Far East Journal of Mathematical Sciences (FJMS)
Volume 29, Issue 2, 2008, Pages 341-352
Published online: June 30, 2008
This paper is available online at http://www.pphmj.com

ON WHO AND AIDS DEATHS IN INDIA

B. D. AGGARWALA

Department of Mathematics and Statistics
University of Calgary
Calgary, Alberta, Canada
e-mail: aggarwal@ucalgary.ca

Abstract

We discuss the number of AIDS deaths in India in recent years and show that they should be considerably less than the ones estimated by World Health Organisation (WHO).

1. Introduction

According to UNAIDS, the number of AIDS deaths in India was estimated at 380,000 in 2003 and 475,000 in 2005 [9]. During these same years, the number of HIV positive people was estimated at 5.3 million in 2003 and 5.7 million in 2005 [9]. In a paper published in 2006 [2], this writer argued that these estimates of HIV positive people in India were too high and a more realistic estimate would be 2.5 million at the end 2004. While the number of HIV positive people in India was revised downwards to 2 to 3 million after a careful survey in 2006 and this revision was reported in the newspapers [8], there was no such revision of the estimate of the number of AIDS deaths. These estimates are important to know the extent of HIV/AIDS in India and to guide health care policy in that country. In this paper, we shall attempt to provide these estimates.

2000 Mathematics Subject Classification: 92D25.

Keywords and phrases: HIV, AIDS, mathematical models, AIDS deaths.

Received December 4, 2007; Revised January 17, 2008

2. The EPI Model

The EPI Model [6] was developed by Prof. James Chin of the University of California in Berkeley in early nineties and is extensively used by WHO in arriving at these estimates. Prof. Chin headed the epidemiology section of WHO for a number of years in the early nineties. The model depends upon assuming that the HIV prevalence in any country follows a Gamma curve of the type given in Figure l.

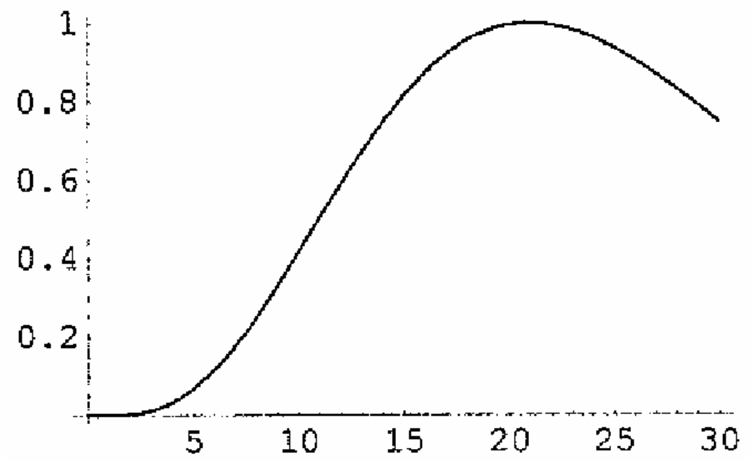

Figure 1. Sketch of $F[t] = At^p Exp[-t/q]$ for $p = 4$ with the peak at $t = 21$.

In EPI model, the beginning of the curve is put at an year when 'significant' prevalence has already developed. This placing does not affect the results very much because in the beginning, the number of hiv positive people is very small any way. The current year, when hiv prevalence is available (through some field estimates), is generally put at or near the peak of the curve (twenty one in the above case) and then the number of AIDS deaths is estimated by known prevalence and by assuming that a certain fraction of people infected in any year i develop AIDS and then die in a later year j. For details, we refer the reader to [6].

We shall assume that patients progress from initial hiv infection to AIDS in about eight years according to a Weibull distribution. The Weibull distribution is gives by

$$F[t, \lambda] = (1 - Exp[-(\lambda t)^{\theta}])$$

and different groups are likely to follow this progression (from seroconversion to AIDS) according to different values of θ and λ. Here are two examples.

