



Statistical Methods for Modeling HIV/AIDS in India

Corresponding Author:

Dr. Brijesh Sathian,

Assistant Professor & Managing and Chief Editor NJE, Community Medicine, Manipal College of Medical Sciences, Department of Community Medicine, Manipal College of Medical Sciences, 155 - Nepal

Submitting Author:

Dr. Brijesh Sathian,

Assistant Professor & Managing and Chief Editor NJE, Community Medicine, Manipal College of Medical Sciences, Department of Community Medicine, Manipal College of Medical Sciences, 155 - Nepal

Article ID: WMC003336

Article Type: Review articles

Submitted on: 05-May-2012, 06:08:59 PM GMT **Published on:** 07-May-2012, 04:24:51 PM GMT

Article URL: http://www.webmedcentral.com/article_view/3336

Subject Categories: BIOSTATISTICS

Keywords: Statistical Modeling, HIV/AIDS, Curve fitting, India

How to cite the article: Sathian B, Sreedharan J. Statistical Methods for Modeling HIV/AIDS in India . WebmedCentral BIOSTATISTICS 2012;3(5):WMC003336

Copyright: This is an open-access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Source(s) of Funding:

Not Applicable

Competing Interests:

No competing interests

Statistical Methods for Modeling HIV/AIDS in India

Author(s): Sathian B, Sreedharan J

Abstract

Deterministic, Stochastic, Statistical and State space models are the statistical models for forecasting HIV/AIDS data. There are also uncertainties associated with these approaches. In addition to this the recent advance in this field used is curve fitting models. Sathian B and Sreedharan J used this method for forecasting several infectious and non infectious diseases. It gives more accurate estimates compared to the other models.

Introduction

Adult HIV prevalence in India declined from 0.41% in 2000 to 0.31% in 2009. The 2008-09 India HIV estimates developed by NACO with support from National Institute of Medical Sciences, National Institute of Health and Family Welfare, UNAIDS and WHO utilised improved methodology and updated epidemiological data from the latest rounds of HIV Sentinel Surveillance and other information on High Risk Groups for more accurate understanding of the Indian epidemic. It is estimated that India had approximately 1.2 lakh new HIV infections in 2009, as against 2.7 lakh in 2000[1]. Statistical methods have expanded spatially in recent years to address large scale worldwide health issues. These methods have a prominent role in the study of the HIV/AIDS epidemic. Deterministic Models, Stochastic Models, Statistical models and State Space Models are the main four categories of statistical modelling. Number of susceptible individuals, infected individuals and number of AIDS cases used as parameters in Deterministic modeling. But in Stochastic models, it will be random variables and superior than deterministic models. Statistical models are better than stochastic and deterministic because it uses the epidemiological and survey data with back calculation methodology. HIV incubation period is the random time between the HIV-infection and the onset of clinical AIDS. Distribution of this non-negative random variable is known as HIV incubation period distribution. The Back-calculation method reconstructs the past pattern of HIV infection and predicts the future number of AIDS cases with the present infection status. It depends on three important factors: incubation distribution, incidence curve and observed number of

AIDS cases over time. This method is very popular and requires less information and assumptions. Lack of information about incubation distribution, the effect of intervention therapy on incubation period, and errors in reported AIDS incidence leads to uncertainties associated with this method. The incubation distribution is assumed to be exactly known in back-calculation methodology. Incubation period of HIV is very long and highly variable within and between cohorts. The current prevalence of HIV-infection and the corresponding pattern of incidence from the beginning of the epidemic to the present time are mainly estimated by means of back-calculation method. It calculates the most likely temporal distribution of infected individuals compatible with the number of observed AIDS cases starting from the suitable estimate of the incubation period, derived from the available data. State space models have the combined effect of stochastic and statistical models but it is mainly for engineering data and used in forecasting of AIDS. There are also uncertainties associated with these approaches. In addition to this the recent advance in this field used is curve fitting models. Sathian and Sreedharan used this method for forecasting several infectious and non infectious diseases. It gives more accurate estimates compared to the other models [2-28].

Curve fitting method

The Curve Estimation procedure produces curve estimation regression statistics and related plots for 11 different curve estimation regression models. A separate model is produced for each dependent variable. You can also find out predicted values, residuals, and prediction intervals as new variables.

