Indian J Med Res 107, April 1998, pp 159-172

Projection of HIV infection in Calcutta

A. Basu, S. Basu, M.S. