Isabel Griffin
Class Project

Swiping Right for an STD: Social-Networking Smartphone Applications, Risky Behaviors, and Sexually Transmitted Diseases

Abstract
      Each year in the United States, 19 million people are diagnosed with an STD, costing the U.S. healthcare system 16 billion dollars. Since the late 1990s, online chat rooms have facilitated opportunities to meet individuals looking to have sexual encounters. In 2009, social networking applications were introduced which not only facilitated the meeting of two risky individuals, but also increased the circumference of a user's sexual network via their Smartphone. The impact of this introduction on the transmission of STDs has yet to be examined in the context of mathematical models. This paper will aim to examine the relationship between the known factors contributing to HIV transmission and the role of using a dating application to meet a sexual partner.

Background
      Each year in the United States, 19 million people are diagnosed with a sexually transmitted disease (STD), costing the U.S. healthcare system 16 billion dollars.[1] The total number of infectious syphilis, chlamydia, and gonorrhea cases in Florida increased from 359.8 per 100,000 persons in 2005 to 588.7 per 100,000 persons in 2015.[2]

STDs and the Internet
      Since the late 1990s, online chat rooms have facilitated opportunities to meet individuals looking to have sexual encounters. Members would solicit encounters, providing age, weight, and preferred sexual act preferences in chat rooms geared towards particular audiences. What resulted were STD outbreaks—most notably an outbreak of syphilis in 1999 that was traced back to a single chat room in San Francisco, California.[3]

The Introduction of Dating Apps
      In 2009, the Social Networking Smartphone Application (SNSA) Grindr was introduced, an app geared towards a gay, bisexual, and bi-curious male audience. Now with the swipe of their finger, users were able to connect with a complete stranger aided by geospatial mapping software that showed their exact distance (in feet). In 2015, two SNSAs Hzone and Positive Singles were released on the market. For the first time in history, users were able to match with partners using SNSAs based on their STD status, a practice commonly referred to as serosorting. This web-based technology not only facilitates the meeting of two risky individuals, but also increases the circumference of a user's sexual network via their Smartphone. App user profiles use abbreviations such as "SWC", meaning "sex without a condom," which may promote risky sexual behaviors and increase this individual's likelihood of being infected with an STD. According to a study by Winetrobe *et. al*, dating app users that included these abbreviations in their profiles were more likely to have had unprotected anal intercourse with their last Grindr-met partner (Winetrobe, 2014). The purpose of this paper is to determine the influence of dating apps in HIV transmission**.**

Risk factors for STD transmission

Known risk factors for STD transmission include: age, gender, race/ethnicity sexual orientation, condom usage, and the total number of sexual partners within a given timeframe.

Public Health Significance

Most adults weigh the cost and benefit of safe sex practices, the fact that HIV is no longer a death sentence and the common belief that STDs are treatable. However, our ability to treat these infections is changing as these viruses and bacteria are becoming resistant to antibiotics, our first line of treatment. The potential public health impact associated with SNSAs and serosorting behaviors not only affect the health of all sexually active adults in Florida, but may also increase the spread of resistant STDs throughout the state via the technology in our pockets.

Methods

Bayesian causal networks are a type of statistical modeling, which allows researchers to obtain a graphical network of variables and outcomes using empirical data. These models are useful when evaluating a phenomenon for which complete data may not be available. For this project, GENIE will be used to model a causal network examining the relationship between known risk factors and Syphilis. Figure 1 depicts the proposed directed acyclic graph (DAG) of relevant nodes and structural dependencies. This DAG examines the relevant nodes known to be associated with STD status: age, gender, race/ethnicity sexual orientation, condom usage, and the total number of sexual partners within the past 12 months. Variables of interest for the purpose of this project include examining the impact of owning a Smartphone and app usage status on the on the outcome of being HIV positive. It is proposed that condom-usage status and number of sexual partners is converging on HIV status. It is also believe that the relationship between age, owning a Smartphone, and app usage is a serial (linear) relationship. Beymer et al. identified that are app users were more likely to be under 35 years old.[4] In the network condom usage and number of sexual partners is diverging from app-usage. Lehmiller *et al.* found that the lifetime average of sexual partners was much higher among app-users compared to non-app users, 75.5 (SD 102.97) and 37.2 (SD 85.92), respectively.[7] A second DAG model (Figure 2) was created in order to examine vectors between race, gender, and sexual orientation.

