Trend Analysis of HIV Infection Rates amongst Generation X Black African Women in South Africa during the Period 2001 to 2010

Wilbert Sibanda*, Philip D. Pretorius
School of Information Technology, North-West University, Vaal Triangle Campus, Hendrik van Eck Boulevard, Vanderbijlpark, South Africa 1900.

* Corresponding author. Tel.: +27 16910 3038; email: Wilbert.sibanda@nwu.ac.za
Manuscript submitted February 25, 2015; accepted May 18, 2015.
doi: 10.17706/ijbbb.2015.5.4.221-231

Abstract: This study investigated the changes in HIV and syphilis prevalence rates amongst Generation (Gen) X black African women attending antenatal clinics across the nine provinces of the Republic of South Africa during the years 2001 to 2010. Gen X women were defined to be individuals born during the period 1961 to 1981. In South Africa, the Gen X period was characterized by numerous social protests, academic boycotts and Soweto student uprisings of 1976. A democratic government was elected in 1994, resulting in major changes in social and economic circumstances of most South Africans. The new government promoted free education resulting in a major change in the demographics of the country's public service. Amongst the new working and middle class, sociological literature reports a remarkable increase in materialism and consumerism. However, the social privileges of a few existed hand-in-glove with unprecedented levels of youth unemployment. Youths from impoverished backgrounds found it increasingly difficult to attain a descent education. Most of these youths grew in homes characterized by physical violence and emotional abuse. In view of the widely held tenet that a generation is one of the fundamental social classifications in a society, this research aims to examine trends in HIV and Syphilis amongst Generation X black African women attending antenatal clinics in South Africa.

Key words: Demographic characteristics, Gen X, HIV.

1. Introduction

In South Africa, annual antenatal HIV seroprevalence surveys are the main indicators of changes in HIV and syphilis infection rates. Antenatal surveys were started in 1990 and remain the most comprehensive HIV surveys [1]. The survey is an important tool for tracking the spatial and temporal trends of the epidemic amongst pregnant women and using sophisticated algorithms the trend is extrapolated to the entire population [2]. It is widely believed that the antenatal clinic data closely approximates HIV prevalence rates in the adult population [3]. However, the exact relationship between HIV prevalence rates amongst antenatal clinic attendees and that of the general population remains uncertain [4]. Numerous demographic characteristics are captured for each pregnant woman attending an antenatal clinic in South Africa, such as the pregnant women's age, educational level, gravidity (number of pregnancies), parity (number of children), male sexual partners' age, race, syphilis and HIV status. This study aims to analyze the trend in HIV and syphilis prevalence rates amongst Gen X black African women attending antenatal clinics

annually during the period 2001 to 2010. This study does not aim to study the effect of antiretroviral treatment.

2. Literature Review

Generation is defined as a cohort of individuals within a selected population who experience the same significant events over a period of time [5]. A generation is therefore one of the most important social classifications in a society [6]. It therefore refers to an entire group of human beings born and living at almost the same age, exhibiting similar challenges and attitudes. Therefore, the three fundamental aspects of a generation are shared temporal, historical and socio-cultural locations [7]. However, this research recognizes that there are numerous variations in attitudes, values, behavior and lifestyles within a given generation.

By definition, Gen Xers are people who were born after the Western post World War II baby-boomers, during the period 1961 to 1981 [8]. In the South African context, Gen Xers are individuals that were born during a period characterized by angry social protests, academic boycotts and Soweto student uprisings of 1976 [9]. The political change in South Africa in 1994 heralded a new political dispensation characterized by major changes in social and economic circumstances of most South Africans [10]. Formerly disadvantaged black Africans were admitted to institutions of higher learning, resulting in a rapid transformation in the racial demographics of senior public service within the country [11]. Amongst the new working and middle class, sociological literature reports a gradual increase in materialism and consumerism [12]. However, the social privileges of a few existed hand-in-glove with unprecedented levels of youth unemployment [13]. Youths from poor neighborhoods faced huge challenges in attaining a descent education. An increasing number of youths grew up in homes characterized by physical violence and emotional abuse. In view of the widely held tenet that a generation is one of the fundamental social classifications in a society, this research aims to describe HIV and syphilis trends amongst women attending antenatal clinics over the 10 years from the year 2001 to 2010.

