Statistical Modeling of the AIDS Virus

**New Mexico Adventures In
Supercomputing Challenge
Final Report
April 4, 2005**

**Team Number 080
Silver High School**

**Team Members:
Adam Cummings, Senior
David Saxton, Senior
Cyrus Marcum, Junior**

**Teacher:
Mrs. Peggy Larisch**

**Team Mentor:
Berry F. Estes**

Acknowledgements

The authors wish to acknowledge the following individuals for the guidance provided in the selection of a suitable topic, development of the computer code, adaptation to the computer based analyses, review and understanding of the results, and assistance provided in the overall preparation of this report:

- Mrs. Peggy Larisch – Teacher, Silver High School, Advanced Computer Studies
- Dr. Thomas Gruska – Professor, Western New Mexico University, Mathematics and Statistics
- Mr. Berry F. Estes – Mentor, Preparation of technical documents and presentation of computer projects.
- Charlie Voss – Principal, Decision and Risk Analysis - System Simulation Group, Golder Associates Inc. Creation of an equation that accurately models the AIDS virus.

Table of Contents

Locations of Sections

Locations of Figures

E.0 Executive Summary

The topic that was chosen for this computer-modeling project is the statistical forecasting of the AIDS/HIV virus. The topic was chosen due to the fact that the AIDS/HIV epidemic is currently a very large threat facing the world today. Cases of this disease are found all over the world. This computer program is designed to model a population that is statistically representative of the population of the United States. The population's starting size is the same as the United States', as well as contact rates (the chance that an infected individual will infect another), and age distribution.

After inputting these beginning variables, the program runs them through a modified version of the Susceptible-Infected (SI) model. The SI model employs an equation to predict the spread of the infected (people with the disease) and susceptible (people able to become infected) portions of a population, and the relation between them. This equation, however, was too simple to accurately model the spread of a disease as complex as AIDS/HIV. Due to it's inability to model many non-static variables (changing population, changing death rate, the randomness of the real world) the model was extensively modified in this computer program. The modified computer program has a birth and a death rate, and changes the contact rate depending upon the percentage of the population infected with the disease. The program also keeps track of the ages of the people in the population, and changes their contact rates based on which age group each member of the population is in. Another portion of the population, the unsusceptible, are also taken into account in the program.

After observing the results and predictions output by the program it was noticed that the population modeled did not die out due to the disease, but was affected in a detrimental manner. The program was also able to show that even in the event of a cure or vaccine being found for the disease, the population would still be negatively affected for several years afterwards. The effectiveness of the cure developed is not as important as when the cure is discovered. Also using an immunization to treat the disease is much more effective than using a cure.

1.0 Introduction

1.1   Purpose

The purpose of this mathematical modeling project is to develop a model based on facts known about the spread of the Acquired Immune Deficiency Syndrome (AIDS) virus, and to predict its spread throughout a population that is statistically representative of the average population. This subject was selected due to the fact that AIDS is adversely impacting populations throughout the world. A number of dangerous viruses have been eradicated (i.e. smallpox) in recent history via improved medical technology and treatment; however no cure has yet been discovered for AIDS. Because there is no positive treatment for the cure of AIDS, additional in-depth research will be required in both developed and undeveloped countries. Mathematical modeling can be applied to evaluating multiple aspects involved in exposing a population to the AIDS virus and predicting the spread of AIDS within this population. These mathematical models can be readily converted into a computer project that evaluates effects of the disease and potential spread of the disease.

The spread of this disease was examined by creating a mathematical model, applying the model to numbers obtained from historical statistical data, and then running the model through a computer program to mimic the behavior of the disease[7]. The equation that this program is based on is known as the SI (Susceptible Infected) model which is a mathematical formula that is commonly used to calculate the spread of infectious diseases. The model takes into account the numbers of people with a disease (infected), and people who can still get the disease (susceptible). Using this model the interaction between these two groups can be evaluated by monitoring the infection of those who are susceptible to the disease. Using the SIR (Susceptible-Infected-Recovered) model was considered but HIV/AIDS currently has a recovery rate of zero. The SI model is therefore superior to the SIR model in this case because it more closely models the actual behavior of AIDS and it has all of the basic components for modeling AIDS.

1.2   Scope

The scope of this program is limited to developing a mathematical model based on historical data that describes the current health problems associated with AIDS; and

creating a set of computer based methods that predict the future behavior of this infectious disease. The program provides a basis for examining the effects of the AIDS virus on a given population and examines the spread of AIDS through a specified population representative of the population of the United States of America. The mathematical models and computer program which were developed by this team to analyze the impact of the AIDS virus in a selected population can also be used, by changing specific input parameters, to predict the behavior of the AIDS virus in other possible situations such as: discovery of a vaccine, a more infectious strain of the virus, or a mutation that causes a part of the population to be immune to AIDS.

1.3   Computer Program

This project makes use of the C++ programming language. Before being written in C++, the original SI model was tested in Excel to insure its accuracy and applicability in this situation. C++ is applicable to this situation because of its mathematical capabilities and its simplicity. The computer program was developed by the team and is based on accepted analytical models used to simulate the spread of the AIDS virus. The original SI model used in the program was obtained from a mathematical modeling textbook. This model has been modified as appropriate to accurately represent the threat to a population posed by AIDS for the specified input conditions defined in the document. These modifications, the testing of them, and their final use in this program were accomplished solely by team 80 with input and guidance from the acknowledged sources

.

2.0   Project Proposal

2.1   Description of Project

The purpose of this computer-modeling project is to develop, using the C++ programming language, a model that accurately predicts the spread of the AIDS/HIV epidemic. The computer program accomplishes this by taking in a number of real-world statistics and running them through a modified version of a mathematical equation known as the SI model. The SI model is a mathematical equation that is used to describe the dynamic effects of a disease and track the susceptible and infected members of a population. This mathematical equation was extensively modified for use in the computer program, due to the fact that the original equation did not take into account certain dynamic factors such as population growth and decline, nor did it take into account different risk factors for different age groups. The computer program takes these variables into account, as well as a population's reaction to a dangerous disease. Using this diverse version of the SI model allows the program to accurately predict the future behavior of the disease by calculating the change in the variables as a function of time.

The topic for this computer project was selected because AIDS/HIV is a very widespread and dangerous epidemic. Cases of this disease are found all over the world, including in every state in the United States, which is the reason that the United States in particular was selected to be modeled by the program. The vast area affected by the disease, as well as the fact that AIDS/HIV is as yet incurable makes it ideal to base a computer modeling program on.

3.0   Analytical Methodology

3.1   Mathematical Bases

     The Susceptible-Infected(SI) model, as shown below, was recommended for use in the application of this computer program by a team mentor (Charlie Voss, Golder Inc.). Mr. Voss is a specialist and works with modeling complex systems both natural and engineered.

$$IR = (ciS)(I/N)$$

The basic SI model consists of five fundamental parameters; these are:

- contact rate(c)
- infectivity(i)
- susceptible population(S)
- infected population(I)
- and total population(N)

These variables are used to calculate the newly infected members of the population (IR) during each of the time steps. As the equation is run the susceptible population goes down and the infected population goes up. At first, the number of new infected increases each time; because a higher percentage of the population is infected more susceptible individuals come in contact with infected individuals. Eventually the susceptible population becomes small enough such that the new infected group within the population in each of the iterations cannot remain constant and begins to decline as shown in Figure 1 [1].
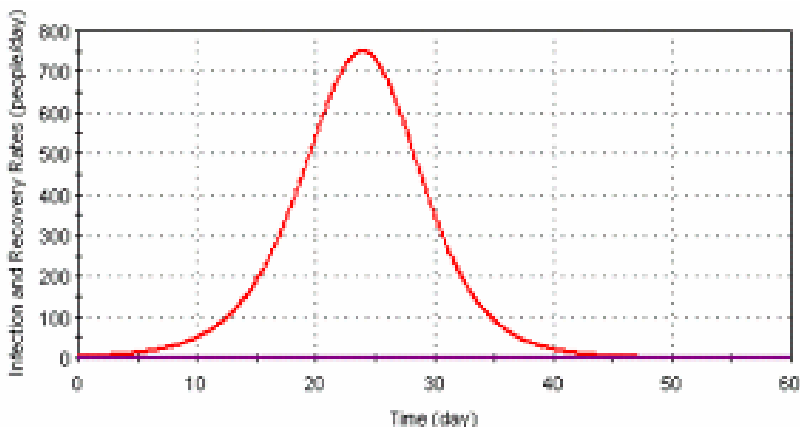

Figure 1 Time History of Infection Rate.