Figure 1: Bayesian Network Structure with literature-based conditional probabilities.
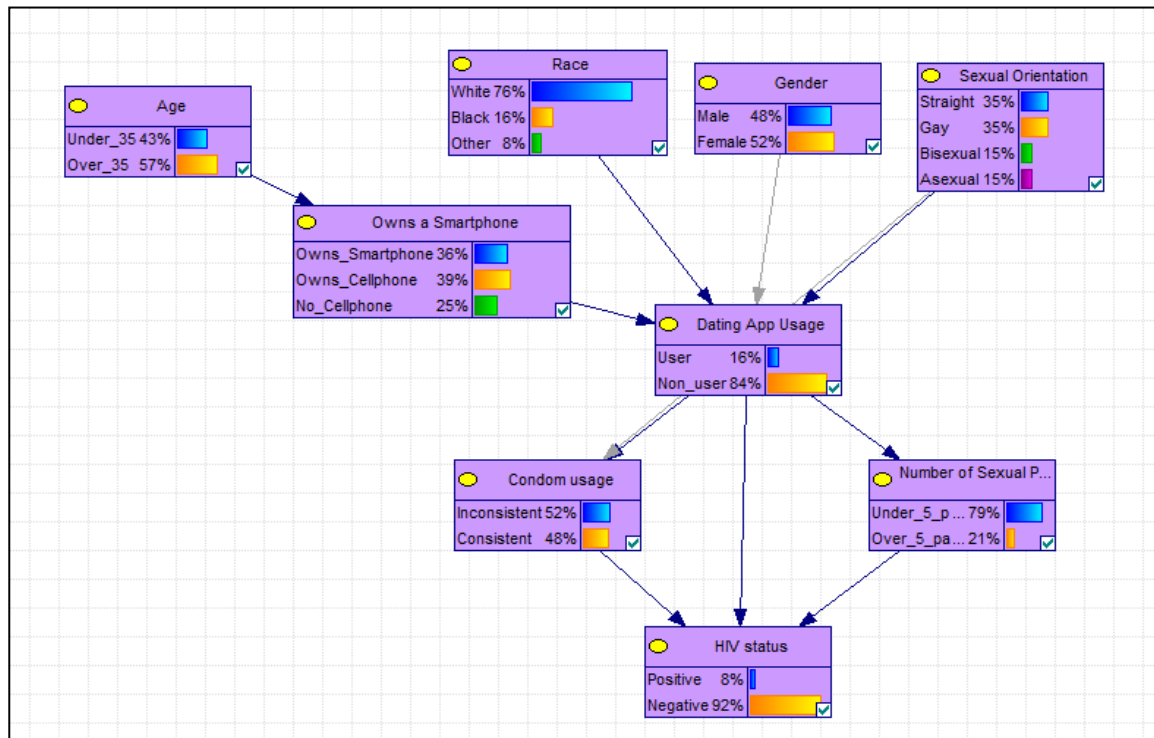
Figure 2: Adjusted Bayesian Network
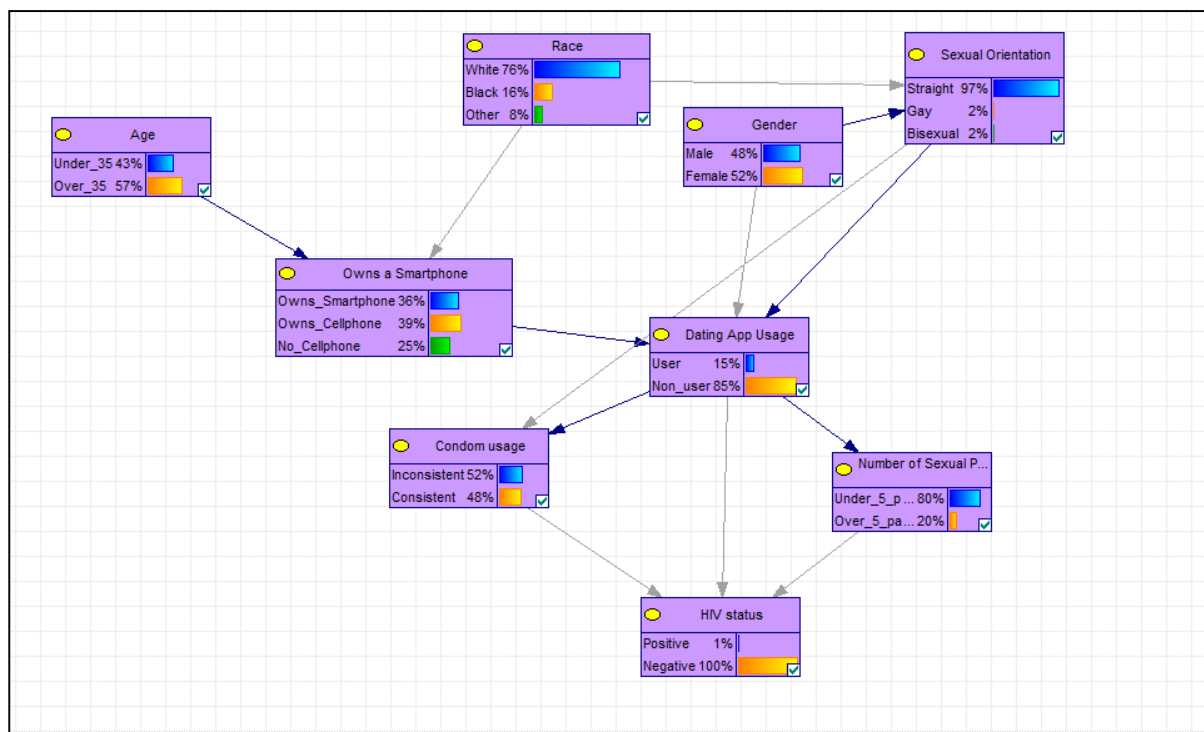
For each node in the DAG, there needs to be a conditional probability distribution. Probability distributions were obtained from studies identified during a literature review conducted between January 1, 2007 and April 21, 2016 using PubMed. Table 1 shows the aggregated data, which was used to calculate conditional probabilities in the model. In the model, each child node probability was adjusted to reflect combination values of parent node. If the child node probability was unknown it was set as a normal distribution. The Genie package was used to not only create the model, but also to examine the impact of having evidence in the model and evaluating the impact on HIV risk.

Aim of Analysis

The goal of the proposed study is to examine the role that dating applications play in facilitating the transmission of sexually transmitted diseases. The main STD outcome of interest addressed in this project is HIV.

Data Source

At this point in time, access to the University of Florida's Florida Cohort Study dataset is pending. As an alternative to evaluating the model using the cohort data, literature review data was aggregated and used to calculate conditional probabilities. A Bayesian network was created and aggregated data obtained from a literature review was used as a guide. A total of five studies were included from the literature review.[4-8] Data obtained from these studies can be found in Table 1 (Appendix). Calculated probabilities included in the GeNIe model can be found in Table 2 (Appendix). Aggregated data was used to calculate probabilities for age (<35 years or ≥35 years), race (white, black, or other), app usage status (yes or no), condom usage status (yes or no), and number of sexual partners (<5 or ≥5 partners). Seeing as all five publications from the literature review involved study populations of MSM, gender probabilities were obtained from the US Census American Community Survey for Florida. The remaining two nodes (owning a Smartphone and sexual orientation) were estimated based on background knowledge (i.e. an educated guess) for the first DAG model. The second DAG used different sexual orientation probabilities from Gates et al. were obtained and applied in the network.[9] HIV prevalence for the state of Florida was obtained from the Florida Department of Health.[10] It was predicted that app users were more likely to be younger and report having a greater number of sexual partners, and thus be at a higher risk for HIV.

Results

As previously mentioned, Figures 1 and 2 depicts the proposed directed acyclic graph (DAG) of relevant nodes and structural dependencies. Figure 3 shows the examination of HIV risk among individuals who both owned a Smartphone and reported being app-users. Thirty-six percent of individuals who owned Smartphones and were also app users would reporting being HIV positive. This is compared to only 8% being HIV positive when there was no evidence set in the model. Figure 4 examines HIV prevalence among Non-users, in which only 3% would report HIV positive. Figure 5 examines HIV risk among smartphone owners who report using apps and that also consistently reported using condoms. When the previously mentioned evidence was set, there was a significant reduction in probability of HIV (from 36% to 3%). Figure 6 shows HIV prevalence

among individuals reporting <5 sexual partners. The model shows that fewer sexual partners may be associated with lower HIV prevalence, as only 4% were HIV positive. On the other hand, Figure 7 shows HIV risk among individuals reporting ≥5 sexual partners. This clearly shows that a large number of sexual partners may be linked to being more likely to report being HIV positive (36% are HIV positive). Figure 8 shows that individuals who both own a Smartphone and use dating apps had the same likelihood of reporting being HIV positive as individuals who did not meet this criteria.

Conclusions

The model suggests that there may be an interaction between dating app usage and HIV status due to the number of sexual partners being reported by app users. It would be worthwhile to continue analyzing this model by conducting sensitivity analysis and model averaging. Bayesian networks are great for showing relationships between variables—even when individual-level data is unavailable. These networks can also be used to show the impact of a public health interventions (for example, condom usage to prevent STDs).

Future analysis

Future steps involving this model include sensitivity analysis and model averaging with the BN learned from background knowledge. In addition, future analysis will include creating a similar Bayesian network using data from the Florida Cohort Study. The data will be analyzed in Banjo. ROC and cross validation tests will be created from the model. To evaluate the network, sensitivity analysis will also be performed. At this point in time, it hasn't been determined whether this sensitivity analysis should be one-way deterministic (varying one parameter at a time over a specified range), or as a probabilistic sensitivity analysis (varying all parameters of the network at once over a specified probability distribution). Maximum a priori probability (MAP) will be examined to determine the probability of the network structure given the data. To examine the strength of the model, cross-validation tests will be performed in which the dataset will be divided into four equally sized sets, and each set will be individually tested again the other 3 datasets. In addition to cross-validation, goodness-of-fit tests will also be performed to determine whether the data behave as though they were sampled from the distribution. A comparison of both the training data set to properties of the simulated data sets of the same size that was generated previously. If the training data behaves differently from what was observed in the majority of simulations, then the data may not have been generated from the learned distribution.

This dataset will allow us to examine the role of serosorting and app usage. Serosorting, in which two HIV positive individuals have unprotected sex, is a common practice among HIV positive individuals. The influence of dating applications in serosorting practices is currently unknown, and has yet to be examined in the context of Bayesian network models.

References

1. Centers for Disease Control. Division of STD Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention. Sexually Transmitted Diseases - Reported Cases and Rates of Reported Cases per 100,000 Population, United States, 1941-2013 Retrieved from: http://www.cdc.gov/std/stats13/tables/1.htm
2. Florida Department of Health. Florida CHARTS Reportable and Infectious Diseases. Retrieved from http://www.floridacharts.com/charts/CommunicableDiseases/default.aspx
3. Dworkin, Mark. "16/Tracking a Syphilis Outbreak Through Cyberspace." *Cases in Field Epidemiology: A Global Perspective*. By Jeffrey D. Klausner. Sudbury, MA: Jones & Bartlett Learning, 2011. 163-71. Print.
4. Beymer, M. R., Weiss, R. E., Bolan, R. K., Rudy, E. T., Bourque, L. B., Rodriguez, J. P., et al. (2014). Sex on demand: Geosocial networking phone apps and risk of sexually transmitted infections among a cross-sectional sample of men who have sex with men in los angeles county. *Sexually Transmitted Infections, 90*(7), 567-572.
5. Bien, C. H., Best, J. M., Muessig, K. E., Wei, C., Han, L., & Tucker, J. D. (2015). Gay apps for seeking sex partners in china: Implications for MSM sexual health. *AIDS and Behavior, 19*(6), 941-946.
6. Chow, E. P., Cornelisse, V. J., Read, T. R., Hocking, J. S., Walker, S., Chen, M. Y., et al. (2016). Risk practices in the era of smartphone apps for meeting partners: A cross-sectional study among men who have sex with men in melbourne, australia. *AIDS Patient Care and STDs, 30*(4), 151-154.
7. Lehmiller, J. J., & Ioerger, M. (2014). Social networking smartphone applications and sexual health outcomes among men who have sex with men. *PloS One, 9*(1), e86603.
8. Rendina, H. J., Jimenez, R. H., Grov, C., Ventuneac, A., & Parsons, J. T. (2014). Patterns of lifetime and recent HIV testing among men who have sex with men in new york city who use grindr. *AIDS and Behavior, 18*(1), 41-49.
9. Gates, G. *How Many People Are Lesbian, Gay, Bisexual, and Transgender?* Los Angeles, CA: Williams Institute, UCLA School of Law, 2011. N. pag. Print.
10. "Florida Charts." *Florida CHARTS Reportable and Infectious Diseases Data*. Florida Department of Health, n.d. Web. 06 Dec. 2016.

Appendix

Table 1: Aggregated data from Literature Review on Dating Apps and STDs

	Beymer	Bien	Chow	Lehmiller	Rendina	**TOTAL**	**5488**
Race							
White	2439	NA	NA	97	180	2716	0.494898
Black	337	NA	NA	NA	114	451	0.082179
Other	2233	NA	NA	NA	85	2318	0.422376
Refused	3	NA	NA	NA	NA	3	0.000547
Age	Beymer	Bien	Chow	Lehmiller	Rendina	**TOTAL**	**10734**
<35	4397	1062	1321	NA	230	7010	0.653065
≥35	2714	280	581	NA	149	3724	0.346935
Gender	Beymer	Bien	Chow	Lehmiller	Rendina	**Total**	**10919**
Male	7184	1342	1902	112	379	10919	1
Female	0	0	0	0	0	0	0
Dating Apps	Beymer	Bien	Chow	Lehmiller	Rendina	**Total**	**10831**
User	2587	545	1097	53	241	4523	0.417598
Non-user	4524	797	805	44	138	6308	0.582402
Condom Usage	Beymer	Bien	Chow	Lehmiller	Rendina	**Total**	**1782**
Inconsistent	NA	NA	947	NA	NA	947	0.531425
Consistent	NA	NA	835	NA	NA	835	0.468575
Sexual Partners in	Beymer	Bien	Chow	Lehmiller	Rendina	**Total**	**378**
<5	NA	NA	NA	NA	NA	197	0.521164
≥5	NA	NA	NA	NA	181	181	0.478836
HIV	Beymer	Bien	Chow	Lehmiller	Rendina	**Total**	**6833**
HIV +	117	49	NA	3	47	216	0.031611
HIV -	4895	1281	NA	109	332	6617	0.968389

**Table 2: Probabilities by app-usage status using aggregate data from literature review**

Condom usage	User	Non-user
Inconsistent	692	255
Consistent	598	237
	User	Non-user
Inconsistent	0.536434	0.518293
Consistent	0.463566	0.481707
	1290	492
Sexual Partners	User	Non-user
<5	97	100
>5	144	37
	241	137
<5	0.40249	0.729927
>5	0.59751	0.270073

Figure 3: Examining HIV risk among an individuals who own a Smartphone and are app-users.
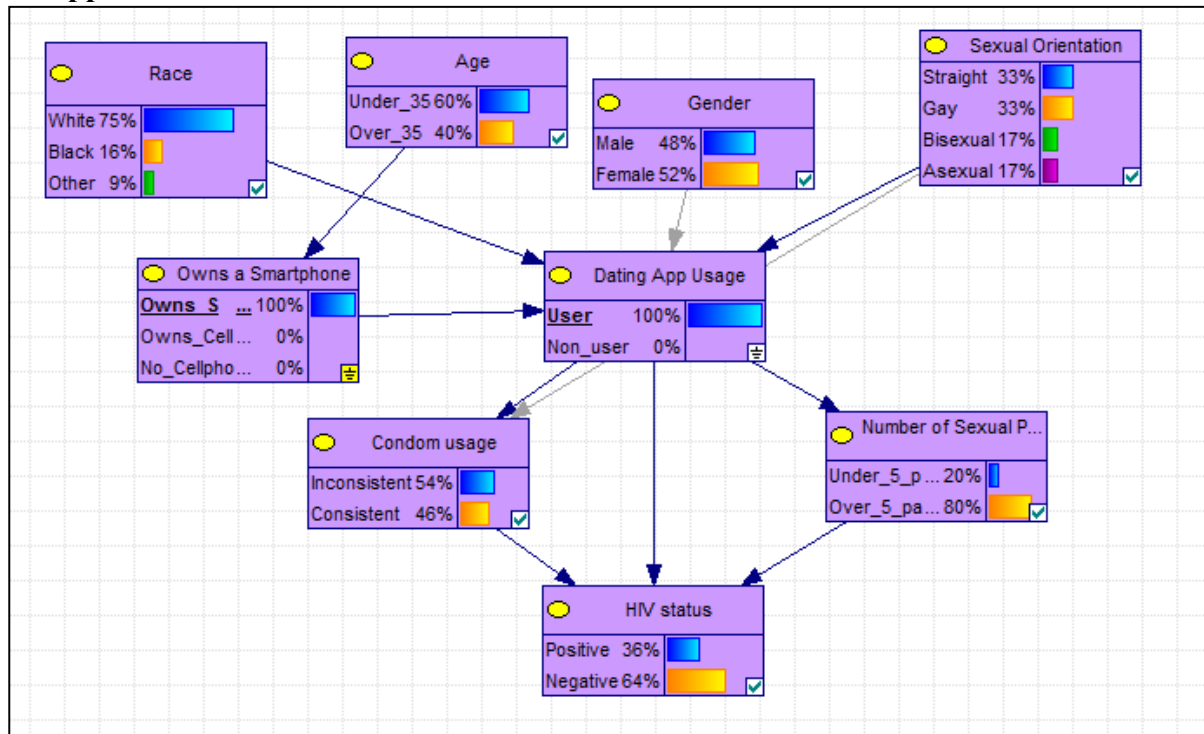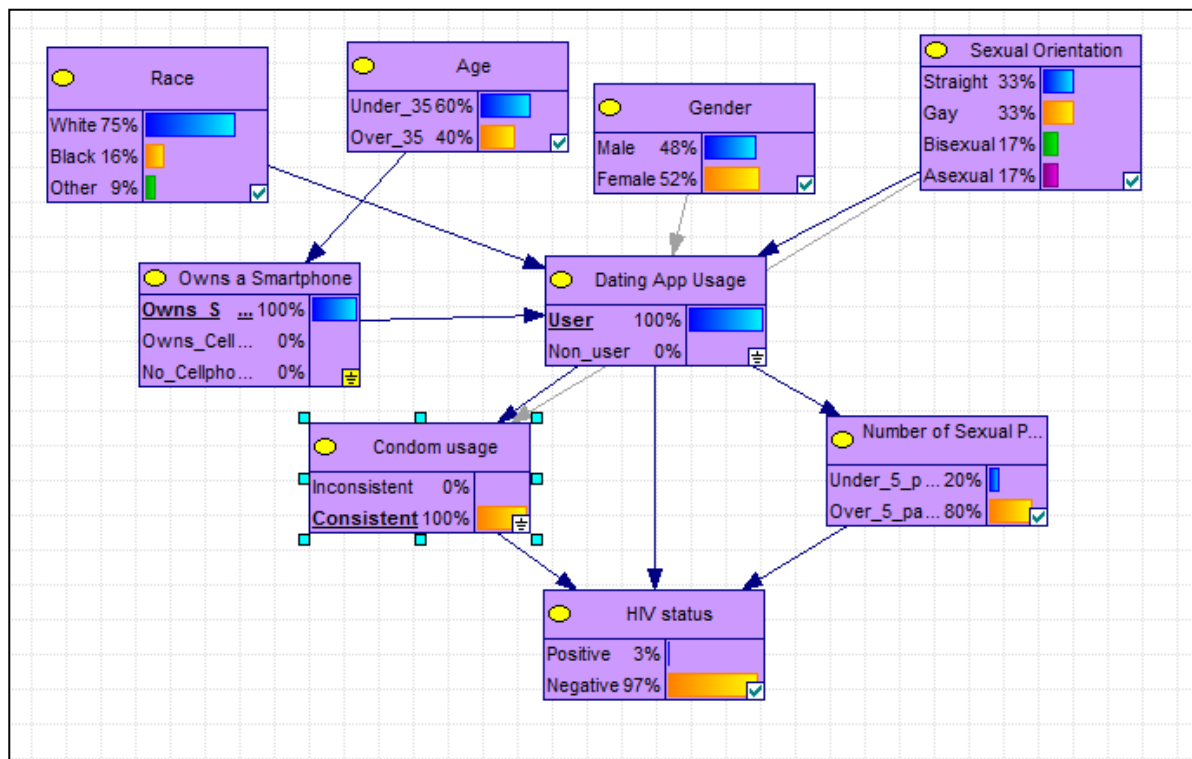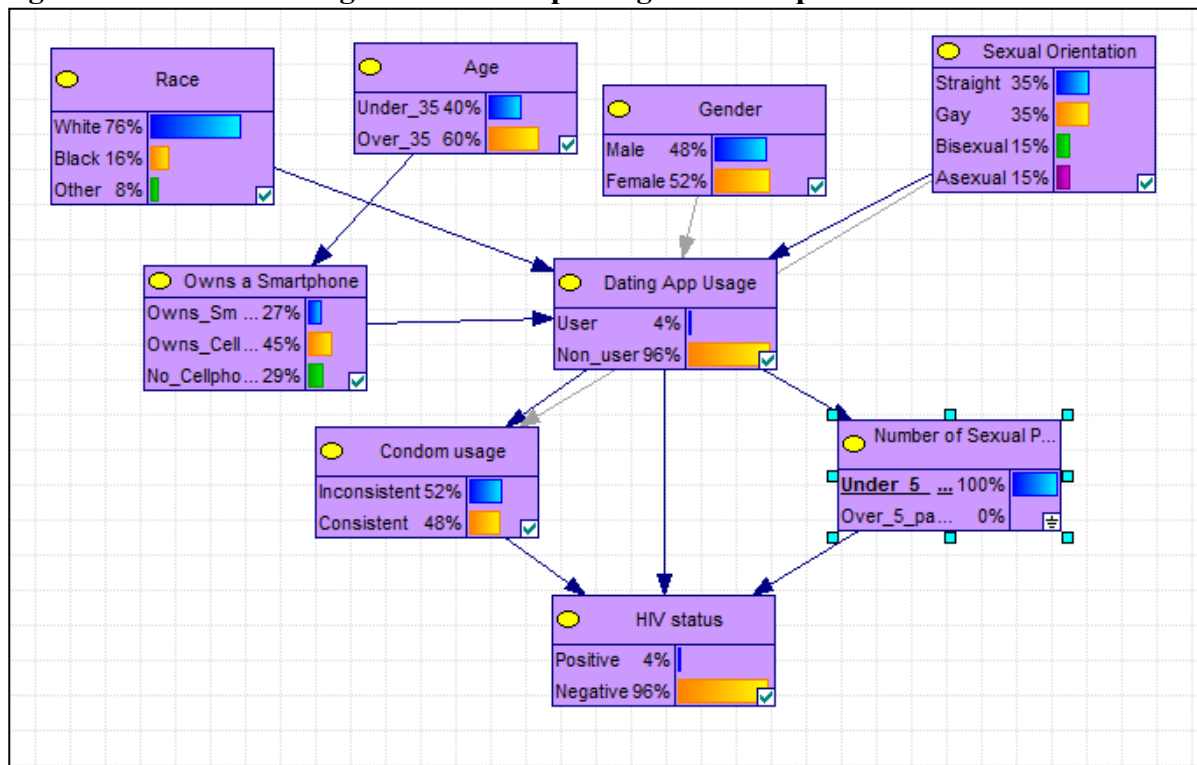
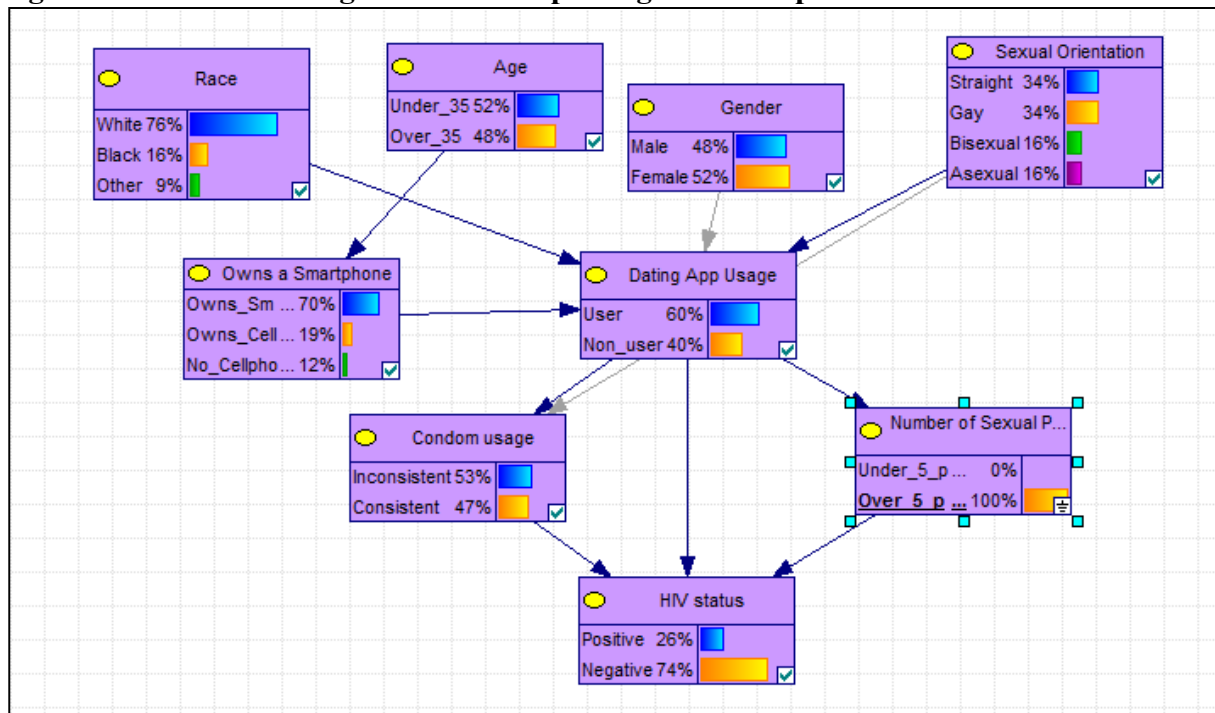**Figure 4: HIV risk among Non-users.**
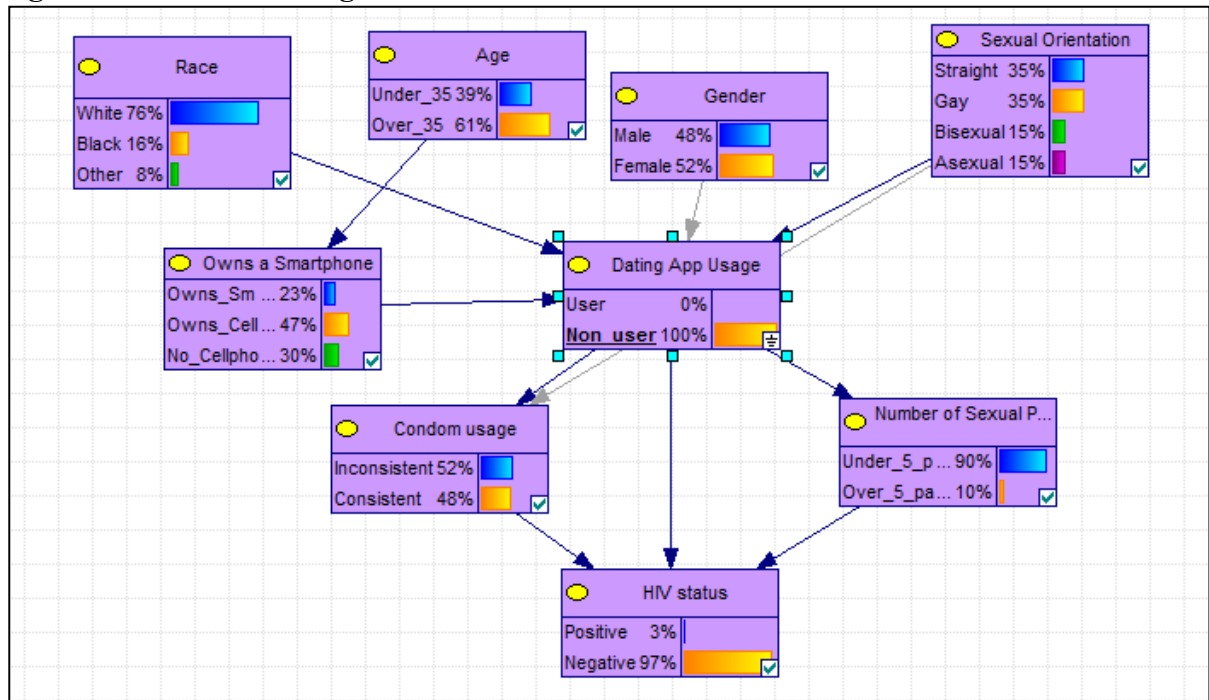
Figure 5: HIV risk among Smartphone owners, using apps, and consistently used condoms.

Figure 6: HIV risk among individuals reporting <5 sexual partners.

Figure 7: HIV risk among individuals reporting ≥5 sexual partners.

Figure 8: Adjusted Bayesian Network among Smartphone owners who report using dating apps.