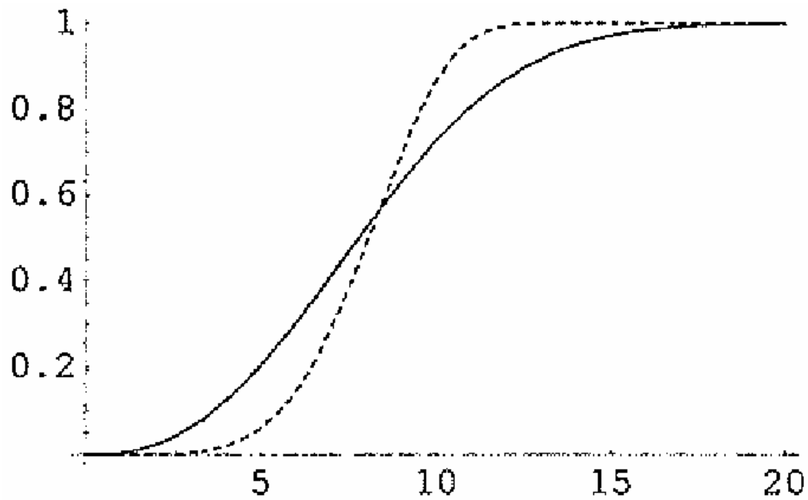

Figure 2. Graphs of the Weibull Distributions for $\lambda = .111$ and $\theta = 2.5$ (solid line) and for $\lambda = .115$ and $\theta = 5$. Both give the median time of development to AIDS as about eight years. In the beginning, progression to AIDS is faster for lower values of θ.

The first few values of $F[t]$ (= $F[t, \lambda, \theta]$ with $\lambda = .111$ and $\theta = 2.5$) are given by

$$\{F[0], F[2], F[4], F[6], F[8], F[10], F[12]\}$$

$$= \{0.0, 0.0229535, 0.123096, 0.303704, 0.52431, 0.726949, .870965\}.$$

The graphs in this figure give the cumulative probability distribution of development to AIDS. The first few given values say that, out of a hundred patients say, most stay asymptomatic for about two years (for $\lambda = .111$ and $\theta = 2.5$), after which some patients start developing AIDS, and by about twelve years, most of them have done so. Notice that $F[7.78] = .4999$, so that approximately eight years is the median interval for development to AIDS. The mean in this case turns out to be 7.99 years. It follows that if $g(s)$ is the number of people who become hiv positive at times s, and $A(t)$ is the number of patients who have developed AIDS by the time t, then the total number of patients who are diagnosed with AIDS in the time interval $(0, t)$ is given by the so called back calculation method [5]

$$A(t) = \int_0^t g(s) F(t - s) ds, \tag{1}$$

where $F(t - s)$ is the probability that a person has developed AIDS in $(t - s)$ years or less. It should be noticed [1] that just as the time to development of AIDS may be distributed as a Weibull distribution, so may the time from onset of AIDS to death by a different Weibull distribution $G(t)$. We shall assume that all people who develop AIDS subsequently die. In a previous paper [1], we have shown that $AD(t)$, the number of people who die of AIDS in the time interval $(0, t)$ is given by

$$AD(t) = \int_0^t g(s) k(t - s) ds, \tag{2}$$

where

$$k(t) = \int_0^1 tG((1 - w)t) F'(wt) dw. \tag{3}$$

It follows that $Pr(t)$, the hiv prevalence at any time t, is given by

$$Pr(t) = \int_0^t g(s)(1 - k(t - s)) ds. \tag{4}$$

If we also assume that the hiv prevalence is given by some curve like in Figure 1, then $g(s)$, the infection rate, may be determined by back calculation. It should be noticed that in this case, the values of $g(s)$ towards the end $(s = t)$ have more weight in the back calculation process than the ones in the beginning $(s = 0)$, so that this process should give more accurate values of $g(s)$ towards the end than towards the beginning. However, in the beginning, they are very small anyway, and do not affect our results in a serious way. This is in contrast to the usual application of back calculation method based on equation (1) where the values of $g(s)$ towards the end have very small weight and therefore are hard to evaluate.

The default value of p in the curve in Figure 1 is usually taken to be four [6]. This is because the Gamma curve with this value has fitted the reports from many countries where such reporting is reliable. However, the value of p may well be different in different societies. As an example, if the patients develop AIDS fast (MSM, IDU), then $p = 3$ may be more appropriate. The most important parameter to determine now is q which determines the location of the peak at $t = pq$. As we shall see, different values of q will give vastly different estimates of AIDS deaths. We would argue that the peak will occur at different points of time in different societies depending upon a variety of factors. The remaining parameter A may be determined by knowing hiv prevalence at any one point on the curve, 2.5 million in 2006 in India. We shall now give details for the case of India.

