Modeling the Spread of Preventable Diseases: Social Culture and Epidemiology

Ahmed Y. Tawfik[1] and Rana R. Farag[2]

Abstract   This paper uses multiagent simulation to examine the effect of various awareness interventions on the spread of preventable diseases in a society. The work deals with the interplay between knowledge diffusion and the spreading of these preventable infections in the population. The knowledge diffusion model combines information acquisition through education, personal experiences, and the spreading of information through a scale-free social network. A conditional probability model is used to model the interdependence between the risk of infection and the level of health awareness acquired. The model is applied to study the spread of HIV/Aids, malaria, and tuberculosis in the South African province Limpopo. The simulation results show that the effect of various awareness interventions can be very different and that a concerted effort to spread health awareness through various channels is more likely to control the spread of these preventable infections in a reasonable time.

1 Introduction

Social simulation models have been used to study knowledge diffusion through a social network [2] as well as modeling the spread of epidemics [3]. These models, unlike techniques that focus on cross sectional statistical analysis [8], also capture the factors affecting the individual and at same time allow global trends to emerge in the simulation environment. The bridging of the gap between the micro and macro levels is a useful feature in the study of knowledge diffusion and epidemics. Moreover, hypothetical scenarios are more easily evaluated using social simulations than other epidemiological models. In addition, the relative simplicity of the

[1] Ahmed Y. Tawfik
 The University of Windsor, Windsor, ON N9B 3P4, Canada, atawfik@uwindsor.ca

[2] Rana R. Farag
 The German University in Cairo, New Cairo, Egypt, rana.hassanfarag@student.guc.edu.eg

"agent" in a social simulation makes it possible to explore complex interactions over substantial time durations that are difficult to capture in purely mathematical models.

One such interaction is the interaction between knowledge about disease prevention and the effectiveness of various awareness spreading media in controlling the spread of a preventable disease. In this context, a disease is considered preventable if immunization, life style changes, and other means of prevention are effective in controlling the spread of the disease.

Knowledge about disease prevention may be acquired through formal education, personal experiences, and through advice from a social network of friends, relatives, and neighbors. A physician or a licensed health practitioner would have acquired good knowledge in the area of disease prevention through education. In fact, there is a continuum representing the validity of preventive health knowledge acquired through education. At one end of this continuum, we find the physicians and health care professionals, and on the other end we find individuals with no formal schooling.

Disease prevention knowledge acquired through personal experiences is typically unreliable especially for diseases with long incubation periods like HIV/Aids, tuberculosis, or malaria as individuals are not likely to correlate risky behavior to a latent effect. However, personal experiences could guide an individual trying to avoid mosquito bites. By trying insect repellent or mosquito net, the individual may acquire knowledge that helps in preventing malaria.

Mass media outlets like newspapers, radio, and television are considered useful in promoting healthy life styles and combating diseases. However, the effectiveness of mass media depends on their degree of penetration, the consistency of the message, and the degree of trust that the individuals accords to the media outlets.

Word-of-mouth advice from friends and relatives is also an important means for knowledge diffusion. However, the quality of the information usually deteriorates as it spreads, particularly among the less educated groups.

This paper presents a social simulation model that can be used to assess the effectiveness of various awareness interventions. Section 2 presents the elements of the social simulation model including the agents, the social networks, and the structure of the rules governing the spread of information. Section 3 introduces the simulated society in Limpopo, South Africa and the specific parameters and rules used in the simulations. Section 4 examines the effect of various health awareness interventions on the spread of the diseases and the size of the population. Section 5 is the conclusion of this work.

2 The Social Simulation Model

Initially the model starts by generating a set of agents that represent the demographics of the society to be simulated. The agents are grouped into households according to a set of predefined criteria. If agents in the household share knowledge the health awareness level becomes the same for all agents in the household.

