems dynamic modelling (SDM) is an iterative process of scope selection, hypothesis generation, causal diagramming, and quantification; it consists of an interlocking set of differential and algebraic equations developed from a broad spectrum of relevant data. A completed SDM model may contain scores or even hundreds of equations along with the appropriate numerical inputs. Importantly, epidemiologic SDM models are designed to reproduce historical patterns and capable of generating useful insights. The data extrapolated from these epidemiological models are useful
not only to study the past, but are reliable also to explore predictive and intervention possibilities (Forrester, 1960, 1985). With this in mind, a SDM model incorporating various HIV/AIDS-risky behaviors has been developed to model HIV/AIDS.

3.2.1 A Macro-epidemiologic Model of HIV Transmission

The dynamic epidemiologic model developed in the macro level of the transmission of HIV and its progression to AIDS relies on a set of multiple determinants that affect the epidemiology of HIV/AIDS in populations. At the macro level, the population is divided into three sub-populations based on their health status. These include those who are susceptible (S), infected with HIV (I), and advanced state of HIV infection or full blown AIDS (A). The transitions between the states of health are regulated by the respective rates such as birth rate, infection rate, progression rate to AIDS and death rate respectively see Figure 2. The macro model considers five ethnic populations: Whites (not Hispanic), African Americans, Hispanics, Asian/Pacific Islanders, and American Indian/Alaska Natives. Within each ethnic group, each individual has the demographic characteristics (age, gender, etc.) and HIV/AIDS risky behaviors: male to male sexual contact (MSM), injection drug use (IDU), male to male sexual contact and injection drug use (MSM/IDU), high risk heterosexual contact, and others (include hemophilia, blood transfusion, perinatal exposure, and risk factor not reported or not identified). The HIV/AIDS infection rate in a given susceptible population directly depends on the proportion of MSM, IDU, MSM/IDU, high-risk heterosexual contact, and other risk factors. Manipulation of one or several of these variables changes the behavior of the system dynamics and result in an increase or decrease of the incidence of HIV/AIDS; thus allowing critical evaluation of alternative disease control strategies.

3.2.2 Micro-epidemiologic Modelling

Cellular level modeling (CD4+ lymphocyte/HIV viral population dynamics): The host populations are CD4+ lymphocytes, the agent is HIV viral population and the environment is the cellular and intracellular/molecular ecosystems. The new state that represents infected CD4+ lymphocytes, referred to as “HIV infected CD4+ cell subpopulation”, is further subdivided into four sub-states (Coffin 1995, Fauci 2014): Productively infected, Latently infected, Defectively infected, and Chronically infected. Molecular level modeling: The CD4+ lymphocyte cell is assumed as infected as soon as the virion enters the host cell. Ordinary differential equations were used to mathematically represent the viral kinetics as they move from reverse transcription through progeny formation and maturation.

4 Mathematical Epidemiology and Risk Assessment Model Applied in HIV/AIDS

Based on the conceptual model developed, a mathematical model which consists of several sub-models is developed. A mathematical model is more detailed, and is more based on data, than the conceptual model. If developed carefully, mathematical and statistical models can serve as tools to better understand the epidemiology of HIV/AIDS. Mathematical models of HIV/AIDS transmission dynamics also play an important role in understanding the epidemiological patterns and methods for disease control as they provide short and long term predictions of HIV and AIDS incidence and prevalence trends, and its dependence on various factors.
Figure 2: An epidemiologic systems model of macro-micro (human-Cellular) Population dynamics.
4.1 Mathematical Epidemiology

The framework for representing the integrated components of the HIV/AIDS epidemiologic model including macro level population dynamics as well as micro level population cellular/molecular level biological dynamics is shown in Figure 2. The dynamics of the model is described using a system of partial differential equations (Tameru et al., 2010).