3. Application to India

The first case of hiv positivity in India was diagnosed in a Chennai hospital in India in 1982, and the first case of AIDS was diagnosed in 1986. In a large country like India, it would be unrealistic to think that hiv was detected after only a few months of its occurrence anywhere in the country even though in the beginning, hiv positivity is asymptomatic. We shall therefore assume that the first case of hiv positivity occurred in India in 1981 and that hiv was 'significantly' present in 1986.

We shall assume that, on the average, hiv progresses to AIDS in India in a span of about eight years and that AIDS results in death in approximately one year. These numbers are appropriate for a developing country like India [6, p. 117]. Consequently, we shall take $\theta = 2.5$, $\lambda = .111$ for F and $\theta = 2.5$, $\lambda = .89$ for G, thus giving the mean of 7.9934 years from seroconversion to onset of AIDS and .9969 years from AIDS to death.

With 1986 as the year one in Figure 1, 2006 is the year 21. The peak in the Gamma curve occurs at $t = pq$. If we therefore take $p = 4$ and $q = 21/4$, we are putting the peak in 2006 while if we take

$$q = (21 - k)/4 \tag{5}$$

we are putting the peak k years before 2006. We feel that in a large country like India, while hiv positivity may already be widespread in high risk groups (FSW, MSM, IDU etc.), it is still slowly creeping into the population at large (husband bringing it home after a visit to a prostitute or from a foreign visit, child born to an hiv positive woman, and so on). Considering that the National AIDS Control Organisation of India (NACO) estimates of hiv prevalence in India were 3.86, 3.97, 4.58 and 5.1 million in the years 2000 to 2003, and that WHO's were perhaps even higher, one would think that the peak is yet to come. If we try to fit a gamma curve with $p = 4$ to this data, it is seen (see Figure 6 below) that the peak should be somewhere near 2011. While it is true that there were voices, this author's [2] and others' [6], which opined that these estimates were too high, and subsequently they were revised, one would think that the NACO and WHO estimates were equally wrong in all the preceding years as well and therefore, the peak is still far away.

To integrate the right hand side of equation (4), with an unknown $g(s)$, we write $g(i) = p[i]$ for $i = 0, 1, 2, 3, \ldots$ and join these $p[i]$'s by straight lines to obtain the function $g(s)$. With $k(t)$ known, we integrate the right hand side by a suitable numerical method. We used Simpson's method with sufficiently large number of subintervals. We now have the right hand side as a function of $p[i]$'s. As an example, in one particular case for $t = 10$, we get the value of $Pr[10]$, the prevalence in year ten, as

$$\begin{aligned} Pr[10] = {} & 0.2020989p[0] + 0.476123p[1] + 0.58648p[2] + 0.694146p[3] \\ & + 0.791737p[4] + 0.873037p[5] + 0.934012p[6] + 0.973457p[7] \\ & + 0.993335p[8] + 0.999364p[9] + 0.499994p[10]. \end{aligned}$$

This says that, in our model, at the end of 10 years, approximately 20.2% of the people who got infected in year zero and 47.6% of the people who got infected in year one, and so on, are still alive after year 10. The mean being 7.99 years, these percentages are reasonable. Also, because of $\theta = 2.5$, people are getting AIDS very fast in the beginning but more slowly later on as compared to a higher value of θ. The coefficient .499994 of $p[10]$ comes from the Simpson's method.

We now assumed the prevalence before 1986 in such a manner that the number of infected people in any given year was non-negative. Since this number is very small, it is not likely to affect our final results in any significant way. Since we now know the prevalence in every year, we can solve for $p[i]'$s by back calculation. We shall now give the results.

As our first example, we take $k = -10$ in equation (5) and put the peak of hiv prevalence in 2016. We take $\theta_1 = \theta_2 = 2.5$ with $\lambda_1 = .111$ and $\lambda_2 = .89$. The prevalence in 2004, 2005 and 2006 turns out to be 2.168, 2.34 and 2.5 million respectively. The cumulative incidence turns out to be 2.86, 3.21, and 3.56 in these years. The difference is the cumulative number of deaths which is 695 thousand, 865 thousand and 1.06 million. The number of deaths in these years turn out to be 148 thousand, 170 thousand and 192 thousand respectively. The following figure gives the cumulative hiv incidence against hiv prevalence in India from 1981 to 2015.

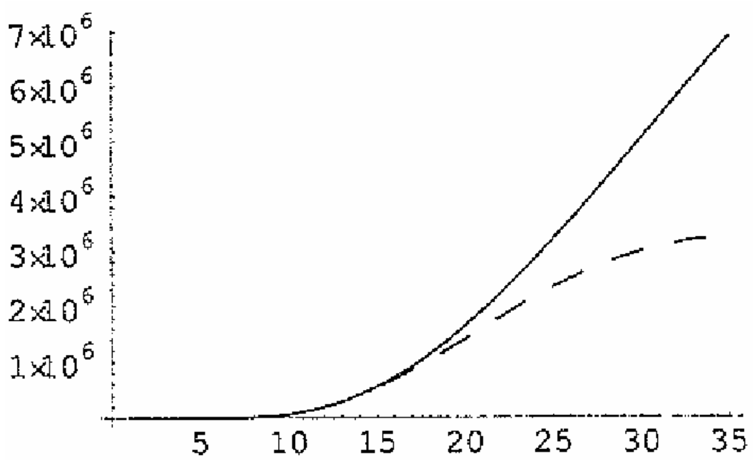

Figure 3. This graph shows cumulative incidence (solid line) against prevalence in our model from the year 1981 to 2015. The difference is the number of AIDS deaths which start increasing dramatically after 1995.

It appears from Figure 3 that the number of AIDS deaths increases dramatically as we approach and then go beyond the peak on the prevalence curve, the year 2016 in the above figure. However, the validity of numbers beyond the peak is always very contentious. Thus the number of deaths in 2006 are 192 thousand for the peak in 2016 and 332 thousand for the peak in 2001. We should point out that with the peak in 2016, the number of AIDS deaths in 2004 turns out to be 148288 which compares with our estimate of 122997 arrived at by a completely different method [3]. It also substantiates our belief that the peak in India is still far away.

Of course, if the peak is put somewhere else, then these numbers will all be different. So, where should the peak be placed? According to one source, "Where on the HIV epidemic curve should the HIV reference year be placed? This question can be answered only by analysis of all the available epidemiologic data. Such data and other observations may suggest that incidence is generally increasing or decreasing [6, p. 116]." We believe that in our case, the incidence is increasing and the reference year (2006) should be placed quite a bit before the peak. Thus for the peaks in 2008, 2011, 2016, and 2032 the number of deaths in 2006 turns out to be 227, 209, 192 and 158 thousand respectively. If therefore, we put the peak somewhere between 2011 and 2016 (a likely scenario), then the number of AIDS deaths in India in 2006 turns out to be close to 200,000. This would be our estimate of the number of AIDS deaths in India in 2006 as against the estimate of UNAIDS which is considerably higher [9].

If, on the other hand, as many epidemiologists believe, the Indian epidemics are in their post-peak phase, and we put the peak in 2001, then the number of AIDS deaths in 2006 turns out to be 332 thousand which compares with WHO estimate of 380 thousand in 2003. It would appear that, according to WHO, India is in a post peak phase with peak having occurred way before 2001. We strongly beg to differ. We believe that both the HIV incidence and the HIV prevalence in India are still on the rise.

4. Discussion

It should be noticed that the estimates of AIDS deaths in India that we have derived depend rather strongly upon the location of the peak on the gamma curve. They also depend upon the gamma curve itself. We expect that in different societies, different gamma curves will be the appropriate ones. Instead of $p = 4$ as in the above case, perhaps p should be equal to three if the number of hiv positive people goes up faster in the beginning (than for $p = 4$), or p may be equal to five if it goes up slower. It stands to reason that if we measure hiv prevalence in a group of high risk people wherein TB and STD's are widely present (FSW's in Mumbai?) with another group where these factors are absent, then hiv prevalence will move along different curves in these two groups. In Figure 4, we show the gamma curves for $p = 3, 4$ and 5 for comparison with peak at the same point.

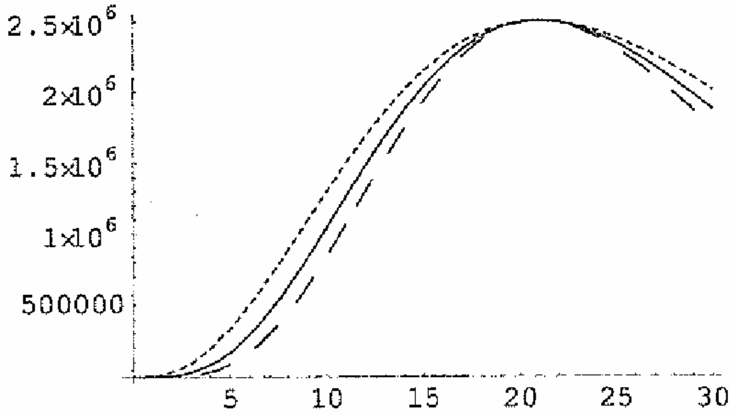

Figure 4. Graphs of the gamma curves for $p = 3$ (Short dashes), $p = 4$ (solid line) and for $p = 5$ with the peak located at $t = 21$ in each case. Notice that for $p = 5,$ the prevalence rises slower (as compared with the standard case for $p = 4$) and goes down faster after the peak, while for $p = 3,$ it is the other way around.

In the next figure, we calculate hiv prevalence assuming that the cumulative incidence till 2004 was 3 million and the total number of AIDS cases was 500,000 and calculating prevalence by E-M (expectation maximisation) technique (see [3] for details). We compare hiv prevalence thus calculated with gamma curves for $p = 4$ and for $p = 5$. Notice that the one for $p = 5$ is much closer.

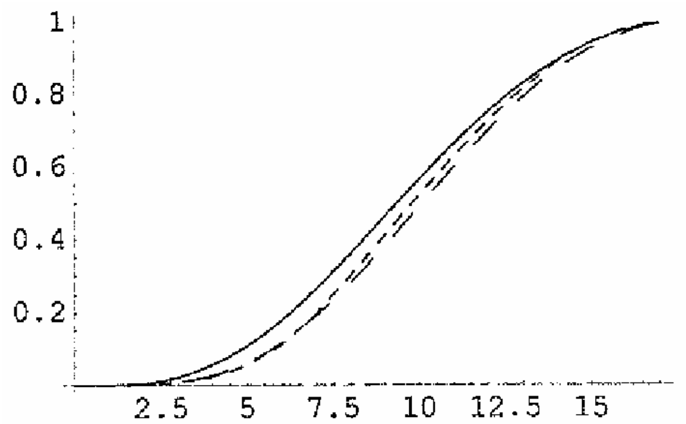

Figure 5. Prevalence curve in one particular case (short dashes) compared with gamma curves for $p = 5$ (long dashes) and for $p = 4$. The one for $p = 5$ is much closer.

We also show a comparison with hiv prevalence calculated by NACO before 2005 (which numbers were revised downwards later on) with appropriate gamma curves for $p = 3$ and for $p = 4$ with peaks in different years. This time the curve for $p = 3$ seems to be closer!

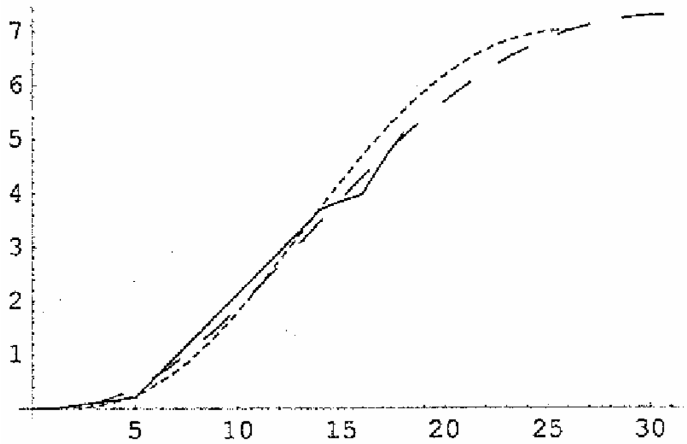

Figure 6. This figure shows the actual HIV prevalence in India as measured by NACO before 2005 (solid line) with Gamma curves for $p = 3$ (long dashes) with peak in 2016 and with $p = 4$ with peak in 2011. The value of A was chosen in each case so as to give the best fit (visually).

In view of all this, we shall now calculate the HIV incidence and the AIDS deaths in 2006 for various other cases. The results are given in the following graphs. It is to be noticed that the value of p affects the results rather more than the value of θ. However, in both cases, the differences are minor compared to the differences due to the location of the peak.