This rollover effect shown above is a trend that is repeated in most all epidemics. It is caused by the interaction of the susceptible and infected populations as the size of their populations increases and decreases throughout the spread of the disease. Although this trend can be found in epidemics in real world situations the curve is effected by many more variables i.e. a population with immunity to the disease, certain population members who are less likely to contract the disease. This curve is most applicable to epidemics that spread rapidly through a population, kill many of the population's members, and finally leave mostly immune members of the population.

The HIV/AIDS epidemic is unique because it can take years before an infected individual shows symptoms of the virus. This means instead of modeling the epidemic for days it must be modeled over many years. Since it needs such an extended modeling time other variables must be taken into account. These variables include birth and death rates for all populations (different age groups, infected, uninfected) in the simulation. Another problem large time steps created in the creation of the program is that people, after being born have to grow old and eventually die. To solve this problem different death rates were applied to each population group to reflect not only a population member's ability to die but also taking into account that individual's age. Also, there is approximately a 1% chance a person is born with immunity to HIV; therefore, in addition to adding age groups, a new population group called "unsusceptible" was added to the program.

In order to make the model more realistic during each time step, each person has a chance (determined using input variables) to have a child, die or become infected. Then each child born has a chance to be born susceptible, unsusceptible, or infected (if the parent is infected). When the program makes a decision it goes through this process: If a randomly generated number is less than the chance defined by the input variable then the event in question (birth, death, infection) takes place if the random number is greater than the chance the event in question does not take place. For example if the birthrate is 17 per 1000 and the random number (between 1-1000) is 15 a baby is born.

A few more examples of events that take place in the program are:

- (Sbr*2)*(uS/N*10)*100) The chance in 100000 a baby is born unsusceptible(uS). This finds the percentage of the population that is uS so if there are more uS in the population more uS babies are born.
- (IR0/S)*100000) The chance in 100000 this individual becomes infected. IR0 is calculated by the SI model.

See Appendix 4 for a complete list of the equations used in the program.

3.2   Computer Applications

The program requires input from a file called "input.txt" (see Appendix 2).   This file includes all of the parameters needed to run the program.   It also allows you to control what is output from the program. All of the output from the program is put into a text file called "output.txt". This is done so the results can be imported into Excel to be graphed/analyzed.   This program was developed entirely by Team 80. Most of the output was used for debugging.

For a complete explanation of the process of the program see Figure 2 for a flowchart of the program and Appendix 2 for the source code.
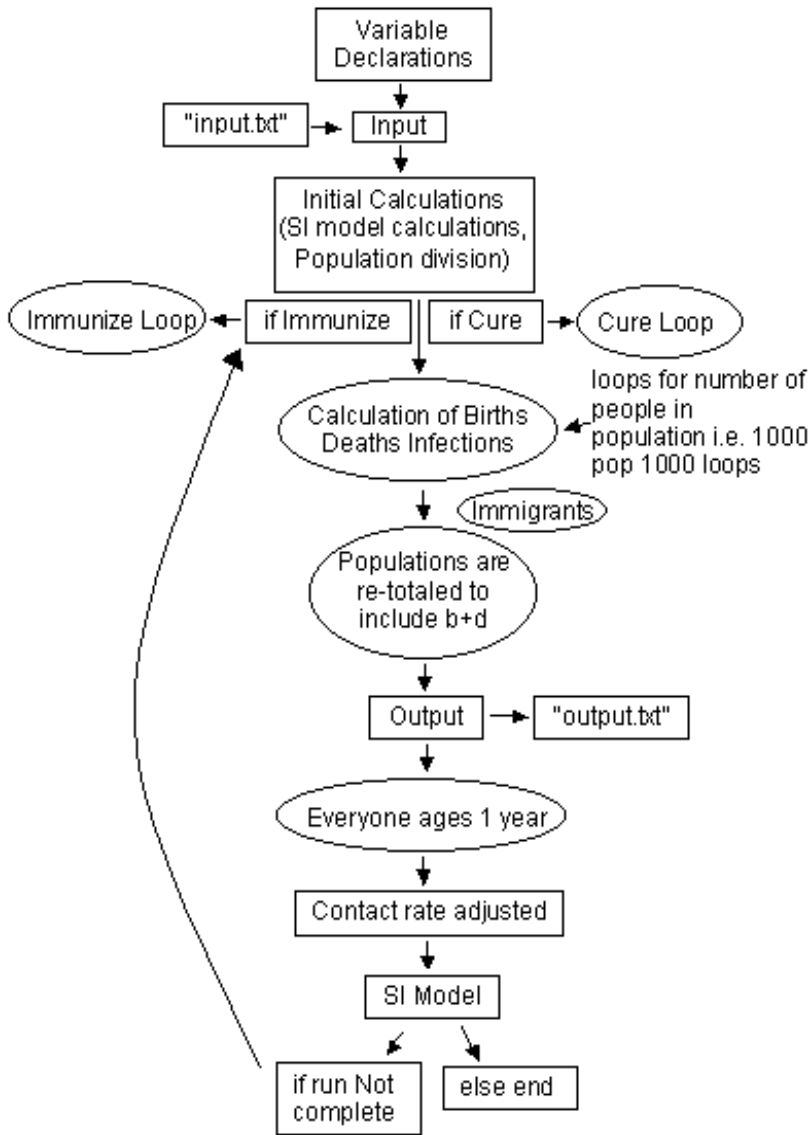
Figure 2

Many assumptions had to be made for the project.   Most of them involve people's activities remaining consistent with how they are today.   The following is a list of the assumptions that were made:

- Birth, death and immigration rates remain similar to what they are currently.

- There are no major population changes (due to war or a different disease).

- The virus does not adapt to changes in the populations.

- All members of specific age groups participate in similar activies.

- Immigration laws remain similar to what they are today.

4.0   Results

4.1   Computer Calculations

The program was run many times so only a summary of one of the runs is given in this section. This is an example of the control run done for comparisons. For this trial there is no cure, no immunization and the populations response is set to 5. The initial parameters are as follows (initial populations are 1/10 of the United States' population in 1990):

- Susceptible Population: 24,815,492
- Infected Population: 3,523
- Susceptible Birth Rate (per 1000): 17
- Susceptible Death Rate (per 1000): 6
- Infected Birth Rate (per 1000): 2
- Infected Death Rate (per 1000): 150

The populations reached a maximum of:

- Susceptible Population: 55,412,044
- Infected Population: 700,972

The final totals for the populations after 150 years are:

- Susceptible Population: 55,412,044
- Infected Population: 387,789

Annual infection rates:

- Average New Infected: 30,788
- Highest New Infected in One Year: 67,583
- Lowest New Infected in One Year:

Initially a different method for modeling AIDS was attempted using statistical curve fitting [2]. It was after this that the SI model was chosen to be the base model. .The SI model is very simple and had to be modified extensively. At this point decisions had to be made regarding how the SI model would be customized. The most difficult decision was which demographics should be included and how to include them. Race was not included as a demographic because the race of each child would have to be determined based on the child's parents. This problem is beyond the scope of this

project. Age was included in the model so people of different age groups partake in different activities (i.e. newborns do not have children, old people do not participate in dangerous activities). Another problem encountered is how to implement cures/immunizations. In the end the decision was made to treat cures and immunizations as percentages, which can be changes through the input file.