A social network links households based on a set of criteria. The network includes special highly connected nodes representing influential individuals. Knowledge spreads along the social network. Based on their educational background agents may acquire knowledge from other sources including formal education and newspapers. Agents also acquire information from radio and television. As the simulation progresses, children grow up and new households are formed while other households dissolve as agents die. Children in a dissolved household are adopted by other households in the social network. The probability of infection of agents at any time depends on age, gender, and their awareness level.

2.1 The Agents

Each agent has a health status that represents whether the agent is healthy, susceptible, infected and incubating, infected and symptomatic, recovering, immune, or dead. Each health state, except dead, lasts for a period of time that depends on the infection and age. Agents may suffer from more than one infection at the same time. There are some interactions between concurrent infections. For example, any infection is more severe if the patient has HIV.

The agent also has an education level that evolves for children according to the population's demographic data. Education affects awareness in more than one way; it determines the awareness obtained from formal education and the reliability of the agent is relaying information.

Agents also have other demographic characteristics including gender, race, age, lifestyle parameters (e.g. number of sex partners), and a variable representing the degree of religiosity. The degree of religiosity determines the influence of information from preachers on the individual. Preachers are members of a group of highly influential agents that also includes physicians, other health care workers.

2.2 The Household

A household is formed as a result of a marriage. Members of the same household share the same level of health awareness. Children are born in a household and every agent is a member of a household. Unmarried adult children continue to be members of the household. A household dissolves if all adult member of the household die. Any children in the dissolving household are adopted by other households. The ability of a household to accommodate orphans depends on the health state of its members.

The household is the unit of membership in the social network. Media exposure is assumed to be the same for members of the same household. Radio and television ownership is assigned per household. Household members also share income; however, the current version of the simulation model does not take in consideration these economic factors. The interaction between household income and health has been studied previously [1]. Other household attributes include its location as it affects the social network, which typically includes some neighbors.

2.3 The Social Network

The social network is a scale free network where most nodes representing ordinary households with a small number of connections to other households. However, there are some high degree nodes representing influential agent such as the physicians, health care workers, and preachers. Fig. 1 shows a social network where ordinary households are shown in blue while more influential nodes are marked by a box around them.
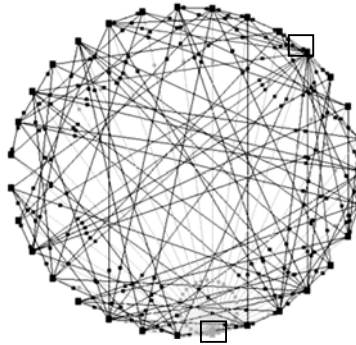

Figure 1 The social network

The network is dynamic to accommodate changes such as household creation and dissolution. There are some compatibility criteria to guide the creation of links between households including educational background and household location.

2.4 Acquisition and Spread of Information

While personal experiences are excellent source of knowledge [4], they are less relevant in this application due to the long incubation periods as discussed earlier. Other sources of knowledge are formal education, peers in the social networks, and influential agents in the social networks.

Each agent is assigned an initial awareness level based on education. Table I provides the interpretation of awareness levels. The higher the level of awareness, the more informed the agent is. Level 9 is assigned to physicians and Level 8 is assigned to other health care workers based on their formal education. The education attainment of a child is determined based on the education attainment of adults in the household.

The increase in awareness level from media differs according to the media and the frequency of the agent's exposure to it. Such that, daily television watching, radio listening, or weekly newspaper reading increases the awareness of the agent by a fixed amount monthly.

Every month the social links of each household are visited and knowledge is shared with a fixed probability assuming that individuals share preventive health

information occasionally. If two friends share knowledge, then their relative awareness is readjusted as follows:

Considering the awareness A_u and A_v of agents u and v, respectively, whenever $A_u > A_v$ then the readjustment is done as: $A'_v = A_v + \gamma A_u$, and $A'_u = A_u - \beta A_v$, where $1 > \gamma > \beta > 0$ to reflect that knowledge diffusion is partial and that the correct information is typically more convincing than incorrect information.