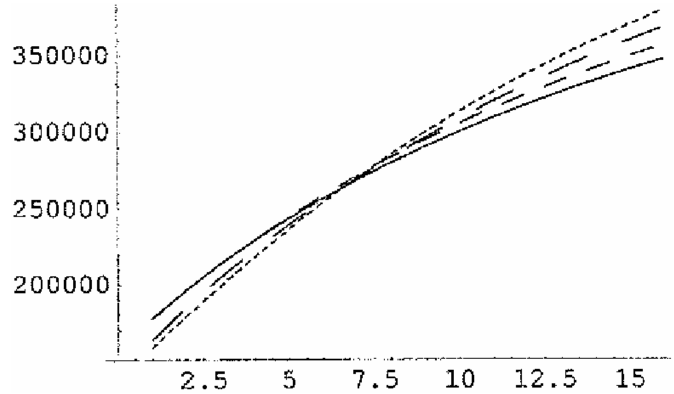

Figure 7. This graph shows HIV Incidence in India in 2006 with peaks (along x-axis) anywhere from 2001 to 2016 in four cases (1) $\theta = 2.5$, $p = 4$, solid line, (2) $\theta = 2.5$, $p = 5$, short dashes (3) $\theta = 5$, $p = 4$ medium dashes and (4) $\theta = 5$, $p = 5$, long dashes. The curves are closer for the same value of p.

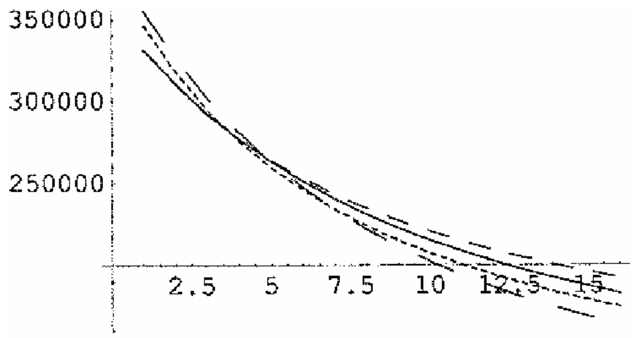

Figure 8. This graph shows AIDS deaths in 2006 with peaks (along x-axis) anywhere from 2001 to 2016 in four cases (1) $\theta = 2.5$, $p = 4$, solid line, (2) $\theta = 2.5$, $p = 5$, short dashes (3) $\theta = 5$, $p = 4$ medium dashes, and (4) $\theta = 5$, $p = 5$, long dashes. Notice that the effect of p is more than that of θ. Once again, for the same value of p, the curves are closer.

5. Conclusion

The estimate of the number of AIDS deaths in India depends rather strongly upon the location of the peak on the HIV prevalence curve. While WHO appears to think that India is in a post peak phase with peak having occurred way before 2001, and consequently, the number of AIDS deaths in India is extremely high (about 400,000 in 2003), we think otherwise. We believe that the peak in India is still a few years away, perhaps five or more, and the number of AIDS deaths in India was of the order of 200,000 in the year 2006. Our previous statement [4] that "a total of 122997 AIDS patients died in the year 2004" still stands.

References

[1] B. D. Aggarwala, On estimating HIV prevalence in Canada and in the United States, Far East J. Appl. Math. 20(3) (2005), 335-354.

[2] B. D. Aggarwala, On estimating HIV prevalence in India, Far East J. Math. Sci. (FJMS) 21(2) (2006), 197-210.

[3] B. D. Aggarwala, Calculating HIV prevalence in India-again, Proceedings of the World Congress on Engineering and Computer Science, Berkeley, California, Oct. 24-26, 2007.

[4] B. D. Aggarwala, On HIV Prevalence and AIDS Deaths in India, Current Themes in Engineering Technologies: World Congress on Engineering and Computer Science: Selected Papers, American Institute of Physics, 2008.

[5] R. Brookmeyer and M. H. Gail, AIDS Epidemiology, A Quantitative Approach, Oxford University Press, 1994.

[6] J. Chin, The AIDS Pandemic, The Collision of Epidemiology with Political Correctness, Radcliffe Publishing, Oxford, 2007.

[7] Private Communication with Prof. Chin.

[8] The Pioneer, March 2007.

[9] UNAIDS Epidemiological Fact Sheets on India, Oct. 16, 2007.