\[
\frac{\partial S_{ijk}(a,t,\bar{u})}{\partial t} + \frac{\partial S_{ijk}(a,t,\bar{u})}{\partial a} = \left[ \sigma_{ijk}(\bar{u}(t),a) \left[ 1 - \gamma_{ijk}(\bar{u}(t),a) \right] - 1 \right] S_{ijk}(a,t,\bar{u}) \tag{1}
\]

where \( S_{ijk} \) denotes the number of susceptible individuals of drug use status \( k \), sex related status \( j \), ethnic group \( i \), age \( a \) at time \( t \); \( \sigma_{ijk}(\bar{u}(t),a) \) is the age-group specific survival rate of individuals of drug use status \( k \), sex related status \( j \) in ethnic group \( i \); \( \gamma_{ijk}(\bar{u}(t),a) \) is the HIV infection rate of individuals of age \( a \) at time \( t \), drug use status \( k \), sex related status \( j \) in ethnic group \( i \), which depends on time of infection and the number of CD4+ count and the viral load at time \( t \) and cellular level dynamics of CD4+, \( \bar{u}(t) \), with \( \bar{u} = (C_i(t,\tau),V_i(t,\tau)) \) \( C_i \)'s are infected cells and \( V_i \) is the viral population, and age \( a \) of the individual.

\[
\frac{\partial C_i(t,a,[dNTP])}{\partial t} + \frac{\partial C_i(t,a,[dNTP])}{\partial a} = -d_i C_i(t,a,[dNTP]), \tag{2}
\]

with

\[ C_i(t,0,[dNTP]) = \beta C_i(t,[dNTP])V(t,[RNA_{cor}]) \text{ for } i \in \{L,P,C,D\}, \]

and

\[
\frac{\partial V(t,[RNA_{cor}])}{\partial t} = \sum_{i \in \{L,P,C\}} \int_{a_{iL}}^{a_{iP}} \gamma_i(a)C_i(t,a,[dNTP])da + u_i V(t,[RNA_{cor}]),
\]

where for \( i = D \), defectively infected CD4+ cells, for \( i = L \), latently infected CD4+ cells, for \( i = P \), productively infected CD4+ cells, and for \( i = C \), chronically infected CD4+ cells and \( a \) is the age of the CD4 cell. A selected equation representing the molecular kinetic level interactions is shown below.

\[
\frac{d[RNA_{cor}]}{dt} = -\frac{V_m[RNA_{cor}]}{K_m[RNA_{cor}]} \frac{[dNTP]}{K_m[dNTP] + [dNTP]} - \left( k_{RNA_{cor}} + \phi_{RT} \right)[RNA_{cor}], \tag{3}
\]

where \([RNA_{cor}]\) is the concentration of genomic RNA present in the viral core and \([dNTP]\) is the concentration of the dNTP pool of the host cell. \( K_m[RNA_{cor}] \) and \( K_m[dNTP] \) are the Michaelis constants for reverse transcriptase with the substrates \( RNA_{cor} (= 2*V_F) \) and the dNTP (Deoxyribonucleoside triphosphate) pool, respectively. \( k_{RNA_{cor}} \) is the degradation rate constant of the genomic RNA and \( \phi_{RT} \) is the efficacy of the drug for reverse transcription inhibitor.

Other equations representing the various other states were developed in a similar fashion. The equations that describe the dynamics in HIV infected populations of drug use status \( k \), ethnic group \( i \), age \( a \) at time \( t \) are defined as follows:

\[
\frac{\partial I_{ijk}(a,t,\bar{u})}{\partial t} + \frac{\partial I_{ijk}(a,t,\bar{u})}{\partial a} + \frac{\partial I_{ijk}(a,t,\bar{u})}{\partial \bar{u}} = \left[ \sigma_{ijk}(\bar{u}(t),a) \left[ 1 - tr(\bar{u}) \right] - 1 \right] I_{ijk}(a,t,\bar{u}), \tag{4}
\]
where \( tr((\bar{u}(t),\tau)) \) is the probability that an individual infected by HIV at time \( t-\tau \) becomes an AIDS patient at time \( t \), which is assumed to be the same for all ethnic groups. A similar equation can be given for the dynamics of AIDS patients populations.

\[
\frac{\partial A(A_{ijk}(a,t,\bar{u}))}{\partial t} + \frac{\partial A(A_{ijk}(a,t,\bar{u}))}{\partial a} + \frac{\partial A(A_{ijk}(a,t,\bar{u}))}{\partial u} = \{\mu(\bar{u}(t),a) - 1\} A_{ijk}(a,t,\bar{u}). \tag{5}
\]

The role of individual characteristics from macro (age, gender and race), socioeconomic status (level of education, level of income and employment status), and psychosocial factors to the cellular level CD4+ count, HIV viral load, and the kinetics inside the cell are also incorporated in this model. We let \( F_{ijk}(t) \) denote the events that an individual of drug use status \( k \), sex related status \( j \), in ethnic group \( i \) is infected by HIV during \([t, t+dt)\) due to sexual contact. An individual may have sexual contacts with partners from different ethnic groups. The probability of HIV transmission due to sexual contacts is formulated in terms of the number of partners, number of sexual contacts with each partner, the probability that a partner is infected, the HIV viral load of the infected and the probability that one contact with an infected partner will result in infection. In this study, since we consider six ethnic groups, each consisting of five risky behaviors related sub groups, the HIV prevalence differs from group to group. The probability that an individual of drug use status \( k \), risky behaviors related status \( j \) in ethnic group \( i \) is infected by HIV at time \( t \) due to sexual contacts is given by:

\[
P[F_{ijk}(a,t,\bar{u})] = 1 - \prod_{e=1}^{3} \{1 - q_e(t)\}, \tag{6}
\]

where

\[
q_e(t) = 1 - \left\{1 - p_e(t)\left[1-(1-r)^{m_{ijk,e}}\right]\right\}^{n_{ijk,e}}
\]

is the probability that an individual of drug use status \( k \), risky behaviors related status \( j \) in ethnic group \( i \) is infected by HIV during \([t, t+dt)\) due to sexual contacts with partners from ethnic group \( e \), \( r \) is the probability of HIV transmission associated with a single sexual contact, \( n_{ijk,e} \) is the number of sexual partners from ethnic group \( e \), \( m_{ijk,e} \) is the number of sexual contacts with a partner from ethnic group \( e \), and \( p_e(t) \) is the probability that a partner from ethnic group \( e \) is infected at time \( t \).

### 4.2 Risk Assessment

Quantitative risk assessment (QRA) was performed after developing the risk pathway scenario tree based on a comprehensive review of published literature that have examined psychological, social and interpersonal variables as correlates of sexual risk behaviors in people who know they are HIV positive (Gerbi, et al., 2012). The main focus in this chapter is on individual level factors influencing HIV/AIDS risky behaviors. QRA provides the methods of measuring risks and provides decision makers with the information needed to make decisions which are scientifically based. The QRA process involves: (a) identification of the hazard, (b) developing a scenario tree outlining the pathway of expected events and all the failures that are likely to occur, (c) gathering and documenting evidence, (d) developing equations or functions, (e) performing calculations to summarize the likelihood of the hazard occurring, (f) considering risk management options and (g) preparing a written report.
The risk pathway presented in Figure 3 consists of a sequence of specific events. For each node or event, a specific question related to the risk of unprotected sex is asked. The product of the probabilities of these answers to these questions determined the final risk related to HIV transmission through unprotected sex. In constructing the risk pathway, the epidemiology and determinants of HIV/AIDS transmissions provided the framework from which the risk assessment was conducted. Quantitative Risk Assessment requires that each parameter should be described and scientific evidence should be presented to justify the parameter estimates. This risk assessment relied on studies that have examined psychological, social, interpersonal, and medical variables as correlates of sexual risk behavior in people who know they are HIV positive. The quantified parameter values of the model were presented in terms of beta distributions which are used to determine the likelihood of HIV infection in men through unprotected sex. The beta distribution is seen as a suitable model in risk analysis and it is widely used to model probability distributions of variables in many areas of research (Soumyo, 1990).
5 Experimentation using the Computational Models

Once the development and integration of the computer modeling methodologies are completed, several simulations need to be conducted with varying initial and boundary conditions to test the validity of the model. The next step is the enhancements adjustments and modifications that need to be made until the model exhibit outputs which are biologically and mathematically reasonable and plausible. Sensitivity analysis must be performed to examine model stability too.