Table 1. Awareness Levels

Level	Interpretation
0	False information
1	
2	
3	*Awareness*
4	*Increases*
5	
6	
7	
8	
9	Correct information

The awareness levels of the physicians and health care workers are not affected because they are confident of their knowledge while the awareness of ordinary individuals is increased as a result of the interaction with them. The amount of the increase is proportional to the agent's education.

2.5 Spread and Effects of Epidemics

The spread of diseases is based on a probability model that takes into account typical epidemiological factors such as age, gender, and lifestyle. However, we add a factor that represents the effect of awareness. For example, the rate of occurrence of infections can be presented as a proportional hazard model [5], such that the factors affecting the hazard rate $h(t)$ include both epidemiological and awareness factors represented by the vector $X = [x_0, x_1, ..., x_n]$ weighed by the coefficients $\alpha_0, \alpha_1, ..., \alpha_n$, or $h(t) = h_0(t)\, e^{\Sigma \alpha_i x_i}$ where $h_0(t)$ represents the unconditional hazard.

Naturally, the coefficient for awareness should ensure that the probability of infection decreases as the awareness increases. Of course, the choice of the survival model and its coefficients will have a great impact on the results.

Once infected, an agent develops symptoms after a random incubation period that depends on the disease, and subsequently may recover or die after another time period. Agents who recover or do not get infected will die at an older age as a result of other causes. A secondary infection to an agent suffering from HIV acts as a time accelerator and the agent die earlier as a result.

3. Case Study: Limpopo

As a case study, we have simulated a village in Limpopo Province in South Africa, which is a well studied example [1] of a community facing great difficulties because of preventable diseases. These problems include high levels of infections with HIV/AIDS, tuberculosis, and malaria. As in other regions in sub-Saharan Africa, the spread of these preventable diseases has had a significant impact on the life expectancy, and on the regional economy.

3.1 The Population

The model starts with 500 individuals whose demographic profile is consistent with recent census data[3] with respect to gender and age. The percentage of the married couples in the actual Limpopo society, which is approximately 22% of adults aged 15 and over, determines the number of households. Throughout the simulation, 1.08% of men marry each year according to official recorded marriage statistics. To account for common law relationships, the model considers that 2.2% of men marry per year, and hence the same number of households constructed. Despite the fact that, in Limpopo's real community, more than one couple may live in the same household, due to the expenses of constructing a new one, the model considers each couple to be living in a separate household so that their friendship relations with other households can be separate. Matching a couple is done randomly, but for a couple to marry, they should meet the following criteria:

- Both partners should be over 15 years old.
- The husband is not older by more than 10 years and the wife is not older with more than 5 years.
- They should be of the same race with a probability of 90%.
- The husband's education level can be higher than the wife's with at most three levels or the wife's education level can be higher with at most one level.

While constructing the community, the model assigns each couple a random number of children between 0 and 3 of each gender. Children inherit the ethnicity of the parents. Each year, a percentage of 4.2% are born (average of birth rates from 1999 to 2003) and the model randomly chooses households for new babies. Single adults in the initial population are also assigned randomly to households.

3.2 Sources of Knowledge

Education levels and initial health awareness is also assigned to individuals based on available statistics. Table 2 shows the statistics[4] used. Health awareness levels within the ranges specified in Table 2 are assigned to agents randomly.

---

[3] Statistics South Africa, Stats in Brief, 2006.

[4]UNESCO, Estimates and projections of adult illiteracy for population aged 15 years and above, by country and by gender 1970-2015.

As the number of physicians in Limpopo is one for every 11,000 individuals and the ratio of nurses to physician is 5.3 to 1, our initial population of 500 is assigned a part time physician who can only see 30 patients a month and a nursing staff that can see 159 patients each month. Priority is given to symptomatic patients but they also see others.

The model assumes that 70% of the population is somewhat religious but the influence of the preacher on a particular agent is proportional to the religiosity factor of the agent.