4.2 Graphs, Tables and Figures

Figure 3 shows the infected deaths in relation to different cure rates and cure times. This shows the importance of when the cure is discovered and how effectively the cure can be manufactured. The dark blue control group shows what would happen if no cure were discovered. The longer it takes for a cure to be discovered the more people die. Eventually a cure being discovered does not have a dramatic effect on the deaths in the population because such a large portion of the population is infected. Similarly, if a better (cheaper, more easily manufactured) cure is discovered fewer people die. Year of discovery has a more beneficial effect on the deaths caused by infection.
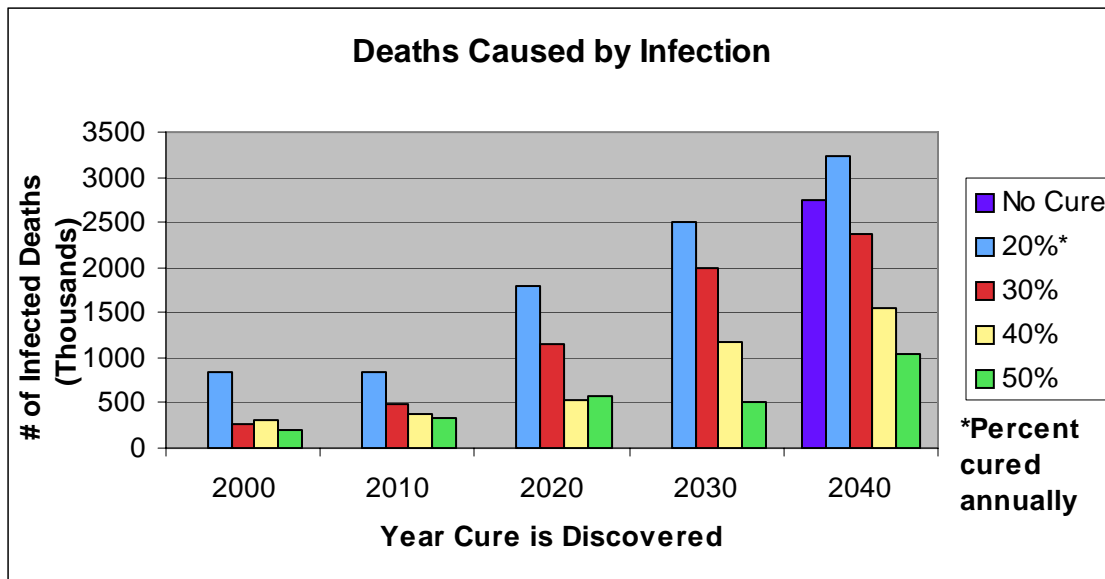
Figure 3

Figure 4 shows the effects of a cure that can be manufactured for only 20% of the population annually. Notice a cure discovered after 40 years does not change the population significantly from the control group. This shows the importance

discovering a cure early. As the cure is discovered later and later it has a smaller and smaller effect on the end results.
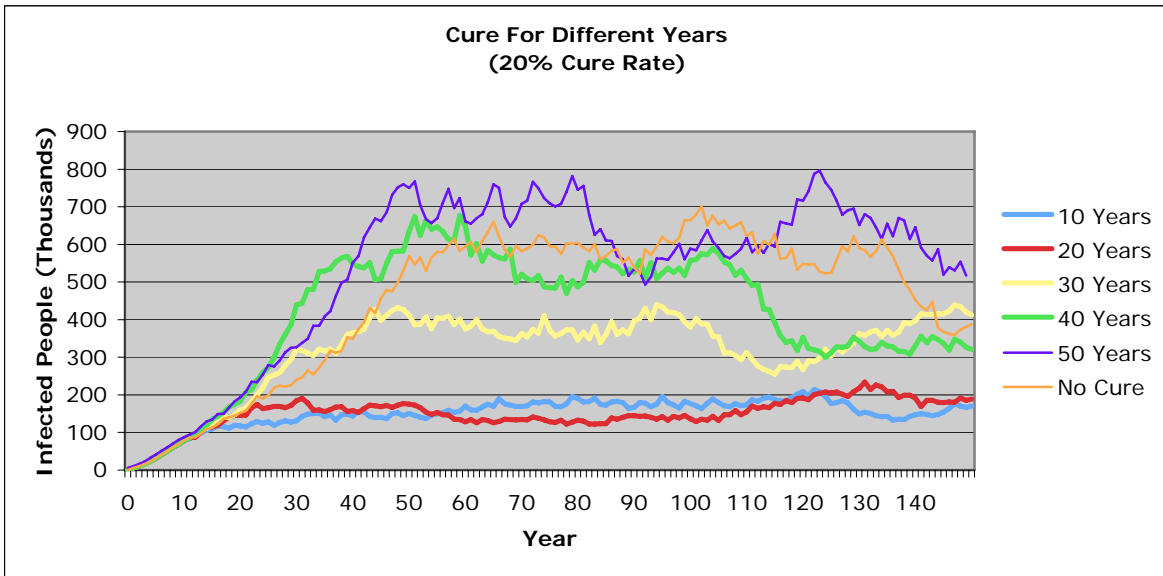

Figure 4

Figure 5 shows the effects of a cure that can be manufactured for 50% of the infected population annually. A cure rate of 50% has a dramatic effect on the infected population regardless of the year of discovery.
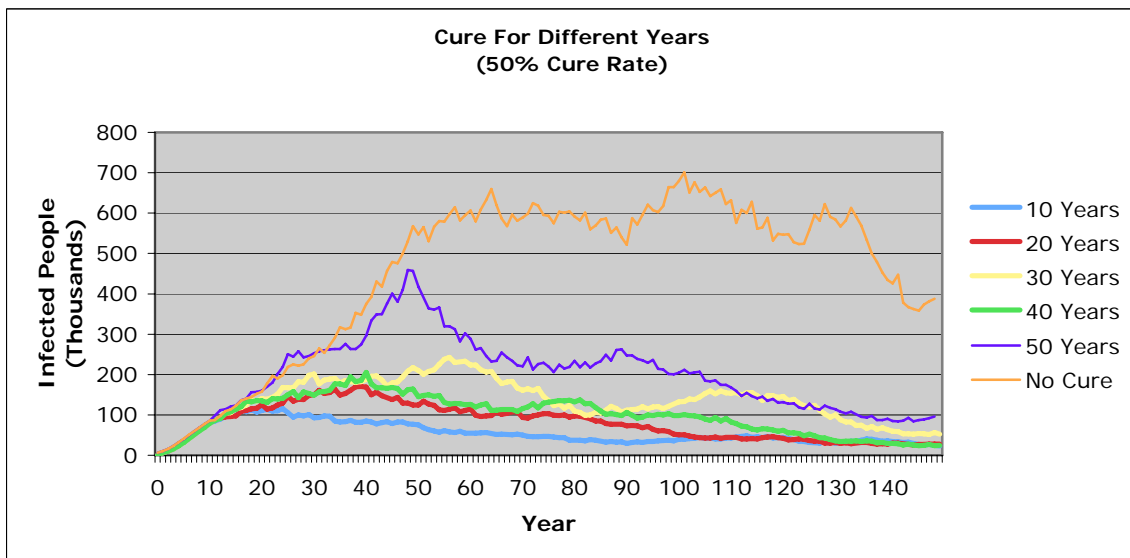

Figure 5

Figure 6 shows the effectiveness of an immunization. The figure shows different immunization rates for an immunization discovered after 50 years (2040). Nearly all of the infected die off, whereas if a cure is used the infected population lingers for much longer.

Immunization Rates After 50 Years
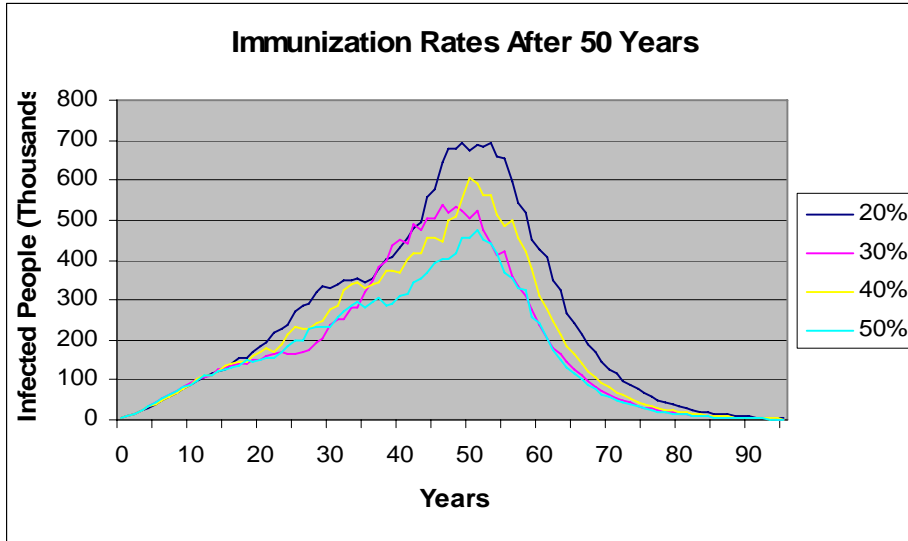
Figure 6

4.3   Other Data and Comparisons:

The statistical data plugged into this program was collected from the CDC (Centers for Disease Control and Prevention), which is an organization that has been working to provide up to date information so that people can more effectively combat AIDS/HIV. The CDC has been putting out an AIDS/HIV surveillance report for the past twenty years, ever since the virus was first found to be a major threat to the population. Unfortunately, as the means of diagnosing AIDS was very inexact in the early years that the disease was around, some of the earlier data in these reports is inaccurate. The HIV/AIDS Surveillance Report contains many tables, as well as graphical representations of U.S. AIDS and HIV case reports, including data by state, metropolitan statistical area, mode of exposure to HIV, sex, race/ethnicity, age group, vital status, and case definition category. The report is usually issued in the early fall[3].