5.1 Macro Level

The macro level computational model developed integrated three stochastic and dynamic sub-models; one to represent AIDS at the human level, the second to represent CD4+ population dynamics at the cellular level, and the third to represent the kinetics at the intracellular viral kinetics level. Using the integrated model as the experimental medium, computer simulations were conducted to examine and answer specific scientific questions. The results of computational experiments showed that the prevalence of infection will be decreased if: a) the number of sexual partners per person is minimized; b) injecting drug use is decreased; c) condom use is increased; d) the number of sexual contacts per partner is decreased; and, e) if injecting drug needles are not shared. As a case example below is given for part c) effect of condom use (Habtemariam et al., 2001).

![Figure 4: Projections of AIDS Cases in Blacks, Hispanics and Whites (under various levels of condom use).](image)

In this predictive model, the focus was on the use of condoms and its impact on reducing the incidence of HIV in sexually active adults in the USA. The model (Figure 4) shows that if active HIV/AIDS prevention and control interventions are not pursued, the HIV/AIDS incidence in the black population would increase from 60 per 100,000 in 1990 to 110 per 100,000 in 2020. In the Hispanic population, it would increase from 40 per 100,000 to 68 per 100,000. In the white population, it would increase from around 16 per 100,000 to 23 per 100,000. These predictions show that there are significant increases for all populations but much more devastating for the Black subpopulation. Condom use in 25% (the status
up to 1995 or so), 50% and 75% of sexually active adult populations was evaluated. The baseline of 25% was used in our model although the rates of condom use varied from low levels (5 to 10%) to 50% or more in surveys (CDC 1996, Douglas et al., 1997; Peipert et al., 1997). Figure 4 shows that increased condom use in 50% - 75% of the sexually active population, can decrease the rates to the pre 1991 levels, which were 47.9% for Blacks, 27.5% for Hispanics, and 11.6% for Whites. By the year 2020, the percentage reduction of AIDS will be 53% in Blacks, 49% in Hispanics and 43% in Whites. Our simulation only examined the proportion of condom use up to 75%, but if higher levels are evaluated, the rate of reduction will be higher and more consistent with the reported findings in the meta-analysis.

5.2 Micro level

The simulations of micro level model describe and quantitatively represent the cellular level dynamics between HIV virus and CD4+ lymphocytes (Figure 5 a) and b) (Habtemariam et al., 2002). Several assumptions are relied upon and some of the parameter estimates that undoubtedly will improve over time. Of greater interest also is the intracellular dynamics that represent the molecular level dynamics and interactions between HIV and the biochemical and RNA kinetics. There is a need for and to see the importance of using computational models to represent complex biomedical systems. Consequently these systems can best be studied cohesively and rationally using integrative systems dynamics modeling.

![Figure 5 (a): Computer simulation of the model with drug intervention. (b): Computer simulation of the model for the virus without drug intervention.](image)