For each model, we can find out the regression coefficients, multiple R, R^2 , adjusted R^2 , standard error of the estimate, analysis-of-variance table, predicted values, residuals, and prediction intervals. Models: linear, logarithmic, inverse, quadratic, cubic, power, compound, S-curve, logistic, growth, and exponential.

Linear Model

Model whose equation is $Y = b_0 + (b_1 * t)$. The series values are modeled as a linear function of time.

Logarithmic Model

Model whose equation is $Y = b_0 + (b_1 * \ln(t))$.

Inverse Model

Model whose equation is $Y = b_0 + (b_1 / t)$.

Quadratic Model

Model whose equation is $Y = b_0 + (b_1 * t) + (b_2 * t^{**2})$. The quadratic model can be used to model a series that "takes off" or a series that dampens.

Cubic Model

Model that is defined by the equation $Y = b_0 + (b_1 * t) + (b_2 * t^{**2}) + (b_3 * t^{**3})$.

Power Model

Model whose equation is $Y = b_0 * (t^{**b_1})$ or $\ln(Y) = \ln(b_0) + (b_1 * \ln(t))$.

Compound Model

Model whose equation is $Y = b_0 * (b_1^{**t})$ or $\ln(Y) = \ln(b_0) + (\ln(b_1) * t)$.

S-curve Model

Model whose equation is $Y = e^{*(b_0 + (b_1/t))}$ or $\ln(Y) = b_0 + (b_1/t)$.

Logistic Model

Model whose equation is $Y = 1 / (1/u + (b_0 * (b_1^{**t})))$ or $\ln(1/y-1/u) = \ln(b_0) + (\ln(b_1) * t)$ where u is the upper boundary value. After selecting Logistic, specify the upper boundary value to use in the regression equation. The value must be a positive number that is greater than the largest dependent variable value.

Growth Model

Model whose equation is $Y = e^{*(b_0 + (b_1 * t))}$ or $\ln(Y) = b_0 + (b_1 * t)$.

Exponential Model

Model whose equation is $Y = b_0 * (e^{*(b_1 * t)})$ or $\ln(Y) = \ln(b_0) + (b_1 * t)$.

The annual numbers of HIV patients will be plotted in y-axis against the corresponding year in the x-axis. Curve fitting, also known as regression analysis, will be used to find the "best fit" line or curve for a series of data points. F-test should be used for selecting the best fitting curve for the testing of hypothesis. P-value must be taken as significant when < 0.05 (two-tailed). R^2 value > 0.80 should be taken as significantly better for prediction. The decision regarding the selection of a suitable prediction approach is governed by the relative performance of the models for monitoring and prediction. It should also adequately interpret the phenomenon under study.

Conclusion

This paper has reviewed a novel method of using the curve fitting method in HIV/AIDS data. The approach is simple to understand and apply, and is capable of

curve fitting a whole range of different models. It also has the advantage that several different models, for a given data series, can be easily investigated, thus easing the model selection dilemma.