Other statistical data that was used in this program was derived from the United States Census Bureau. This organization makes record of the population of the United States, and this data was used, in conjunction with the data taken from the CDC, to

derive many of the variables used by this program, such as the number of people that a diagnosed AIDS patient will come into risky contact with in a year [4]. Other data that was used by this program, such as the one to two percent chance that a person can be born immune to AIDS, was collected from the various online and printed news articles that have to do with the subject [5].

5.0 Conclusions

5.1   Mathematical Models

The SI model was chosen to be used in this program because it is used to model an infectious disease without any recovery rate. However, the original SI model was not completely applicable to the problem faced by this program due to the fact that the original equation only took into account a static population (one that does not change its total population over time) and it is obvious that real-world populations never remain static for any length of time. Due to the fact that the SI model lacked the complexity to model population change or population response to a dangerous disease it was modified extensively by Team 80 to suit the purposes of this program.   These modifications (discussed further in Appendix 4) caused it to be more applicable to the situation, as it allowed the program to use a changing contact rate, calculate a population's response to a disease, and also take into account age, adjusting the contact rate based on that as well [1].

5.2   Computer Program

The computer program that was the final product of the work of Team 80 more than adequately runs the modified SI model.   Along with running the model it also implements the use of random numbers to make its operations more similar to reality. The program also simulates a population's response to a disease that is perceived as a threat.   Using these features this statistically based program operates in a very realistic way. The program in this study produces results that are similar to observed AIDS virus activity in a selected country (United States). The population selected for this study was not observed to die off; however AIDS is having a detrimental effect on the population by causing many people to have shortened lives and a declining birth rate in the general population.

5.3   Results

The computer program developed by Team 80 was able to produce results that were very similar to real-world data. The program did this through use of an extensively modified version of the mathematical equation known as the SI model. The SI model,

when unmodified, takes the number of infected (I) and susceptible (S) people in a population, as well as the contact rate and infectivity of the disease, and uses these number to generate the number of new infected for the next year[1]. This program also takes into account the birth and death rates of the population being modeled, the age division, and also the chance that a person is immune to the disease (either through a cure/immunization or a genetic immunity to the disease). The program also uses randomly generated numbers to determine the chance that a person will die, be born, or be infected, as well as for many other of this program's functions. This allows for varied, yet statistically accurate results each time this program is run. Another way the program achieves accurate results is by allowing for a population's possible response to an epidemic. This assumes that once a critical percentage of the population becomes infected, it will trigger more education and preparation against the disease, thus lowering the contact rate.

The results of the repeated tests of the program indicate that AIDS/HIV will have a detrimental effect on any population it is introduced to, causing much loss of life. The results showed that while the disease is highly unlikely to eliminate the entire population, those who are susceptible to it are affected negatively. The unsusceptible population was mostly unaffected, and grew relatively smoothly throughout the tests. The results also show what the effect of a cure/immunization will have on the spread of the virus, based on how effectively the cure is distributed and at what point it is developed. These results show that the later a cure is developed, the less influence it will have on the disease, showing the importance of developing an early cure. The infected population is directly dependent upon the susceptible population so the development of an immunization would be extremely beneficial. An immunization stops the problem at its source; as a result; it is significantly more beneficial to use an immunization than a cure. This project shows the importance of taking preparations against the AIDS epidemic in a timely and efficient manner in order to prevent the greater loss of human life.

5.4   Word Processor Program

Several programs were used in the formation of the program written by this team. Dev C++ provided an adequate compiler and bug checker for the C++ programming language.   Microsoft Word and Open Office were used for word processing and document formatting.   Microsoft paint was used multiple times for the preparation of graphics used in the report.   Notepad was used for reading and editing input and output files in this program.   Microsoft Excel created the charts shown in this report and was used to test the equations used in the program.   Scientific Notebook was used to type out the formulas used in a graphically presentable way.   Finally three separate G-mail accounts provided by www.google.com were used by the members of this group for correspondence with each other and data storage.

6.1   Future Work

For future work it would be beneficial to add to the program more demographics such as sex, race, location, lifestyle, etc. Each of these would effect how often a person takes part in an activity which could possibly allow him/her to become infected, and would thus allow the program to more accurately predict the spread of the disease. The SI model could become nearly obsolete due to being replaced by a much more complex series of equations. Other variables used by the program could also be allowed to change as the program runs, such as birth rate, death rate, or infection chance, depending on the program's generated statistics. This fully completed program would allow people to be more prepared for the possible future effects of the disease by giving them an accurate model that shows how the disease will affect the country as a whole.

References

1. John Sterman, John D. Sterman "Business Dynamics: Systems Thinking and Modeling for a Complex World", McGraw-Hill/Irwin, 2000.

2. Douglas Downing, Jeff Clark, and Jeffrey Clark, "Forgotten Statistics: A Self-Teaching Refresher Course", Barron's Educational Series, 1996.

3. Center For Disease Control and Prevention. <u>HIV/AIDS Surveillance Report</u>, 1-14 1987-2002.

4. "Monthly National Population Estimates", http://www.census.gov/popest/national/NA-EST2004-01.html, January 20, 2005

5. Randy Dotinga, "Genetic HIV Resistance Deciphered", http://www.wired.com/news/medtech/0,1286,66198,00.html, January 20, 2005

6. "WHO HIV/AIDS Programme", http://www.who.int/hiv/en/, December 9, 2004.

7. "AIDS", http://en.wikipedia.org/wiki/AIDS, December 9, 2004

8. "AIDS", http://www.free-definition.com/AIDS.html, December 9, 2004.

9. Mirko D. Grmek, "History of AIDS: Emergence and Origin of a Modern Pandemic", Princeton University Press, 1990

Appendix 1 – Understanding The Model and Program

Appendix 1 – Understanding The Model and Program

1.1 Breakdown of Population Groups

Total Population (**N**)
├── Infected (**I**)
│   ├── Diagnosed (**D**)
│   └── Undiagnosed (**unD**)
└── Uninfected (**unI**)
    ├── Susceptible (**S**)
    └── Unsusceptible (**unS**)
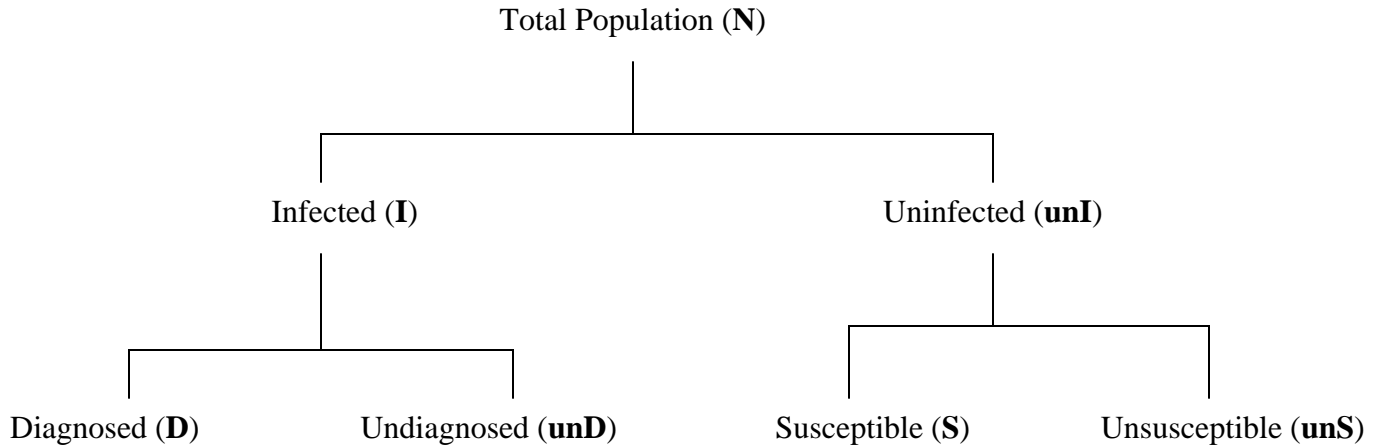
Each of the four main groups that compose the total population have different qualities that separate them from the other three population groups. These qualities include like likelihood of dieing, ability and likelihood of having children, chance of death, and ability to infect or be infected.

Diagnosed Population: higher death rate and lower birth rate than uninfected, can infect susceptible individuals
Undiagnosed Population: birth and death rate are the same as uninfected, can infect susceptible individuals
Susceptible Population: lower death rate and higher birth rate than infected, can be infected
Unsusceptible Population: lower death rate and higher birth rate than infected, cannot be infected

Each of the four major groups that make up the population are divided into age groups from 0 to 99. These age groups determine whether the individual can be infected, infect others, have children, or his/her likelihood of dieing.

0-14 Year Old Infected: cannot have children, have twice the normal infected death rate, cannot infect susceptible individuals
15-50 Year Old Infected: can have children, normal infected death rate, can infect susceptible individuals
50-99 Year Old Infected: cannot have children, twice infected death rate, cannot infect susceptible individuals

0-14 Year Old Uninfected: cannot have children, half the normal uninfected death rate, cannot be infected
15-50 Year Old Uninfected: can have children, normal uninfected death rate, can be infected if susceptible
50-99 Year Old Uninfected: cannot have children, twice the uninfected death rate, cannot be infected

1.2   Program Flow Chart

Variable Declarations

"input.txt" ← Input

Initial Calculations …. (Explain SI Model, Population division into age)

Immunize Loop ← if Immunize   if Cure → Cure Loop

Calculation of Births Deaths Infections ← loops for number of people in population i.e. 1000 pop 1000 loops

Immigrants

Populations are re-totaled to include b+d

Output → "output.txt"

Everyone ages 1 year

Contact rate adjusted

SI Model

if run Not complete   else end

1.3   Variables Used In the Program

Below is shown a list of the variables used in the program that are necessary to understanding the model.

Populations
N – Total Population
uS – Unsusceptible Population
S – Susceptible Population
I – Total Infected Population
dI – Diagnosed Infected Population
udI – Undiagnosed Infected Population

Births
Ib – Total Infected Births
dIb – Diagnosed Infected Births
udIb – Undiagnosed Infected Births
uSb – Unsusceptible Births
Sb – Susceptible Births

Deaths
Id – Infected Deaths
Sd – Susceptible Deaths
uSd – Unsusceptible Deaths
dId – Diagnosed Infected Deaths
udId – Undiagnosed Infected Deaths

Infections
IR0 – New infected for that year

Various other Variables
Sbr – Susceptible Birth Rate
Sdr – Susceptible Death Rate

Idr – Infected Death Rate
Ibr – Infected Birth Rate
C – Contact Rate/Infectivity
firstYearInfected – People infected during the current time step of the program
popResponse – Changes how the population responds to the infection
percentCured – Percent of the infected population cured annually
percentImmunized – Percent of the susceptible population immunized annually
totalInfectedDeaths – Keeps a total of all deaths caused by infection
startPopulation – Initial uninfected population
infectedAge – An array that holds population totals for each age.
susceptibleAge – An array that holds population totals for each age.
unSusceptibleAge – An array that holds population totals for each age.
yearsSinceInfection – An array that keeps track of how long a person has been infected.

1.4   Unmodified SI Model Illustration

The graphic shown below illustrates the many factors that are dealt with in the process of modeling the two populations in the original SI Model [1].
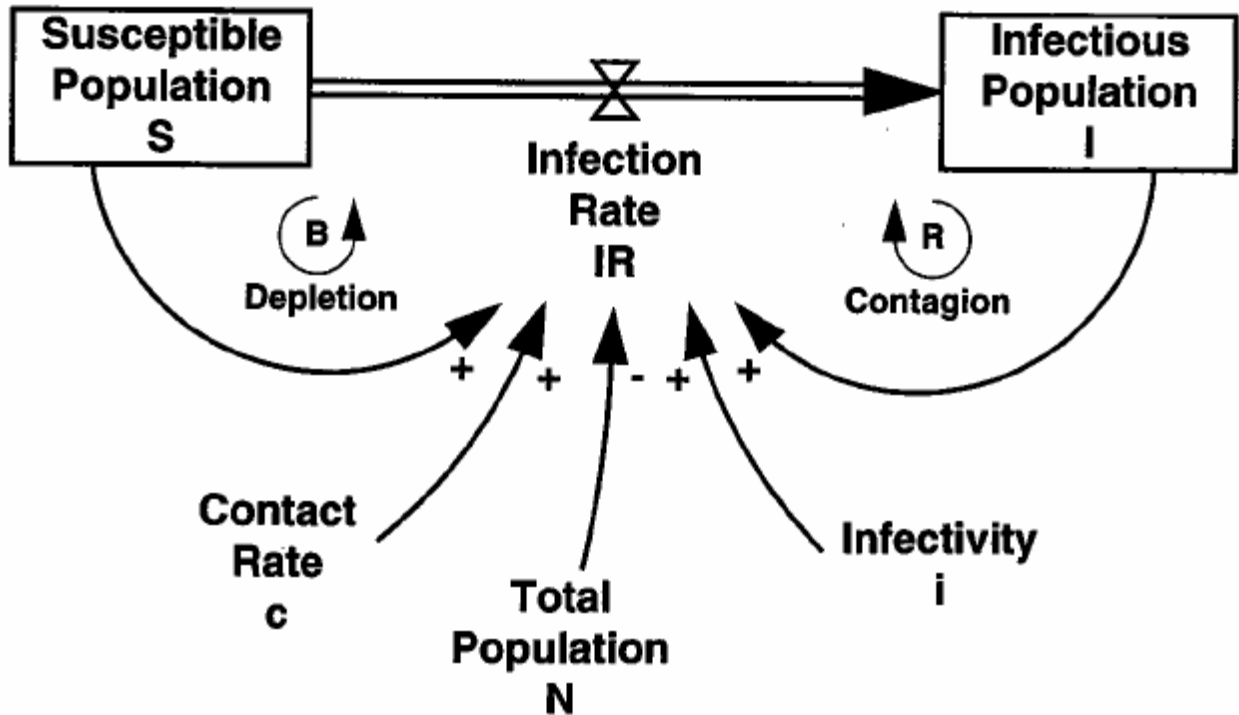
The SI Model

$$IR = (ciS)(\tfrac{I}{N}) \quad N = I + S$$

The original SI Model calculated the number of newly infected people(**IR**) each year. It took into account the size of the total population(**N**), the infected population(**I**), the susceptible population(**S**), the contact rate(**c**) of the uninfected with the infected, and the infectivity(**i**) of AIDS[1].