3. Research Methodology

3.1. Data Sources

This study uses South African annual antenatal HIV seroprevalence data from year 2001 to 2010. The national antenatal sentinel HIV and Syphilis prevalence survey is conducted in all the nine provinces of the Republic of South Africa. HIV testing is conducted as per standard requirement from the World Health Organization (WHO). The South African antenatal survey is aimed at pregnant women between the ages of 15 and 49. In order to reduce the likelihood of individuals being included in the survey more than once, only pregnant women attending clinics for the first time were interviewed. The survey has grown over the years from an annual average of 16 000 in 2001 to 34 000 in 2010. This survey covers all the 52 health districts in all the nine provinces of the Republic.

3.2. Data Analysis

3.2.1. Gen X demographic characteristics

As stated in the introduction, the study considered Gen X to be the period from the year 1961 to 1981. However, the available data for the study was for the period 2001 to 2010. On that basis, considering the year 2001, it means that the youngest Gen X individual is 20 years, while the oldest is 40 years old. This therefore means that in the year 2001, pregnant women between the ages of 20 and 40 who attended antenatal clinics were considered to be Gen X. The same approach was applied to the subsequent years of 2002 to 2010.

3.2.2. Syphilis and HIV trend, 2001-2010

This part of the research involved producing tables and plots of observed data along with their confidence limits. The study also involved assessing whether there were extreme observations and outliers in the observed data. A log transformation of the data was also conducted, in order to compare it with 3-year moving averages (MAs). In general, 3-year moving averages (MA) increase the stability of observed data with minimal loss of changes in HIV prevalence rates for the observed and untransformed data. Average annual percentage changes (AAPC) in syphilis and HIV prevalence rates were used to compute projected prevalence rates and their confidence intervals using linear extrapolation method.

3.2.3. Comparative trend analysis, 2001-2010

A simple linear regression of the response on time was used to compare the time series trends of HIV and syphilis among Gen X black African women over the study period. All analyses were conducted using SAS 8.1 statistical software package (SAS Institute, Cary, NC).

4. Results

4.1. Black African Gen X HIV Trend, 2001-2010

The observed HIV prevalence rates amongst Gen X black African women, increase gradually from 285.8 per 1000 pregnant women (28.58%) in 2001 to about 436.95 per 1000 women (43.70%) in 2010, as shown in Table 1 and Fig. 1.

Table 1. Black African Gen X HIV Prevalence Rates 2001-2010

Year	HIV Rate/1000	95% Confidence Interval
2001	285.8	274.17-297.43
2002	328.57	316.11-339.80
2003	320.87	308.55-353.61
2004	371.04	357.79-384.27
2005	393.5	379.85-409.80
2006	389.98	376.40-400.50
2007	394.07	380.42-405.04
2008	403.37	389.56-415.22
2009	417	402.95-429.98
2010	436.95	422.58-451.55

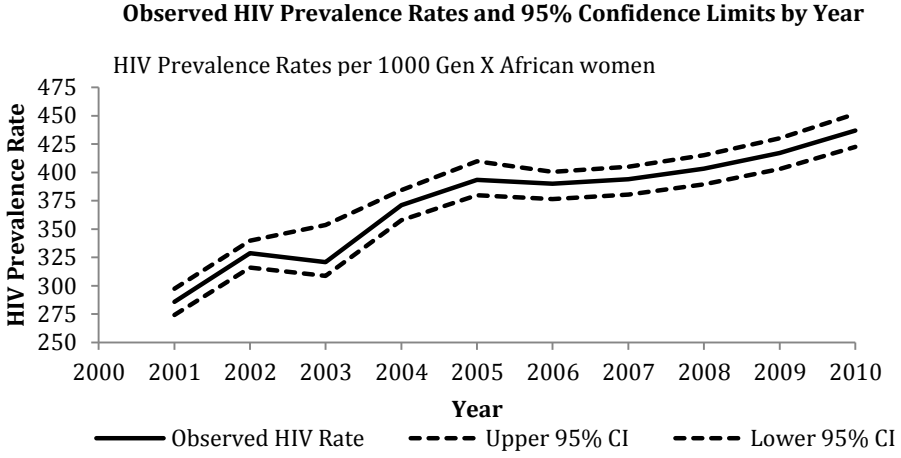

Fig. 1. Observed HIV data trend.

