

BAYESIAN MODELS OF HIV/AIDS SENTINEL SURVEILLANCE TRANSMISSION OF HIV/AIDS IN THE UPPER EAST REGION OF GHANA

¹Clement Ayarebilla Ali M.Sc & ²Francis T. Oduro PhD

¹University of Education, Winneba, Ghana

²Kwame Nkrumah University of Science
and Technology Kumasi, Ghana

Abstract

The first HIV/AIDS Sentinel Surveillance (HSS) surveys started in Ghana in 1992 and Upper East Region in 1994. Since then, the prevalence rates have been the most dominant methods for reporting, monitoring and predicting the pandemic. The purpose of this study was therefore, to use the Bayesian models to analyze the structural transmission of the HIV virus in the HSS sites of the Upper East Region of Ghana. The study gathered secondary data of the infected persons from the four HSS (Bawku, Bolgatanga, Navrongo and Builsa) sites from the Upper East regional administration of the Ghana Health Services, via the office of the Regional HIV/AIDS Coordinator. In the Bayesian analyses, the prior transmission rates represented the initial transmission probability matrix, derived from the assumptions and the likelihood represented the initial proportion of those already being transmitted to from 1994 to 2012. We then used the Matlab software to simulate the posterior transmission probabilities to arrive at the subsequent and steady transmission probabilities. We noticed that the urban sites of Bawku, Bolgatanga and Navrongo would remain the higher than the rural site of Builsa. We therefore, concluded that urban dwellers were more promiscuous to transmit and receive the HIV/AIDS virus than their rural counterparts and recommended that strategic programmes be more focused in the urban areas than the rural ones.

Keywords: *Bayesian, HSS, Posterior, Prior, Steady*

Background to the Study of HIV Reporting in Ghana

Reports of Goubar, et. al. (2007), Eze (2009), Ali & Oduro (2012), Ghana Aids Commission (2012) show that the HIV/AIDS Sentinel Surveillance (HSS) surveys started in Ghana in 1992 and Upper East Region in 1994. Since then, the prevalence rates have been the most dominant methods for reporting, monitoring and predicting the pandemic. These authors however, contend that there is a continuing interest in understanding the evolution of the HIV epidemic and those initial methods have not matched the new definition of the virus. That is Ali and Oduro (2012) had earlier adopted the Markov chain models to analyze the age-gender dynamics of the HIV virus and now seek to widen the scope of the analysis to cover the Bayesian models on the structural transmission of the four HSS sites.

Statement of the Problem of HIV Reporting in Ghana

The researchers observed that the prevalence rates and caseloads continue the preferred methods for reporting, monitoring and predicting the HIV virus infection rates by stakeholders in Ghana. The Prevalence rates however, pose a lot of challenges to predictions because they are just yearly percentages. One cannot use percentages to establish trends with data. The Bayesian methods can adequately address this structural trend defect of the prevalence rates. Another inadequacy and associated challenges originating from the use of prevalence rates (percentages) is that, they report only qualitative data and cannot address quantitative data arising from HIV/AIDS phenomena. Prevalence rates are not appropriate and robust enough to represent HIV/AIDS data. Even if can adopt the prevalence rates to do this, we still modern advanced methods to tackle the complex nature of HIV/AIDS infections across the various towns and villages of the Upper East Region of Ghana.

Hypotheses

Ho: The Bayesian Analysis models are inappropriate methods in monitoring, reporting and predicting the prevalence rates of HIV virus in Upper East Region of Ghana at 5% level of significance.

Ha: The models are more appropriate in analyzing the HIV virus prevalence rates.

Research Questions

1. How are the models appropriate in analyzing the HIV virus spread across the various HSS sites?
2. What HSS site(s) will become dominant(s) in the future?

Purpose of the Study

The purpose of this research was to apply the Bayesian Analysis models to analyze the structural transmission of the HIV virus in the HSS sites of the Upper East Region of Ghana.