1.5 Modified SI Model Illustration

The graphic shown below illustrates the many factors that are dealt with in the process of modeling the four populations in the modified SI Model used in this program.
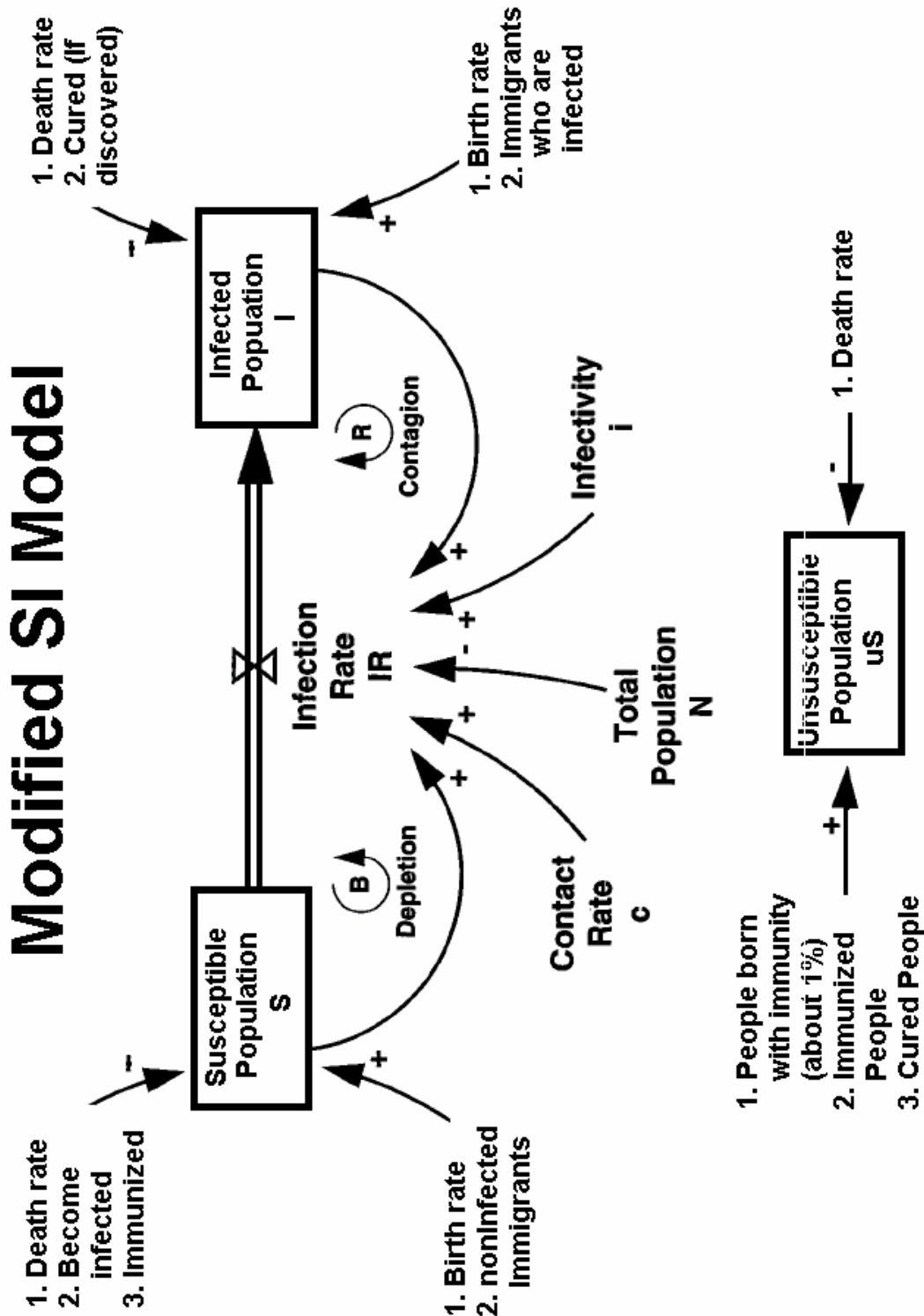
Appendix 2 - Code and Computations