Even though there is a drop in HIV rate during the years 2002-2003 and 2005-2006, the overall HIV prevalence rates appeared to gradually increase towards the year 2010. However, the confidence intervals around each HIV rate were observed to be narrow.

4.1.1. Data transformation and smoothing

Logarithmic transformation was applied on the observed HIV prevalence data. The advantage of the log transformation is that the overall shape of the plot is unchanged, while decreasing the jagged nature of the curve. Based on the log transformation of data, HIV prevalence rates were observed to increase from the year 2001 to 2010, as in Fig. 2.

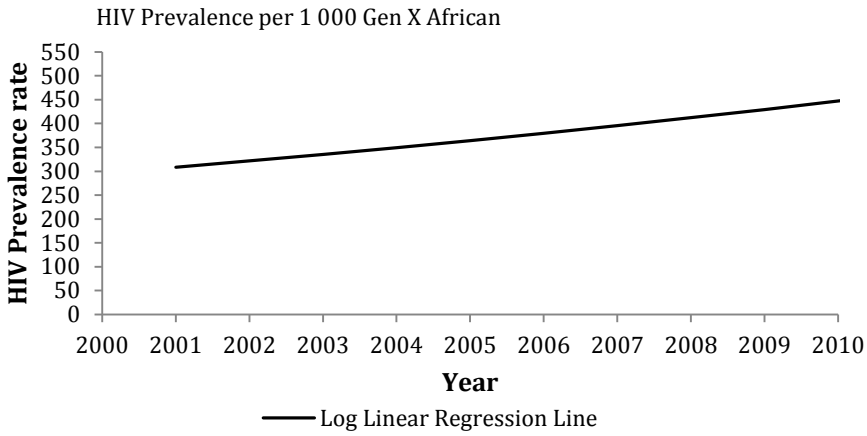

Fig. 2. Log-linear HIV trend.

In addition to logarithmic transformation, a 3-year overlapping MA smoothing technique was used to study the HIV trend amongst the Gen X black African women. The 3-year MA technique demonstrated a progressive and sustained increase in HIV prevalence rates from 300 infections per 1 000 women (30%) in 2001 to about 400 infections per 1 000 women (40%) in 2010, as shown in Fig. 3.

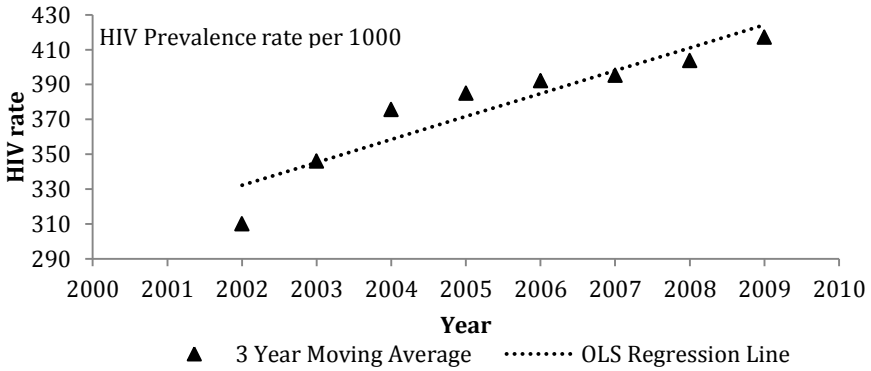

Fig. 3. HIV rate based on 3-year MA and OLS linear regression model.