Objectives of the Study

The objectives of this study were:

- To showcase the Bayesian Analysis models as more plausible than the traditional methods of monitoring, reporting and predicting HIV virus spread.
- To discover those HSS sites that are highly volatile in transmitting HIV virus.
- To make recommendations to policy makers to help combat the pandemic in the study areas.

Theory of Bayesian Analysis

Applying Bayes' Theorem in HIV Virus Reporting

Bruyninckx (2002), Scott (2003), Wolpert (2006), Goubar et al. (2007), Baffour (2011), and Ofosu & Hesse (2011) explain the conditional probability of iT_{given} (OS_{as follows}):

$$P(T_i | S_0) = \frac{P(S_0 \cap T_i)}{P(S_0)} \dots\dots\dots(1)$$

In the equation (1) above, $P(S_0)$ represent the initial likelihood probabilities of those already infected, $P(T_i)$ represent the conditional initial transmission probability matrix of those assumed to be infected and $P(S_0 | T_i)$ represent the dependent probabilities of those likely to be infected in future. If the assumptions of T_i continue, then the Bayes' theorem of the HIV prevalence rates can be stated as follows:

$$P(T_i | S_0) = \frac{P(T_i)P(S_0 | T_i)}{\sum_{i=1} P(T_i)P(S_0 | T_i)}, i = 1,2,3, \dots\dots\dots(2)$$

In the equation (2) above, the symbols $P(S_0 | T_i)$ is the likelihood given the prior (original rates of infections), $P(T_i | S_0)$ is the prior given the likelihood (assumed rates of infections), $P(T_i)P(S_0 | T_i)$ is the posterior (predicted rates of infections). Wolpert (2006) Eze (2009), and Ali & Oduro (2012) agree that because infected persons interact with the uninfected, $P(S_0 | T_i)$ and $P(T_i | S_0)$ become the dependent $P(S_0)$ and $P(T_i)$ or S_0 and T_i respectively. The $\sum_{i=1} P(T_i)P(S_0 | T_i)$ is the total probability of $P(S_0)$ and serves as only a normalizer. Our discussion will focus on the numerator- $P(T_i)P(S_0 | T_i)$.

Meaning of Bayesian Analysis in HIV Virus Reporting

Bruyninckx (2002), Wolpert (2006) and Xu (2007) explain that any analysis involving prior, likelihood and posterior probabilities is called Bayesian analysis. This analysis complements the classical hypothesis testing methods, which rely on null and alternative hypotheses to make judgments. As applied to the HIV virus, the likelihoods represent the initial proportions of infected persons in the sites, the priors represent new probabilities originating from the likelihoods and the posteriors are their products to produce subsequent probabilities.

Assumptions of Bayes' theorem in the HIV/AIDS Virus Transmission Probabilities

Draper (2005), Wolpert (2006), Goubar et al. (2007), Johnson et al. (2009), Greenwood et al. (2010), and Ali & Oduro (2012) make some basic assumptions to guide this Bayesian analysis as summarized below:

- 1). The uncertainties can be measured with prior transmission probabilities. This will allow us to use the priors as the transmission probability matrix.
- 2). The priors can form a square transmission probability matrix (Ti) and sketch the transmission probability diagram and matrix (iT).
- 3). The Ti can be computed iteratively as a non-singular matrix with the Matlab software.
- 4). First infected persons in each site are highly promiscuous to keep a chain of Ti-transmit and contract the disease.

Representing HIV Virus with Bayesian Transmission Probability Diagram

Barnett, Ziegler & Byleen (2000), Bruyninckx (2002), Cao (2007), Ali & Oduro (2012) assert that a transmission probability diagram is a two-way flow of transmitting to and receiving from the HIV virus. In the case of the four HSS sites, the researchers showed how each site could infect and being infected by the other sites in the analysis below.

Representing HIV Virus with Bayesian Transmission Probability Trees

Wolpert (2006), Ali & Oduro (2012) explain the Bayesian transmission probability tree as a chain sequence of infinite dependent events of the prior transmission probabilities and the likelihoods. The first pairs of branches are the initial likelihoods (S₀) and the next branches as the initial priors (T_i). Because the priors are dependent, we could continue building up infinite number of branches.

Representing HIV Virus with Bayesian Probability Transmission Matrix

Barnett, et al. (2000), Chen (2003), Mathews & Fink (2004), Johnson, Dorrington & Bradshaw (2009), Ali & Oduro (2012) explain the transmission probability matrix as the initial sequences of likelihoods as extracted from the initial Bayesian transmission probability tree as follows:

$$T_i = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1j} \\ t_{21} & t_{22} & \dots & t_{2j} \\ \vdots & \vdots & \dots & \vdots \\ t_{2j} & t_{2j} & \dots & t_{ij} \end{bmatrix} \dots \dots \dots (3)$$

In the equation (3) above, the transmission probability matrix iT represents the initial priors. The following observations can be drawn from the equation (3) above:

- 1). The priors form initial probability matrix at each level.
- 2). The sum of entries of each row in a prior is 1.
- 3). The sum of probabilities in each row of the posterior is also 1.
- 4). Any k-step simulated probability matrix has the same properties as this initial prior.

Computing HIV Virus Subsequent and Steady Transmission Probabilities

Mathews & Fink (2004), Weisstein (2005), Pereira & Stern (2008), Eze (2009), Johnson, et al. (2009), Ali & Oduro (2012) explain that the steady transmission probability is the probability that, in a discrete-time, the Bayesian process will converge to a limiting value, S which is independent of the initial transmission probabilities (S₀). In The Matlab software simulation procedures, the first posterior is the same as:

This is because these four sites have always been used by the Ghana Aids Commission to monitor the prevalence rates and caseloads of the disease across in the region.

Draper (2005) and Goubar et al. (2007) justify that there are no samples and populations selection procedures in the Bayesian methods. Hence, we used the whole to derive the initial transmission probabilities (S_0) as 0.4, 0.3, 0.2 and 0.1, representing initial infected persons of Bawku-Ba, Bolgatanga-Bo, Navrongo-Na and Builsa-Bu sites respectively. This was restated in a row matrix form as below:

$$S_0 = [0.4 \quad 0.3 \quad 0.2 \quad 0.1] \dots \dots \dots (6) .$$

In the equation (6) above, the entries represent the likelihoods or initial proportions of infected persons in the four sites. The researchers then applied some assumptions to each of the four sites to derive the initial transmission probability matrix, iT.

Presentation of Bayesian Analysis Models of HSS Sites Transmissions

Derived assumptions for initial transmission probability matrix

- 1). All the sites interact with one another (where Bawku— aB , Bolgatanga— oB , Navrongo— aN and Builsa— uB).
- 2). If Bawku is the primary source of transmission, then the routes before a first person could be transmitted with HIV in each site would be traced as:

$$B_a \rightarrow B_o \rightarrow N_a \rightarrow B_u$$

$$B_a \rightarrow N_a \rightarrow B_u \rightarrow B_o$$

$$B_a \rightarrow B_u \rightarrow B_o \rightarrow N_a$$

- 3). If Bolgatanga is the primary source of transmission, then the routes before a first person could be transmitted with HIV in each site would be traced as:

$$B_o \rightarrow B_u \rightarrow N_a \rightarrow B_a$$

$$B_o \rightarrow N_a \rightarrow B_a \rightarrow B_u$$

$$B_o \rightarrow B_a \rightarrow B_u \rightarrow N_a$$

- 4). If Navrongo is the source of transmission, then the routes before a first person could be transmitted with HIV in each site would be

traced as:

$$N_a \rightarrow B_o \rightarrow B_u \rightarrow B_a$$

$$N_a \rightarrow B_u \rightarrow B_a \rightarrow B_o$$

$$N_a \rightarrow B_a \rightarrow B_o \rightarrow B_u$$

- 5). If Builsa is the source of transmission, then the routes before a first person could be transmitted with HIV in each site would be traced as:

- $B_u \rightarrow B_o \rightarrow N_a \rightarrow B_a$
- $B_u \rightarrow N_a \rightarrow B_a \rightarrow B_o$
- $B_u \rightarrow B_a \rightarrow B_o \rightarrow N_a$

6). If these assumptions remain the same, then we can derive the initial transmission probability matrix of the HSS Sites.

Initial Bayesian Transmission Probability Diagram of HSS Sites

It is convenient to represent the initial proportions of transmissions on a directed probability transmission diagram as follows:

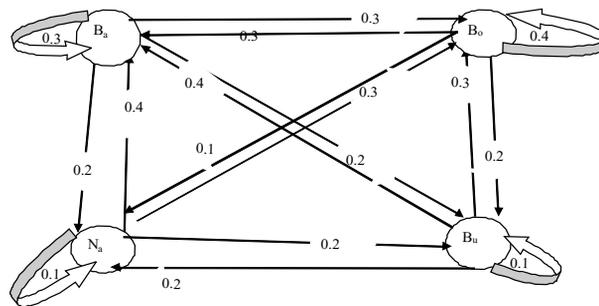


Figure 1: Initial Bayesian Transmission Probability Diagram of HSS Sites

In the figure (1) above, the sites are the nodes and the initial transmission probabilities are the edges. Bawku site initially had 30% of the HIV infected persons and shared 70% with the three sites, Bolgatanga site initially had 40% of the HIV infected persons and shared 60% with the three sites, and each site of Navrongo and Builsa initially had 10% of the HIV infected persons and shared 90% with the other sites. We extracted the figures and formed an initial transmission probability matrix.

Initial Transmission Probability Matrix of HSS Sites

The equation (7) below represents the initial transmission probability matrix extracted from figure (1) above:

$$T_i = \begin{bmatrix} & B_a & B_o & N_a & B_u \\ B_a & 0.4 & 0.3 & 0.2 & 0.1 \\ B_o & 0.3 & 0.4 & 0.2 & 0.1 \\ N_a & 0.2 & 0.3 & 0.3 & 0.2 \\ B_u & 0.1 & 0.2 & 0.3 & 0.4 \end{bmatrix} \dots\dots\dots(7)$$

In the equation (7) above, the first row represents the initial transmissions of Bawku site, the second row represents the initial transmissions of Bolgatanga site, the third row represents the initial transmissions of Navrongo site and the fourth row represents the initial transmissions of Builsa site. We have observed that the sum of entries of each row is 1 and there are no 0 entries in any of the rows. These mean that initially any infected person from any site could infect any first uninfected person. We now have to establish the initial four-route transmission probability tree of HSS to compute the subsequent and obtain the steady transmission probabilities of all the sites.

Initial four-state transmission probability tree of HSS sites

We combined both the initial proportions (S0) and assumptions (Ti) to form the initial transmission probability tree as follows:

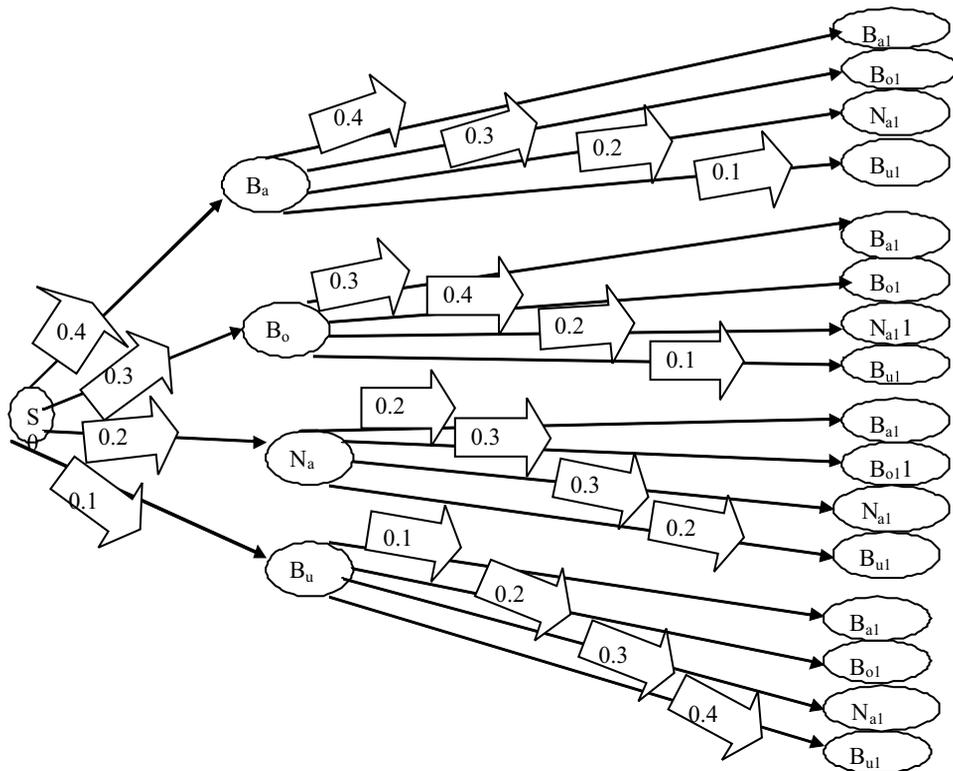


Figure 2: Initial four-state transmission probability tree of HSS sites

In the diagram (2) above, the first note is the initial proportions of the sites and the subsequent branches are the initial prior transmission probabilities that each site was expected to transmit to the other sites. We could build another set of branches on the first set, and continue the build up of the branches infinitely. We have observed that each site can infect every other site. We have extracted the figures to compute the subsequent and steady transmission probabilities.

Steady Transmission Probabilities of HSS Sites

The Matlab software applies the matrix multiplication to the equations 5a and 5b to compute the subsequent and derive the steady transmission probabilities as follows:

$$S_0 T_{ij}^{(k)} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.2 & 0.1 \\ 0.2 & 0.3 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{bmatrix}^{(k)} = [0.2671 \quad 0.3136 \quad 0.2419 \quad 0.1774]..(8)$$

In equation (8) above, the process reached the steady transmission probabilities (S) after only 9 steps. This mean the probability that the first uninfected person to get the HIV virus in Bawku, Bolgatanga, Navrongo and Builsa are 0.2671 , 0.3136 , 0.2419 and 0.1774 respectively. The Chapman-Kolmogorov equation (6c) also computes the subsequent and derives the steady transmission probability matrix as:

$$\lim_{k \rightarrow \infty} T_{ij}^{(k)} = \begin{bmatrix} & B_a & B_o & N_a & B_u \\ B_a & 0.2670 & 0.3136 & 0.2419 & 0.1774 \\ B_o & 0.2670 & 0.3136 & 0.2419 & 0.1774 \\ N_a & 0.2670 & 0.3136 & 0.2419 & 0.1774 \\ B_u & 0.2670 & 0.3136 & 0.2419 & 0.1774 \end{bmatrix} \dots\dots\dots(9)$$

In the equation (9) above, we have observed that the entries of all the four rows are the same. This means we have obtained steady transmission probability matrix. Thus, similarly, over a long period of time, an HSS site with the given four-state transmission probability matrix should reach steady transmission probabilities at about 26.70% in Bawku, 31.36% in Bolgatanga, 24.19% in Navrongo and 17.74% in Builsa.

Properties regular promiscuous transmission probabilities

The long run transmissions possess the regular promiscuous transmission probabilities. This is because each row of equation (10) is just the equation (9). We have also observed that every entry in the rows of equation (10) is greater than 0. These mean that all sites were always promiscuous to transmit and receive the HIV virus. Thus, any infected person from any site could infect any first uninfected person within the four sites.

Discussions of Transmission Probabilities of HSS Sites

We observed that the initial transmission probabilities were $\begin{bmatrix} B_a & B_o & N_a & B_u \\ 0.4 & 0.3 & 0.2 & 0.1 \end{bmatrix}$ as obtained from the Ghana Aids Commission (GAC, 2012). We have computed and derived the final steady transmission probabilities as $\begin{bmatrix} B_a & B_o & N_a & B_u \\ 0.27 & 0.31 & 0.21 & 0.18 \end{bmatrix}$

representing the infected persons from Bawku, Bolgatanga, Navrongo and Builsa sites respectively. This means the incidence of HIV infections would be stabilized in future in the region. We envisaged however, that the urban sites of Bawku, Bolgatanga and Navrongo would remain the highest in future. So, educational programmes and other preventive measures must be swifter in the urban than the rural sites.

That notwithstanding, the rural Builsa site also kept increasing as the years went by and therefore, must be placed on red alert! This is because initially, they represented only 10% of the infected persons in the data. However, their proportion had significantly increased to about 18% at the end of the iterations. This therefore, suggests that gradually the urban sites would shed off a significant portion of their infections to the rural ones.

Summary and Conclusions

We obtained the requisite HIV/AIDS data of the four HSS sites from 1994 to 2012 to represent the initial likelihoods, S0. We successfully adopted strong assumptions to derive the initial prior transmission probability matrix, Ti for the multiplication with the S0. The Matlab software applied the Chapman-Kolmogorov equations to derive the subsequent and steady transmission probabilities and matrix. The subsequent and steady transmission probabilities indicated highly promiscuous and regular HIV infections among the four HSS sites. However, we discovered that the urban areas were more volatile than the rural ones. Hence they must be targeted more while placing eagle eyes on rural ones. Therefore, we have

succeeded in showing that the Bayesian models can better report, monitor and predict the HIV/AIDS in the Upper East Region of Ghana.

Recommendations for policy making

The researchers observed that the Bayesian models very suitable for data analyses in HIV/AIDS. If future studies could involve persons from the Ghana Aids Commission, Regional HIV/AIDS Coordinators, HIV/AIDS Focal Persons and other relevant institutions, much success could be achieved. The analyses also showed that the patterns of HIV/AIDS prevalence rates may not change from its traditional higher records in the urban centres in future. Therefore, we recommend that if future studies could take data directly from the health centres of the region, better analyses could be depicted.

The Matlab software simulations depicted a few number of nine iterations to reach the steady transmission probabilities. This is because all the transmission created an easy situation that the infected could contact the uninfected persons without difficulty. We recommend that if future studies could get access to interview the infected persons to get into details of this volatile infectious transmission matrix.

Suggestions for Further Research on the Bayesian Analysis on HIV/AIDS Virus

- (1). The study applied secondary data to the Bayesian models and could have contained some unforeseen measurement errors. As a result of that, the conclusions may not support the empirical evidence of the pandemic in the region.
- (2). The data was taken from only the four HSS sites of the Upper East Region of Ghana. As a result of that, one cannot generalize the findings to all other regions of Ghana.