Appendix 2 – Code and Computations

```
/**********************************************************************/
/* Team 080:  Adam Cummings, David Saxton, Cyrus Marcum          */
/* Adventures in SuperComputing 2004-2005                   */
/* Programmed by Adam Cummings                      */
/* Version Date: 3-17-2005                     */
/* This program uses models the spread of the AIDS virus         */
/* It uses an extensively modified version of the           */
/*     Si model to accomplish this               */
/* Input is required from the file "input.txt"          */
/**********************************************************************/

#include <iostream.h>
#include <stdlib.h>
#include <fstream>
#include <iomanip>
#include <time.h>
using namespace std;

//Function ProtoTypes:
    void Input();  //Handles all input
    bool ShouldThisOutput(ifstream& fin); //determines what should be output
    void Output(); //handles all output
    void Cure();   //called if cure is started
    void Immunization();  //called if immunization is started
    void LowerContact(int); //determines contact rate for each time step
//Type Definations
    const int MAX_AGE=100;
    typedef int ageCounter[MAX_AGE]; //0-99 each an age group
    struct unDiagType    //struct needed because c++ does not allow agregate
    {              //operations of arrays
     ageCounter unDiagnosedAge;
    };

//Variable Definations:
    int years, yearsUntilCure, yearsUntilImmunization;
    double N=0, S=0, I=0;
    double firstYearInfected=0,IR0=0, Ib=0, Id=0;
    double Sb=0, Sd=0, uSb, uSd, uS;
    double dI=0, udI=0;
    double dIb=0, dId=0, udIb=0, udId=0;
    double Sbr=0, Sdr=0, Idr=0, Ibr=0,c=0, i=0;
    int popResponse;
    int percentCured, percentImmunized;
    int randNum;
    long totalInfectedDeaths=0, startPopulation;
    bool outputN=true, outputI=true, outputS=true, outputuS=true;
    bool outputIR0=true, outputudI=true, outputdI=true, outputSb=true;
    bool outputSd=true, outputudIb=true, outputudId=true;
    bool outputdIb=true, outputdId=true, outputuSb=true, outputuSd=true;
    ofstream fout("output.txt");
    ageCounter infectedAge, susceptibleAge, unSusceptibleAge;
    unDiagType yearsSinceInfection[10];


/*
```

```cpp
	N-total population
	S-susceptible pop; uS-unsusceptible(people with born immunity)
	I-infected pop; dI-diagnosed Infected;  udI-undiagnosed infected

	b -births  br-birthrate
	d-deaths   dr-deathrate
	IR0-infected this year(undiagnosed)

*/


int main()
{

	Input();
	startPopulation=N;
	srand(time(NULL));
	fout << setprecision(3) << "Total_Pop: " << int(N) <<"\tinf_Pop: "
		<< int(I) << "\tContact_Rate: " << c << "\tinfectivity: " << i
		<< "\tPop's_Response: " << popResponse  << "\tCured%: " << percentCured/10
		<< "\tImmunized%: " << percentImmunized/10
		<< "\nSus_Brth_Rate: " << Sbr << "\tSus_Dth_Rate: " << Sdr
		<< "\tInf_Brth_Rate: " << Ibr << "\tInf_Dth_Rate: " << Idr
		<< "\tYears_until_Cure: " << yearsUntilCure
		<< "\tYears_until_Immunization: " << yearsUntilImmunization
		<< endl << endl;

	fout << setw(5) << "Year";             //Outputs Column Headings
	if(outputN)
	  fout << setw(20) << "Totl_Pop";        //Outputs Column Headings
	if(outputS)
	  fout << setw(20) << "Sus_Pop";         //Outputs Column Headings
	if(outputuS)
	  fout << setw(20) << "unSusceptible";   //Outputs Column Headings
	if(outputI)
	  fout << setw(20) << "Inf_Pop";         //Outputs Column Headings
	if(outputIR0)
	  fout << setw(20) << "New_Inf";         //Outputs Column Headings
	if(outputudI)
	  fout << setw(20) << "unDiagnosed_Inf";    //Outputs Column Headings
	if(outputdI)
	  fout << setw(20) << "Diagnosed_Inf"; //Outputs Column Headings
	if(outputSb)
	  fout << setw(20) << "Sus_Birth";       //Outputs Column Headings
	if(outputSd)
	  fout << setw(20) << "Sus_Death";       //Outputs Column Headings
	if(outputudIb)
	  fout << setw(20) << "unDiagnosed_Bth";   //Outputs Column Headings
	if(outputudId)
	  fout << setw(20) << "unDiagnosed_Dth";   //Outputs Column Headings
	if(outputdIb)
	  fout << setw(20) << "Diagnosed_Bth"; //Outputs Column Headings
	if(outputdId)
	  fout << setw(20) << "Diagnosed_Dth"; //Outputs Column Headings
	if(outputuSb)
	  fout << setw(20) << "unSus_Brth";       //Outputs Column Headings
```

```cpp
    if(outputuSd)
      fout << setw(20) << "unSus_Dth";
    fout << setw(20) << "contact_Rate";        //Outputs Column Headings
    fout << endl;


    S = N - I;        //initial susceptible calculation
    dI = I;                //all starting infected are diagnosed
    IR0 = (c*i*S)*(I/N);  //Initial SI
    uS=S*.01;
    S = S*.99;
    for(int j = 0;j<MAX_AGE;j++)  //Divides initial population into age groups
    {
      susceptibleAge[j] = S*((-.0882*j+9.19)/488);
      infectedAge[j] = I*.01;
      unSusceptibleAge[j] = uS*((-.0882*j+9.19)/488);
    }

    for(int counter = 0;counter<years;counter++) //Main Loop
    {

      if (counter>=yearsUntilImmunization)  //starts immunizations
        Immunization();

      if (counter>=yearsUntilCure)        //Cure
        Cure();

      dId=0;
      dIb=0;
      udIb=0;
      udId=0;
      uSb=0;
      uSd=0;
      Sb=0;
      Sd=0;

/***********************************************************************/
/*   The following loops are where most of the calculation take place */
/*   (births, deaths, and infections).  A random number is generated  */
/*   (usually 1-1000) and if the random number is less then rate      */
/*   currently being checked (i.e. if 150 babies are born per 1000 in */
/*   population and the random number is <150 then a baby is born).    */
/*   This is done once for every person in the population             */
/*                                                 */
/***********************************************************************/


      for(int index = 0;index<MAX_AGE;index++)
      {
        for(int temp = 0;temp<susceptibleAge[index];temp++)
        {
          randNum = rand()%100000+1; //1-100000
          if((index>=15)&&(index<=50)) //ages 15-50 can be infected only
          {
            if(randNum<=((IR0/S)*100000))    //infectionloop
                    {       //(IR0/S)*100000 is chance 1 susep becomes infected
```

```
                        susceptibleAge[index]--;
                        yearsSinceInfection[0].unDiagnosedAge[index]++;
                        firstYearInfected++;
                        }
            }
        randNum = rand()%1000+1;
        if(index>=15&&index<=50) //ages 15-50 can have children only
            if(randNum<=(Sbr*2))
            {
                randNum=rand()%100000+1;
                if(randNum<=((Sbr*2)*(uS/N*10)*100)) //chance baby is uS over S
                {
                    uSb++;
                    unSusceptibleAge[0]++;
                }
                else
                {
                    susceptibleAge[0]++;              //SuscepBirths
                    Sb++;
                }
            }
        if(index<15)   //age<15 deaths
            if((randNum>=600)&&(randNum<=(600+(Sdr/2))))
            {
                susceptibleAge[index]--;
                Sd++;                         //SuscepDeaths
            }
        if(index>=15&&index<=50)   // 15<age<50
            if((randNum>600)&&(randNum<=(600+(Sdr))))
            {
                susceptibleAge[index]--;
                Sd++;                         //SuscepDeaths
            }
        if(index>50)      // age>50
            if((randNum>600)&&(randNum<=(600+(Sdr*2))))
            {
                susceptibleAge[index]--;
                Sd++;                         //SuscepDeaths
            }
    }

    for(int temp = 0;temp<unSusceptibleAge[index];temp++)
    {
        randNum = rand()%1000+1;
        if(index>=15&&index<=50)
        {
            if(randNum<=Sbr*2)   //UnSus Births
            {
                randNum=rand()%100000+1;
                if(randNum<=((Sbr*2)*(uS/N*10)*100)) //chance baby is uS over S
                {
                    uSb++;
                    unSusceptibleAge[0]++;
                }
                else
                {
```

```
                    susceptibleAge[0]++;
                    Sb++;
                 }
              }
          }
          if(randNum<=Sdr)      //UnSusDeaths
          {
              uSd++;
              unSusceptibleAge[index]--;
          }
       }
       for(int temp = 0;temp<infectedAge[index];temp++)
       {
          randNum = rand()%1000+1;              //infectedDeaths
          if(randNum<=Idr)
          {
              infectedAge[index]--;
              dId++;
          }
          if((randNum>200)&&(randNum<=(200+Ibr)))   //infectedBirths
              dIb++;
       }
       for(int temp = 0; temp<10;temp++)
          for(int pop = 0; pop<yearsSinceInfection[temp].unDiagnosedAge[index]; pop++)
          {
              randNum = rand()%1000+1;
              if(randNum<Sbr)          //UnDiagnosed Births
              {
                 udIb++;
                 dIb++;
                 yearsSinceInfection[temp].unDiagnosedAge[index]--;
                 infectedAge[index]++;
              }
              if(randNum>200 && randNum<(200+Sdr))   //UnDiagnosed Deaths
              {
                 yearsSinceInfection[temp].unDiagnosedAge[index]--;
                 udId++;
              }
          }
   }

   for(int index =0 ;index<MAX_AGE;index++)     //immigration loop
       susceptibleAge[index]+=(10000*.01);

   infectedAge[0]+=(dIb*.25);
   susceptibleAge[0]+=(dIb*.75);
   dI=0;
   S=0;
   udI=0;
   uS=0;
   for(int index = 0;index<MAX_AGE;index++)  //retotals pop groups
   {
       uS+=unSusceptibleAge[index];
       S+=susceptibleAge[index];
       dI+=infectedAge[index];
       for(int j=0;j<10;j++)
```

```cpp
                udI+=yearsSinceInfection[j].unDiagnosedAge[index];
    }

    I = dI+udI;    //totals infected with and without symptoms(sick and healthy)

    N = S + I + uS;              //Finds the new total population



    totalInfectedDeaths+=dId+udId;

    fout << setw(5) << int(counter);
    Output();
    fout << endl;



    cout << counter << endl;
    for(int index=8;index>=0;index--)
        yearsSinceInfection[index+1]=yearsSinceInfection[index];
        for(int i=0;i<MAX_AGE;i++)
                    if(counter>=yearsUntilCure)
                    {
                            unSusceptibleAge[i]+=yearsSinceInfection[9].unDiagnosedAge[i]*.95;
                            infectedAge[i]+=yearsSinceInfection[9].unDiagnosedAge[i]*.05;
                    }
                    else
                            infectedAge[i]+=yearsSinceInfection[9].unDiagnosedAge[i];
    for(int i =98; i>=0;i--)    //Ages main groups
    {

        infectedAge[i+1]=infectedAge[i];
        susceptibleAge[i+1]=susceptibleAge[i];
        unSusceptibleAge[i+1]=unSusceptibleAge[i];
        yearsSinceInfection[0].unDiagnosedAge[i]=0; //Zeros fst yr infected
        for(int j=0;j<10;j++)
            yearsSinceInfection[j].unDiagnosedAge[i+1] = yearsSinceInfection[j].unDiagnosedAge[i];
    }
    susceptibleAge[0]=0;
    infectedAge[0]=0;
    unSusceptibleAge[0]=0;

    LowerContact(counter);

    IR0 = (c*i*S)*(I/N);            /*Finds New infected
                                    this equation is contact rate*
                                    infectivity*Susceptible pop*
                                    (infected/total population)
                                    this is called SI model
                                    */



    if(I<1)
        break;

}
```

```cpp
        fout << endl << endl;
        fout << "Total infected deaths: " << totalInfectedDeaths << endl;
        fout << "Population Change: " <<N<<'-'<<startPopulation<<'='<< N-startPopulation << endl;


        return 0;
}
//*************************Input*************************//
void Input()
{

    ifstream fin("input.txt");

    fin >> years;
    fin.ignore(100, '\n');
    fin >> N;
    fin.ignore(100, '\n');
    fin >> I;
    fin.ignore(100, '\n');
    fin >> c;
    fin.ignore(100, '\n');
    fin >> i;
    fin.ignore(100, '\n');
    fin >> Sbr;
    fin.ignore(100, '\n');
    fin >> Sdr;
    fin.ignore(100, '\n');
    fin >> Ibr;
    fin.ignore(100, '\n');
    fin >> Idr;
    fin.ignore(100, '\n');
    fin >> yearsUntilCure;
    fin.ignore(100, '\n');
    fin >> percentCured;
    fin.ignore(100, '\n');
    fin >> yearsUntilImmunization;
    fin.ignore(100, '\n');
    fin >> percentImmunized;
    fin.ignore(100, '\n');
    fin >> popResponse;
    fin.ignore(100, '\n');
    fin.ignore(100, '\n');

      outputN=ShouldThisOutput(fin);
      outputI=ShouldThisOutput(fin);
      outputS=ShouldThisOutput(fin);
      outputuS=ShouldThisOutput(fin);
      outputIR0=ShouldThisOutput(fin);
      outputudI=ShouldThisOutput(fin);
      outputdI=ShouldThisOutput(fin);
      outputSb=ShouldThisOutput(fin);
      outputSd=ShouldThisOutput(fin);
      outputudIb=ShouldThisOutput(fin);
      outputudId=ShouldThisOutput(fin);
      outputdIb=ShouldThisOutput(fin);
      outputdId=ShouldThisOutput(fin);
```

```cpp
      outputuSb=ShouldThisOutput(fin);
      outputuSd=ShouldThisOutput(fin);
}
//**************************ShouldThisOutput**************************//
bool ShouldThisOutput(ifstream& fin)
{
   int temp;
   fin >> temp;  fin.ignore(100, '\n');
   return (temp==1);
}
//**************************Output**************************//
void Output()
{
    if(outputN)
     fout << setw(20) << int(N+.5);
    if(outputS)
     fout << setw(20) << int(S+.5);
    if(outputuS)
     fout << setw(20) << int(uS+.5);
    if(outputI)
     fout << setw(20) << int(I+.5);
    if(outputIR0)
     fout << setw(20) << int(IR0+.5);
    if(outputudI)
     fout << setw(20) << int(udI+.5);
    if(outputdI)
     fout << setw(20) << int(dI+.5);
    if(outputSb)
     fout << setw(20) << int(Sb+.5);
    if(outputSd)
     fout << setw(20) << int(Sd+.5);
    if(outputudIb)
     fout << setw(20) << int(udIb+.5);
    if(outputudId)
     fout << setw(20) << int(udId+.5);
    if(outputdIb)
     fout << setw(20) << int(dIb+.5);
    if(outputdId)
     fout << setw(20) << int(dId+.5);
    if(outputuSb)
     fout << setw(20) << int(uSb+.5);
    if(outputuSd)
     fout << setw(20) <<  int(uSd+.5);
    fout << setw(20) << c;
}
//**************************Cure**************************//
void Cure()
{
   for(int index=0;index<MAX_AGE;index++)
     for(int temp =0;temp<infectedAge[index];temp++)
     {
        randNum = rand()%1000+1;
        if(randNum<=percentCured)
        {
           infectedAge[index]--;
           unSusceptibleAge[index]++;
```

```cpp
        }
    }


}
//**************************Immunization************************//
void Immunization()
{
   for(int index=0;index<MAX_AGE;index++)
      for(int temp =0;temp<susceptibleAge[index];temp++)
      {
         randNum = rand()%1000+1;
         if(randNum<=percentImmunized)
         {
            susceptibleAge[index]--;
            unSusceptibleAge[index]++;
         }
      }


}
//************************LowerContact************************//
void LowerContact(int counter)
{
      static int cDown;
      static float cConst;

      if(counter<10)   //Randomizes Contact Rate
      {
         c=c*.80;
         cConst=c;
      }
      else
      {
         float temp;
         temp = rand();
         temp/=RAND_MAX;
         temp/=10;
         c=cConst+(.05-temp)-((I/N)/popResponse);
      }

}
```