4.1.2. Estimation of average annual percent change (AAPC)

AAPC estimates year to year HIV rate changes, instead of computing the overall change from the start to the end of the study period. This estimate provides a wealth of information on how fast the HIV prevalence rate changes over time. The AAPC indicator was calculated using the formula in Equation (1).

$$AAPC = \frac{\sum_{i=1}^{n} \left(\frac{HIV\ rate_i}{HIV\ rate_{(i-1)}}\right) \times 100}{n-1} \quad (1)$$

where *n* is the total number of years, Year 2001 is the first and Year 2010 is the final year.

Using HIV rates on untransformed data with no smoothing, the average of the year to year changes was found to be 4.499% calculated as shown in Table 2.

Table 2. AAPC HIV Rates for Untransformed and 3-Year MA Data

Method	Variable Name	Average annual % HIV rate
Untransformed HIV rate	HIV rate	4.499
Three overlapping MA HIV rate	MA HIV rate	3.85

$$\frac{\left[\left(\frac{328.57}{285.80}-1\right)+\left(\frac{320.87}{328.57}-1\right)+\ldots+\left(\frac{429.98}{415.22}-1\right)+\left(\frac{451.55}{429.55}-1\right)\right]\times 100}{10} = 4.499\% \quad (2)$$

Equation (2) above, illustrates how the average year to year changes were computed. The obtained year to year value of 4.499% implies that the HIV prevalence rates are increasing over time. However, using eight 3-year MA approach, the AAPC was found to be slightly lower at 3.85% as shown in Equation (3) below. The AAPC value derived from expression (3) was considered to be the most accurate estimate of average annual percent change in HIV prevalence rate amongst Gen X black African women.

$$\frac{\left[\left(\frac{346.03}{310.10}-1\right)+\left(\frac{375.72}{346.03}-1\right)+\ldots+\left(\frac{403.75}{395.34}-1\right)+\left(\frac{417.24}{403.24}-1\right)\right]\times 100}{8} = 3.85\% \quad (3)$$

4.1.3. HIV prevalence rate projection

The 3-year MA AAPC value was used to compute the projected HIV prevalence rates and their confidence limits beyond the year 2010, through the use of linear extrapolation as shown in expressions (4) and (5) respectively.

$$Projected\ HIV\ rate_t = 3yearMA_{t-1} + 0.0385 \times 3yearMA_{t-1} \quad (4)$$

$$Projected\ HIV\ rate = \pm 1.96\sqrt{\frac{Projected\ HIV\ rate \times 1000}{Projected\ HIV\ uninfected\ rate}} \quad (5)$$

The projected HIV prevalence rates computed using linear extrapolation of AAPC values demonstrated a gradual increase in HIV rate amongst Gen X black African women from the year 2010 to 2020, as shown in Table 3 and Fig. 4.

The linear extrapolation HIV projection method was compared with an OLS and log linear models. The OLS regression methodology predicted higher HIV prevalence rates than the linear extrapolation model.

However, the log linear model provided HIV prediction values considerably higher than both the OLS and linear extrapolation approaches. The disparity between the predicted rates based on linear and log linear

models was observed to increase as the time periods being forecasted moved further away from the time of observation as shown in Fig. 5.

Table 3. Projected HIV Prevalence Rates

Years	Projected MA HIV rate	95% Confidence interval
2009, 2010, 2011	433.29	379.28-487.30
2010,2011,2012	449.96	394.34-505.57
2011,2012,2013	467.26	409.99-524.54
2012,2013,2014	485.24	426.26-544.21
2013,2014,2015	503.9	443.17-564.64
2014,2015,2016	523.29	460.75-585.83
2015,2016,2017	543.42	479.02-607.82
2016,2017,2018	564.32	498.00-630.64
2017,2018,2019	584.03	517.74-654.32
2018,2019,2020	608.57	538.25-678.90