References

- Ali, C.A. and Oduro, F.T. (2012). "Markov Chain Models of Age and Gender Structure Dynamics of HIV/AIDS Transmission in the Upper East Region of Ghana". A Research Paper Presented at the 5th Annual International Applied Science Research, Koforidua Polytechnic, 25th to 28th June, 2012 (Unpublished).
- Barnett, R.A. Ziegler, M.R. and Byleen, K.E. (2000). "Applied Mathematics for Business, Economics". Life Sciences and Social Sciences (7ed). New Jersey: Prentice-Hall, p.463-509.
- Baffour, E. (2011). "Basic Statistics and Probability". MAB 592. Winneba: Institute for Educational Development and Extension, University of Education: Winneba, p93
- Bruyninckx, H. (2002). "Bayesian Probability".
[<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.4949&rep=rep1&type=pdf>, accessed on 9th October, 2013]
- Cao, X.R. (2007). "Stochastic Learning and Optimization, A Sensitive-Based Approach". New York: Springer [http://www.ee.ust.hk/~eeca/paper/Cao09_annual_reviews.pdf, accessed on 9th October, 2013]
- Chen, J. (2003). "Applied Probability Theory (Lecture Notes)". Fall: Waterloo University [<http://www.stat.ubc.ca/~jhchen/stat461/title.pdf>, accessed on 9th October, 2013].
- Draper, D. (2005). "Bayesian Modeling, Inference and Prediction". University of California, Santa Cruz: Department of Applied Mathematics and Statistics. [<http://users.soe.ucsc.edu/~draper/draper-BMIP-dec2005.pdf>, accessed on 25th June, 2013].
- Eze, J.I. (2009). "Modelling HIV/AIDS Epidemic in Nigeria". A Dissertation Submitted to the University of Glasgow for the Degree of Doctor of Philosophy in the Department of Statistics [<http://theses.gla.ac.uk/642/01/2009EzePhD.pdf>, accessed on 25th June, 2013]

- Ghana AIDS Commission (2012). "2012 Annual Report. National Aids/STI Programme. Accra: Ghana Aids Commission". [<http://www.ghanaims.gov.gh/gac.pdf>, accessed on 9th October, 2013].
- Goubar, A., Ades, A.E., De Angelis, D., McGarrigle, C.A., Mercer, C.H., Tookey, P.H. Fenton, K and Gill, O.N. (2007). "Estimates of HIV prevalence and proportion diagnosed based on Bayesian multi-parameter synthesis of surveillance data". U.K.: Medical Research Council [http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1194947329693, accessed on 25th June, 2013]
- Greenwood, J., Kircher, P. and Tertilt, M. (2010). "An Equilibrium Model of the African HIV/AIDS Epidemic". [<http://personal.lse.ac.uk/kircher/Papers/HIV-AIDS.pdf>, accessed on 25th June, 2013].
- Johnson, L.F., Dorrington, R.E. and Bradshaw, D. (2009). "A Bayesian Approach to Assessing Uncertainty in a Model of the HIV/AIDS Epidemic in South Africa. University of Cape Town: Burden of Disease Research Unit". [<http://www.statssa.gov.za/isi2009/ScientificProgramme/IPMS/0161.pdf>, accessed on 25th June, 2013].
- Mathews, J.H. and Fink, K.D. (2004). "Numerical Methods Using Matlab (4th ed)". New Jersey: Pearson Prentice Hall: Inc, 600-609
- Ofosu, J.B. and Hesse, C.A. (2011). "Introduction to Probability and Probability Distributions". Accra: Excellent Publishing and Printing: 22-35.
- Pereira, C. A. B. and Stern, J. M. (2008). "Can a Significance Test Be Genuinely Bayesian"? [<http://www.ufscar.br/~polpo/cis/resumos/jstern.pdf>, accessed on 9th October, 2013].
- Scott, C. (2003). "Bayesian Estimation" [<http://arxiv.org/pdf/astro-ph/0212497.pdf>, accessed on 9th October, 2013].
- Weisstein, E. W. (2005). "Markov Chains" [://mathworld.wolfram.com/MarkovChain.html, accessed on 14th March, 2013].
- Wolpert, R. (2006). "International Society for Bayesian Analysis (ISBA)". [http://en.wikipedia.org/wiki/International_Society_for_Bayesian_Analysis, accessed on 9th October, 2013].