The Following shows the formatting of "input.txt". It is from this file that the program receives initial numbers to start the calculations. These numbers are changes each time the program is run to simulate different scenarios.

150	Number Of years to run
24815492	Total Population
3523	Infected Pop
.7	Contact Rate
.98	Infectivity
17	Sus Brth Rt
6	Sus Dth Rt
2	Inf Brth Rt
150	Inf Dth Rt
10000	Years Until cure is found
500	People Cured per 1000
10000	Years Until immunization is found
300	People Immunized per 1000
1	Population's response to infection(1 is lowest, 10 is highest, larger numbers increase contact rate)
The following are for outputting data, 1=true 0=false	
1	Total Pop
1	Infected Pop
1	Susceptible Pop
1	unSusceptible Pop
1	New Infected this year
1	unDiagnosed infected pop
1	Diagnosed infected Pop
1	Susceptible born this year
1	Susceptible who died this year
1	unDiagnosed infected who were born this year
1	unDiagnosed infected who died this year
1	Diagnosed infected who were born this year
1	Diagnosed infected who died this year
1	unSusceptible born this year
1	unSusceptible dead this year

Appendix 3 - History, Transmission, and Behavior of the AIDS Virus

Appendix 3 - History, Transmission, and Behavior of the AIDS Virus

AIDS (Acquired Immunodeficiency Syndrome) is a worldwide pandemic that has been growing rapidly since its recognition in the early 1980's. It is only recently that treatment has been developed that has allowed people to live with AIDS for up to 10-20 years. It is estimated that as of the end of 2004, 37.2 million adults and 2.2 million children worldwide were living with AIDS. During 2004, an estimated 4.9 million people contracted AIDS, and 3.1 million died of the disease[6].

AIDS is caused by a virus known as HIV (Human Immunodeficiency Virus). HIV is abundant in the bodily fluids of people affected with AIDS, and is usually spread through them. The most common forms of HIV infection are through sexual contact, and through the sharing of needles in intravenous drug use. The virus can also be transmitted from mother to child during childbirth or breast-feeding, though transmission seldom occurs in the womb[7]. Before much was know about AIDS and HIV, another common form of transmission was blood transfusions. After someone has been infected with HIV they are called HIV+, or HIV positive[8].

HIV causes infection by attacking the CD4+ T cells that are supposed to coordinate the body's immune system to fight off infection. After being infected with HIV a person may experience various flu-like symptoms (this is called "seroconversion illness"), but many people do not show any symptoms[8]. Soon, all newly infected people become symptom-free. During this time the HIV particle count in the patient's body increases by billions every day, and the CD4+ T cell count decreases, with varying speeds. Because the virus specifically attacks the cells of the immune system, the antibody response is low or nonexistent, and the virus is allowed to spread unchecked. Typically, once a person's CD4 cell count is below 200 per cubic millimeter, they are considered to have AIDS[8]. The time from being infected with HIV and being diagnosed with AIDS varies; some have been diagnosed with aids only months after becoming HIV+, and some have remained symptom-free for as long as 20 years. The typical time from infection to AIDS is 8 to 10 years[7].

AIDS is closely related to the simian immunodeficiency viruses (SIV). SIV is endemic to many African monkeys and apes, though most of them do not show symptoms of the virus. Scientists speculate that SIV was transferred to humans

through one of many various methods, such as the consumption of raw flesh [9]. The exact location and time where the transference took place is still unknown, but the earliest known fluid sample known to contain HIV was taken in 1959, from a British sailor who had apparently contracted it in what is now known as the Democratic Republic of the Congo[8].

AIDS became officially recognized in 1981, when the American Centers for Disease Control put out a press release, which described five cases of a rare respiratory disease that people with healthy immune systems were seldom affected with. Afterwards several cases of a form of skin cancer were reported, which was also seldom seen in people with healthy immune systems. Both of these diseases were later classified as opportunistic infections, because they are commonly found in patients suffering from AIDS[7]. Other tests performed on these people also found that they were suffering from several other diseases that were also later classified as opportunistic infections. The patients also had a much lower than average CD4+ T cell count. Because most of the original sufferers of the virus were homosexual men, the syndrome was originally referred to by doctors as GRID (Gay-Related Immune Deficiency), but after it was discovered that other sufferers included Haitian immigrants, intravenous drug users, blood transfusion recipients, and heterosexual women, the virus was officially renamed AIDS in 1982[9].

It wasn't until 1984 that researchers were able to isolate the virus that causes AIDS. The two men who discovered the virus, Robert Gallo and Jean Luc Montagnier, agreed to share credit after a long dispute. In 1986 it was officially given the name HIV. This discovery allowed for the development of an antibody test, which allowed people to know if they were infected and at risk of developing AIDS[8].

Currently, there is no cure for AIDS, though in recent years combinations of drugs, known as "cocktails," have allowed people with AIDS to have a much longer lifespan than previously expected.   Most of these treatments usually consist of combinations of two or more anti-retroviral agents. Such treatment has caused patients to repeatedly test negative for HIV, though if the treatment was discontinued the viral loads quickly increased.   These treatments are not perfect, however, and there is

concern that drug resistance will develop[8]. Also, as these treatments are inevitably very expensive, they are not availiable to the vast majority of people living with AIDS.