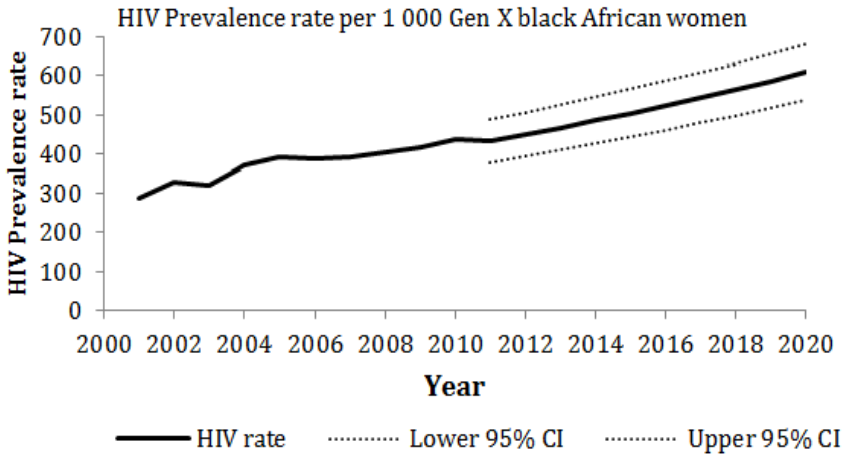

Fig. 4. HIV prevalence rates with confidence intervals predicted by linear extrapolation of AAPC values.

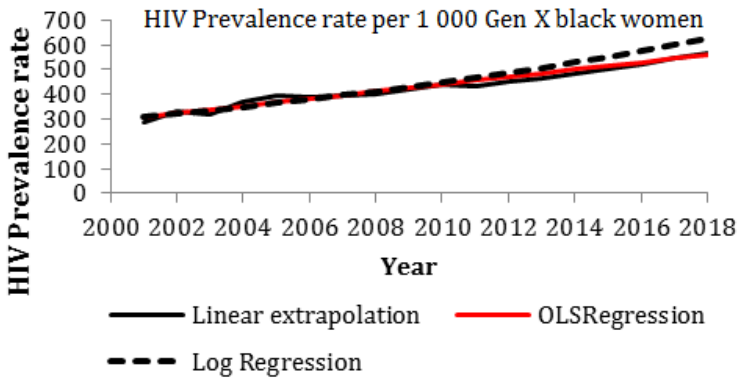

Fig. 5. Projection of HIV prevalence rates using linear extrapolation, OLS and log linear regression models.

5. Time Series Analysis (TSA)

In general time, series analysis (TSA) provide more precise results compared to other regression methodologies. Unlike OLS and log linear models which assume that errors in the modeled observations are uncorrelated, TSA assumes that the errors are correlated. However, TSA tends to generate wider confidence bands than other regression models.

The advantages of TSA series methods include the ability to precisely characterize a trend through identification and adjustment for non-linear components of the trend and ability to differentially weight observations for forecasting. For this research, we used a non-stationary ARIMA (0, 1, 0) model, also called integrated I (1) model. In general, stationary TSA assumes that the mean, variance, autocorrelation are constant over time, which suggests that both the expected values of the series and its auto-covariance function are independent of time. ARIMA (0, 1, 0) is the first difference trend where the first difference is considered to be the changes from one year to the next. The autoregressive (p) and moving averages (q) parameters are considered not to be significant.

$$y_t = \delta + y_{t-1} + \varepsilon_t \tag{6}$$

where δ is the difference and ε is the error term.

The integrated I (1) model predicts a change in HIV prevalence rate as an average change, with some fraction of previous change, a random error, and some fraction of the random error in the preceding period added (6).

The ARIMA (0, 1, 0) model predicted a gradual and sustained increase in HIV prevalence rates during the prediction period during the years 2010 and 2020. However, the confidence intervals were observed to widen over the prediction period, indicating an increase in uncertainty of the predicted HIV values further away from the observed data points (Fig. 6).

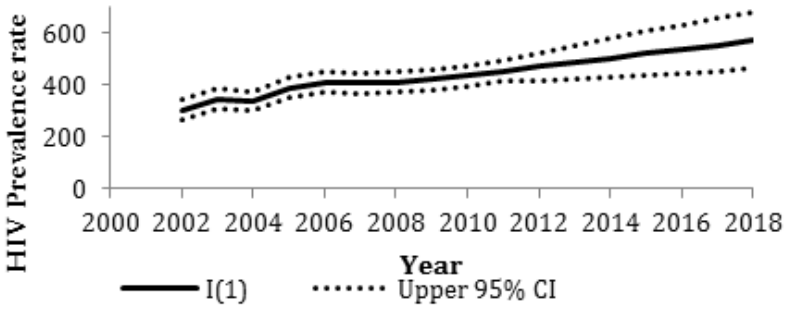

Fig. 6. Prediction of HIV prevalence rates using an ARIMA (0, 1, 0) or I (0) model with 95% CI.