Table 2. Education and health awareness in the population

Education	Health Awareness Level	Percentage
No schooling	0-1	33.4%
Some primary	0-2	14.1%
Completed primary	2-4	5.5%
Some secondary	2-5	26.1%
Completed secondary	3-6	14%
Higher education	5-8	6.8%

According to a recent study[5], the percentage of the South African population that listens to the radio only once a week is 12.8% and the percentage that listens daily to the radio is 79.3%. While the percentage that watches the television once a week is 10.7%, and daily television watchers constitute 67.3% of the population. Moreover, 40.6% of the population read newspapers on weekly basis. These figures were used to incorporate the effect of media in the simulations.

3.3 The Epidemics

HIV/AIDS primarily infects the 15 to 49 age group. The infection rate is set to 21% which is the average between the World Health Organization statistics and UNAIDS statistics. Mother-to-child infection during pregnancy, birth and breast-feeding occurs with a probability of 25% and children infected through this process do not survive more than 4 years.

Throughout the simulation, the model only considered HIV transmission through either sexual contacts or through mother-to-child transmission. A recent study indicated that the rate of male-to-female infection in South Africa is as high as 74% to 100% [7], and UNAIDS[6] estimates that male-to-female transmission during sexual contact is about twice as likely to occur as female-to-male ones. Therefore, the model considers that with each assigned contact if the male is HIV-

---

[5] BBC World Service Trust, Research Summary Report: African Media Development Initiative, 2006.

[6] UNAIDS, Women and AIDS Fact Sheet, 2004.

positive then with a probability of 87% the female may become infected too but if the female is HIV-positive, the male may get infected with a probability of 44%.

In the initial community, the model chooses 0.511% of the agents randomly to be infected with tuberculosis. During the simulation, 0.6% of the population gets infected annually[7] According to available WHO statistics, 0.18% of the population that gets infected with TB is HIV-positive adults, 0.12% is normal adults, and 0.3% is children, since they are more liable to get infected. Also, 7.2% of infected people die annually, and 40% are cured.

Children under 4 who are co-infected with HIV/AIDS and TB die. Older agents also die from the co-infection if they have reached the symptomatic phase of HIV characterized by a very weak immune system.

In Limpopo, 6369 individuals reported Malaria infection in 2006 which is equal to 0.11% of Limpopo's population[8], but still most of the cases go unreported, so the model assumes that 1% of the agents get infected. Children under 5 years represent 60% of the infected population, the remaining 40% is equally divided between school aged children and adults.

Children under 6 years who are co-infected with HIV and malaria die, while 40% of infected children in this age group die as a result of a malaria infection alone. Malaria also claims the lives of 20% of infected children in the 6 to 14 age group and 10% of adults who are not in the symptomatic phase of HIV. It also causes a precipitous death for symptomatic HIV/AIDS patients.

4. Effect of Awareness Interventions

To study the effects of various awareness interventions, we simulated the system described in Section 2 and the case study introduced in Section 3 using Repast [6]. As we could not obtain enough data to properly validate the hazard model and its parameters, the results presented in this section should be considered illustrative of what social simulation models could produce. Also, in these simulations, we did not take into account changes resulting from medical advances, or changes in the education attainment over the simulation period. Under the assumed set of parameters, the simulations evaluate the following scenarios over 100 year period:

- Scenario 1: Base scenario with no awareness intervention such that an agent's awareness is solely determined by the agent's education.
- Scenario 2: Agents use their social network to share knowledge but did not have access to advice from medical professionals nor preachers.
- Scenario 3: Agents get knowledge through mass media outlets only.
- Scenario 4: Medical professionals and preachers are the only ones spreading knowledge through the social network.
- Scenario 5: Agents share knowledge within each household only.

---

[7] WHO, Global tuberculosis control: surveillance, planning, financing: WHO report 2007.

[8] Department of Health: South Africa, Malaria cases in South Africa: 1999-2007.

- Scenario 6: All channels for knowledge sharing listed above are enabled.

In all the above scenarios, we recorded the population growth and the number of agents infected by each of the three epidemics. Each scenario was run several times for cross validation and the average result reported here. As expected, the population in Scenario 6 was healthier than all the others and grew from 500 to 8500 in 100 years. In Scenario 6, HIV/AIDS was eliminated within the first 32 years while tuberculosis and malaria infections went down as well.