5.3 Risk Assessments

After the risk pathway (scenario tree) culminating in the likelihood of unprotected sex leading to HIV infection is presented, parameters for each node were estimated using a review that included 17 English language published articles contributing tests of association for 23 variables associated with sexual risk behaviors. The majority of the studies 82% were conducted in the United States and about 18% of the studies were conducted in Europe. Participants were recruited from HIV outpatient clinics, sexually transmitted disease (STD) clinics, local or state health departments, or other community locations. A Monte Carlo Simulation with the software @Risk (Palisade Corporation) for excel was used for analysis. A total of 20,000 iterations were done for the simulation. The risk pathway consists of a sequence of specific events. For each node or event, a specific question related to the risk of unprotected sex is asked (Figure 3). The product of the answers to these questions determined the final risk related to unprotected sex leading to infection with HIV. There were 23 inputs and three outputs for the simulations. To see the
effect of various inputs on the output, likelihood of HIV infection in men through unprotected sex, a sensitivity analysis was performed using regression tornado graphs (Figure 6). Sensitivity analysis identifies what is “driving” the risk estimates and provides a way to show how the results of the likelihood of unprotected sex due to: 1) less knowledge about HIV/AIDS and beliefs, 2) emotional states and personality would be affected and how sensitive those results would be to changes in the values of specific input variable.

![Tornado graph showing the likelihood of unprotected sex due to: 1) less knowledge about HIV/AIDS and beliefs, 2) emotional states and personality](image)

**Figure 6:** Tornado graph showing the likelihood of unprotected sex due to: 1) less knowledge about HIV/AIDS and beliefs, 2) emotional states and personality

6 Discussion

Models for HIV/AIDS could be conceptual, in-vivo or in-vitro, systems analysis, mathematical, or computational just to name a few. The knowledgebase developed using the EPOA methodology provides a well-organized structured source of information, which is used in the variable and parameter estimations and analysis (biological, mathematical, statistical and computer simulations) that are crucial in epidemio-
logic modeling of HIV/AIDS. These models are cost effective in that they are less time consuming, easily manipulated for different scenarios and not dangerous in comparison to human experimentations especially invasive ones for such a disease like HIV/AIDS. This ethical approach has enabled great strides to be achieved in the area of HIV/AIDS research which would have otherwise taken a long time especially due to the limited availability of suitable animal models. A model is evaluated first and foremost by its consistency to empirical data; any model inconsistent with reproducible observations must be modified or rejected. However, a fit to empirical data alone is not sufficient for a model to be accepted as valid. Other factors important in evaluating a model include; Ability to explain past observations; Ability to predict future observations; Cost of use, especially in combination with other models; Refutability, enabling estimation of the degree of confidence in the model and; Simplicity. Epidemiologic Problem Oriented Approach has become a crucial tool in the development of HIV/AIDS models that are important for decision making in public health as it relates to prevention, control and treatment strategies. As more knowledge becomes available, the knowledge bases can be updated and used to improve on the models thus enhancing better decisions to be considered as pertains to this HIV/AIDS pandemic. Furthermore, the results from this study may be extrapolated to assist in public health policy planning, decision-making, and in education to prevent and/or reduce the HIV/AIDS pandemic burden. In addition, the integration of the macro and micro modeling will help to see the effect of mitigation measures from individual level to the population level and ultimately to the national and international levels.

7 Conclusion

The rationale behind relying upon computational epidemiology and risk analysis modeling is that “Epidemiology has its tentacles in literally all disciplines, such as in mathematics and statistics, molecular epidemiology, socioepidemiology, computational epidemiology etc. In its simplicity, it focuses upon the study of population dynamics, whatever the population may be (biotic, i.e., living or biological entities, as well as abiotic or non-living population entities).” Key to its breadth of application is epidemiology’s extension into computational and risk analysis models, which can be used to examine and communicate a diversity of epidemiological concerns and scenarios. Excitingly, the role that computational and risk analysis models could play in the advancement of the theoretical understanding of disease processes and the identification of specific intervention strategies and scenarios holds the potential to impact and save human life.

Computational epidemiologic models provide insights in examining a variety of computational experimentations. Complimented with computational tools under the frame work of sound epidemiology, risk analysis can help the current and future problems in tackling the problem of HIV/AIDS by considering multiple scenarios and sensitivity analyses. Such experimentations can be of profound help in evaluating scientific questions related to effective strategies in HIV drug therapy interventions and risky behaviors.

Acknowledgement

This work is supported by a Research Centers in Minority Institutions (RCMI) Award, 2G12RR03059-16, from the National Center for Research Resources.
References