References

1. HIV declining in India; New infections reduced by 50% from 2000-2009; Sustained focus on prevention required. Government of India Ministry of Health & Family Welfare Department of AIDS Control National AIDS Control Organisation. Online [2010] Accessed [2012]. Available from: <http://nacoonline.org/upload/HomePage/NACO%20Press%20Release%20on%20HIV%20Estimates.pdf>.
2. Anderson RM. The role of mathematical models in the study of HIV transmission and the epidemiology of AIDS. *AIDS* 1988; 1: 241-246.
3. Hyman JM and Stanley EA. Using mathematical models to understand the AIDS epidemic. *Math. Biosciences* 1988; 90: 415-474.
4. Jager JC and Ruitenberg EJ. *Statistical Analysis and Mathematical Modeling of AIDS*, Oxford University Press, Oxford, 1988.
5. Wilkie AD. An actuarial model for AIDS. *Journal of Royal Statistical Society*, 1988; Series A, 151: 35-39.
6. Hethcote HW, Van Ark JW and Longini IM. A simulation model of AIDS in San Francisco: I. Model formulation and parameter estimation. *Math. Biosciences* 1991; 106: 203-222.
7. Anderson RM and May RM. Understanding the AIDS epidemic. *Scientific Amer* 1992; 266: 58-66.
8. Mode CJ, Gollwitzer HE and Hermann N. A methodological study of a stochastic model of an AIDS epidemic. *Math. Biosciences* 1988; 92: 201-229.
9. Isham V. Assessing the variability of stochastic epidemic. *Math. Biosciences* 1991; 107: 209-224.
10. Tan WY and Xiang ZH. A state space model of HIV pathogenesis under treatment by anti-viral drugs in HIV infected individuals. *Math. Biosciences* 1999; 156: 69-94.
11. Tan WY and Xiang ZH. State Space Models for the HIV pathogenesis. In *Mathematical Models in Medicine and Health Science*. (Eds: Horn, M.A., Simonett, G. and Webb, G.), Vanderbilt University Press, Nashville, TN, 1998. 351-368.
12. Jewell NP, Dietz K and Farewell VT. *AIDS Epidemiology: Methodological issues*. Birkhauser, Basel. 1992).
13. Bacchetti P, Segal M and Jewell NP. Backcalculation of HIV infection rates, *Statistical Science* 1993; 8: 82-119.
14. Brookmeyer R and Gail MH. *AIDS epidemiology:*

- A Quantitative Approach. Oxford University Press, Oxford. 1994.
15. Wu H and Tan WY. Modeling the HIV epidemic: A state space approach. In: "ASA 1995 Proc- the Epidemiology Section". ASA, Alexandria, VA: 1995,66-71.
16. Kalman RE . A new approach to linear filter and prediction problems. J Basic Eng 1960; 82: 35-45.
17. Cazelles B and Chau NP. Using the Kalman filter and dynamic models to assess the changing HIV/AIDS epidemic. Math. Bioscience 1997; 140:131-154.
18. Healy MJR and Tillett HE. Short-term extrapolation of the AIDS epidemic. J Royal Stat Soc, Series A 1988; 151: 50-61
19. Anderson RM, Medley GF, May RM and Johnson AM. A preliminary study of the transmission dynamics of the human immunodeficiency (HIV), the causative agent of AIDS, IMA J Math Appl Med and Biol 1986; 3: 229-263.
20. Isham V. Mathematical modeling of the transmission dynamics of HIV infection and AIDS: A review. J Royal Stat Soc.1988; 151: 5-30.
21. Brookmeyer R and Gail MH. Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States. Lancet 1986; 2: 1320-1322.
22. Brookmeyer R and Gail MH. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. J Amer Stat Asso 1988; 83: 301-308.
23. Sathian B, Sreedharan J, Mittal A, Baboo NS, Chandrasekharan N, Devkota S, Abhilash ES, Rajesh E, Dixit SB. Statistical Modelling and Forecasting of Reported HIV Cases in Nepal. Nepal Journal of Epidemiology 2011;1(3): 106-110.
24. Sathian B. Statistical Modelling of HIV/AIDS in Nepal: A Necessary Enquiry. Nepal Journal of Epidemiology 2011;1(3):74-76.
25. Sathian B, Bhatt CR, Jayadevan S, Ninan J, Baboo NS, Sandeep G. Prediction of cancer cases for a hospital in Nepal: a statistical modelling. Asian Pac J Cancer Prev 2010;11 (2): 441-5.
26. Sathian B, Sreedharan J, Chandrasekharan N, Devkota S, Rajesh E, Mittal A. Statistical modelling in the prediction of kala-azar in Nepal. Journal of Epidemiology and Community Health. 2011: 65.
27. Sathian B, Sreedharan J, Sharan K, Baboo NS, Chawla R, Chandrasekharan N, et al. Statistical Modelling Technique in Forecasting of Palliative Oncotherapy Load in Hospitals. Nepal Journal of Epidemiology. 2010;1 (1): 38-43.
28. Sathian B, Sreedharan J, Sharan K, Baboo NS, Ninan J, Joy T, Abhilash ES. Forecasting Breast Cancer Cases requiring Radiotherapy at a Teaching Hospital in Nepal. Journal of Clinical and Diagnostic Research. 2010; 4: 2378-83.

Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.