Scenario 2 was the least effective intervention as HIV/AIDS infected more than 40% of the adult population. However, the total population grew to 2230 over 100 years. These results suggest that the dilution of knowledge through the social network in the absence of reliable sources does not add to the population's awareness. These results are obtained using the information transfer rate for correct information (γ) as 0.3 and that of incorrect information (β) as 0.3.

Table 3 gives the population at the end of 100 years and the percentage of the population infected with HIV/AIDS. It is clear that the population grows as the mortality rate due to HIV decreases and the death rate because of these infections remained low.

Table 3. Population and HIV infection at the end of the simulation period

Awareness Intervention Scenario	Final Population	HIV/AIDS % of adults	Tuberculosis (%)	Malaria (%)
1. Formal Education	2204	41.5%	0.55%	0.86%
2. Education + Social Network	2229	41.4%	0.54%	0.86%
3. Education + Mass media	3801	25.5%	0.44%	0.76%
4. Education + Reliable community sources	5161	3%	.0.31%	0.63%
5. Education + Household	7209	0% [*]	0.17%	0.25%
6. All awareness interventions	8500	0% [**]	0.11%	0.19%

[*] Reached after 65 years   [**] Reached after 32 years

The fifth scenario gave the best result for a scenario in which only one method is enabled, due to the presence of many adults living in the same household, so if the awareness level of each became the highest of all their awareness levels, because they live together and have much time to convince each other with what they know, a lot of agents will have high awareness, and therefore the information will be disseminated more quickly.

The results highlight the role of information sharing within the household and in the community. However, these results also warn against the spread of misinformation throughout the social network.

5. Conclusions and Future Research

This study confirms that social simulations can be a useful tool to study the effects of awareness interventions. It shows that various interventions can have significantly different results. As such, this work illustrates an application of multi-agent systems to an important problem. It also shows how to combine traditional epidemiological models and social simulation models to study and analyze the spreading of preventable diseases. The validity of these traditional epidemiological models remains of course the domain of epidemiological studies.

It is important to note that the parameters used in the case study relied on available statistics. However, in the simulations, many additional parameters were arbitrarily chosen. To assess how the results are affected by these parameters, we would like to perform some sensitivity analyses.

References

1. Alam, S.J., Meyer, R. , and Zeirvogel, G. Modelling the socio-economic impact of HIV/AIDS in South Africa. In WCSS 2006: First World Congress on Social Simulation, Kyoto, Japan, August 21-25, 2006.
2. Cointet, J.-P. and Roth, C., How realistic should knowledge diffusion models be? Journal of Artificial Societies and Social Simulation, Vol. 10, No. 3, 2007.
3. Huang, C.-H., Sun, C.-T., Hsieh, J.-L., and Lin, H., Simulating SARS: Small-world epidemiological modeling and public health policy assessments, Journal of Artificial Societies and Social Simulation, Vol.7, No.4, 2004.
4. Kobti, Z., Snowdon, A. W., Rahaman, S., Dunlop, T., and Kent, R. D. A cultural algorithm to guide driver learning in applying child vehicle safety restraint. In the 2006 Congress on Evolutionary Computation, pp. 1111-1118. IEEE, July 2006.
5. Leemis, Lawrence M., Reliability - Probabilistic Models and Statistical Methods, Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1995.
6. North, M.J., N.T. Collier, and J.R. Vos, Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit, ACM Transactions on Modeling and Computer Simulation, Vol. 16, Issue 1, pp. 1-25, ACM, New York, New York, USA, 2006.
7. Pettifor, A., Hudgens, M., Levandowski, B., Rees, H., and Cohen, M., Highly efficient HIV transmission to young women in South Africa., AIDS, Vol. 21, No. 7, pp. 861-865, 2007.
8. Poundstone, K. E., Strathdee, S.A., and Celentano, D. D., The Social Epidemiology of Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome, Epidemiology Review, Vol. 26, pp. 22-35, Oxford University Press, 2004.