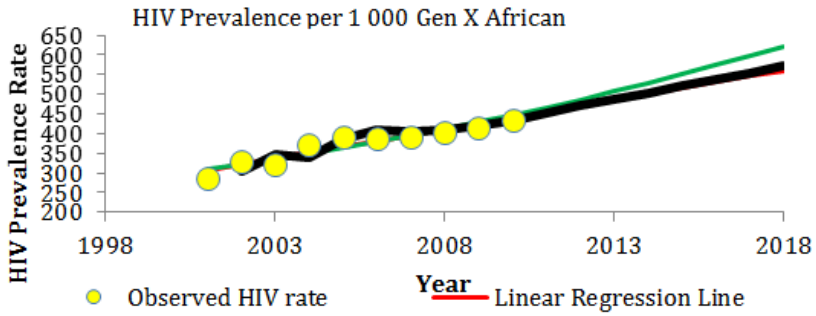

Fig. 7. Comparison of HIV predictions using ARIMA (0,1,0), linear extrapolation and log linear models.

The ARIMA (0, 1, 0) integrated model compared well with the linear model as shown in Fig. 7. However, the log-linear model predicted HIV prevalence rates considerably higher than both the linear and the integrated ARIMA models. However, the three model predictions suggest a disturbing gradual increase in HIV infection rates amongst Gen X black African women during the period of 2011 to 2018.

5.1. Syphilis in Relation to HIV Status, 2001-2010

In general positivity rates of syphilis among Gen X black Africa women were observed to gradually decrease from the year 2001 to 2006, as shown in Table 4. However a spike of 28.68 infections per 1 000 women was observed in 2007, after which the prevalence rates dropped to 17.59 infections per 1 000 women in 2008. Another, albeit smaller spike of 21.67 infections per 1 000 women was observed in 2009, followed by a decrease in syphilis rate to 15.86 infections per 1 000 women in 2010 (Fig. 8). In comparison, the syphilis rates were considerably lower than HIV rates over the same period with mean rates of 24.14 (95% CI, 19.69-28.59) and 374.12 (95%CI, 344.62-403.62) infections per 1 000 women respectively.

Table 4. HIV and Syphilis Rates, 2001-2010

Year	Syphilis Rate/1 000	HIV Rate/1 000
2001	33.74	285.8
2002	36.74	328.57
2003	23.48	320.87
2004	25.03	371.04
2005	22.5	393.5
2006	17.59	389.98
2007	28.68	394.07
2008	17.59	403.37
2009	21.67	417
2010	14.35	436.95

In order to deal with the jagged nature of the syphilis trend, a 3-year overlapping MA graph was generated as shown in Fig. 9.

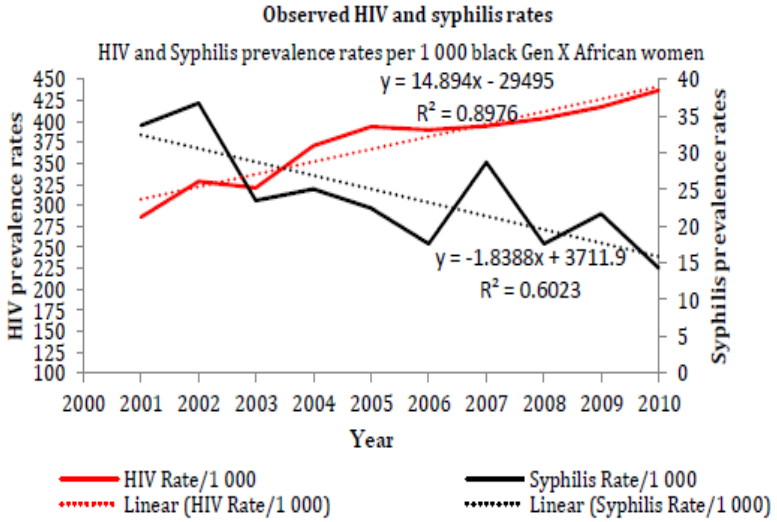

Fig. 8. Observed HIV and syphilis rates amongst Gen X black African women.

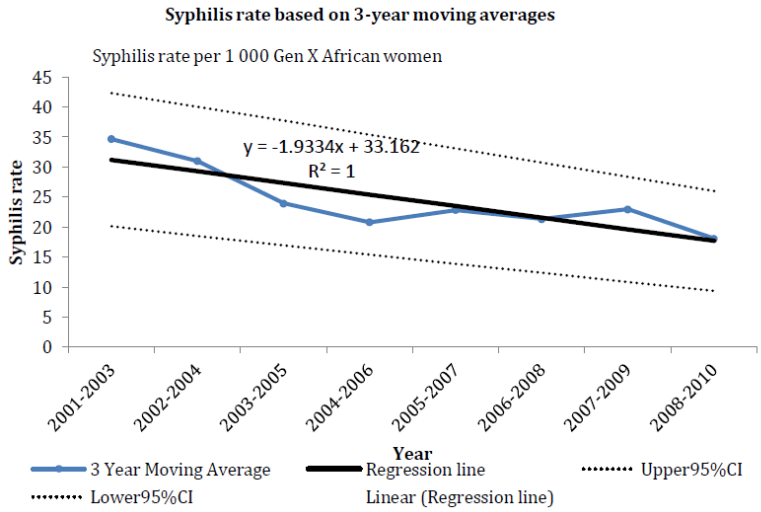

Fig. 9. Syphilis rates based on 3-year MA and a linear regression model.

The 3-year MA plot demonstrated that there was a general decrease in syphilis prevalence rates from the year 2001 to 2006. However, the 3-year average of the years 2005 to 2007, 2006 to 2008 and 2007 to 2009 demonstrated an increase in syphilis prevalence rate. The last 3-year average of the years 2008 to 2010 showed a drop in syphilis prevalence rates to 18.055 infections per 1 000 women.

6. Discussion

This research clearly shows that there was a gradual and sustained increase in HIV infection rates from the year 2001 to 2010. Furthermore, following many years of decreasing syphilis rates, the spikes observed in 2007 and 2009 were concerning in view of the fact that syphilis and other sexually transmitted diseases (STDs) tend to increase the risk of acquiring as well as transmitting HIV infection.

Even though overall, there is a decline in syphilis prevalence rates from the year 2001 to 2010, the 3-year MA model employed in this study indicates an increase in syphilis infection rates from 2006 to 2010. The above research findings demonstrate that despite the numerous ongoing sexual educational campaigns in South Africa, people still engage in unprotected sexual practices. Some researchers have linked the increase in STDs to the increased availability of anti-retroviral treatments (ARVs). This suggests that people on ARVs are engaging in high risk behavior, under the assumption that there would be treatment for HIV [14].

An independent study based on Amsterdam Cohorts demonstrated that anti-retroviral treatment was associated with increased levels of risk behaviors amongst HIV-infected and HIV-uninfected homosexual men [15]. Another concerned finding of this research was the fact that HIV infection rates were predicted to increase from 2010 to 2020. The three models employed in this study demonstrated a sustained increase in HIV prevalence from the year 2001 to 2010 and beyond. The above findings emphasize the need to improve HIV campaigns within the communities in order to curb the spread of the epidemic and associated STDs.

References

[1] National Department of Health. (2012). *The 2012 National Antenatal Sentinel HIV and Herpes Simplex Type-2 Prevalence Survey, South Africa.*

[2] Shisana, O. (2005). *South African National HIV Prevalence, HIV Incidence, Behaviour and Communication Survey*. HSRC Press.

[3] Gouws, E., Mishra, V., & Fowler, T. B. (2008). Comparison of adult HIV prevalence from national population-based surveys and antenatal clinic surveillance in countries with generalised epidemics:

Implications for calibrating surveillance data. *Sexually Transmitted Infections*, *84(Suppl 1)*, 17-23.

[4] Salomon, J. A., & Murray, C. J. L. (2001). Modelling HIV/AIDS epidemic in sub-Saharan Africa using seroprevalence data from antenatal clinics. *Bull World Health Organ, 79(7),* 596-607.

[5] Pilcher, J. (1994). Mannheim's sociology of generations: An undervalued legacy. *British Journal of Sociology, 45(3)*, *481-495.*

[6] Marmot, M., Friel, S., Bell, R., Houweling, T. A., Taylor, S., & Commission on social determinants of health. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *The Lancet, 372(9650), 1661-1669*.

[7] Mannheim, K. (1952). The problem of generations*. Mannheim's Essays on the Sociology of Knowledge,* 276-322.

[8] Strauss, W., & Howe, N. (1991). *Generations*. New York: Harper Perennial.

[9] Sibanda, W., & Pretorius, P. D. (2014). Trend analysis of HIV prevalence rates amongst gen X and Y pregnant women attending antenatal clinics in South Africa between 2001 and 2010. *Mediterranean Journal of Social Sciences, 5 (21), 93*.

[10] Dawes, A., & Finchilescu, G. (2002). What's changed? The racial orientations of South African adolescents during rapid political change. *Childhood, 9(2), 147-165*.

[11] Akoojee, S., & Nkomo, M. (2007). Access and quality in South African higher education: The twin challenges of transformation. *South African Journal of Higher Education*, *21(3)*, 385-399.

[12] Schenk J., & Seekings J. (2010). *Locating Generation X: Taste and Identity in Transitional South Africa,* 1-24. South Africa: Centre for Social Science Research (CSSR), University of Capetown*.*

[13] Jacobs, G., & Smit, E. (2010). Materialism and indebtedness of low income consumers: Evidence from South Africa's largest credit granting catalogue retailer. *South African Journal of Business Management*, *41(4), 11-33.*

[14] Rietmeijer, C. A., Patnaik, J. L., Judson, F. N., & Douglas Jr, J. M. (2003). Increases in gonorrhea and sexual risk behaviors among men who have sex with men: A 12-year trend analysis at the Denver metro health clinic. *Sexually Transmitted Diseases, 30(7)*, *562-567*.

[15] Dukers, N. H., Goudsmit, J., de Wit, J. B., Prins, M., Weverling, G. J., & Coutinho, R. A. (2001). Sexual risk behavior relates to the virological and immunological improvements during highly active antiretroviral therapy in HIV-1 infection. *AIDS, 15*, 369-378*.*

Wilbert Sibanda is an associate member of the pharmaceutical society of South Africa. He has held the following qualifications, BSc (human physiology) (University of the Witwatersrand, Johannesburg), BSc (med) Hons (pharmacology) (University of Capetown), MSc (med) (pharmacy) (University of the Witwatersrand, Johannesburg), PhD (statistical modeling of antenatal HIV data) (North-West University, South Africa). He is involved in statistical modeling of HIV antenatal data in South Africa.

He is a subject specialist (bioinformatics) at the School of Information Technology, North-West University, Vaal Triangle campus, Hendrik van Eck Boulevard, Vanderbijlpark, South Africa. He has published extensively in the field of statistical modeling of HIV antenatal data.

He has chaired a special session at the International Computers and Industrial Engineering (CIE 42) Conference in Capetown, South Africa in 2012. He also chaired the machine learning and databases session at the 2013 IEEE international conference on bioinformatics and biomedicine, at Tongji University, Shanghai, China. He is a reviewer of the *International Journal of Computational Biology, Informatics and Control* (IJCBIC). He is also a technical committee member of the 4th International Conference on Bioinformatics and Biomedical Science (ICBBS) to be held on the 25th to the 25th of June 2015, in Bangkok, Thailand.

Philip D. Pretorius is a researcher in the field of data mining field, applying data mining and statistical methods on various big data sets. He is a promoter and study leader for postgraduate students of industry and academia in the School of IT, and co-ordinate the postgraduate research in the School of IT. He was director of the School of IT till 2014, and previously lectured in statistics and decision support systems.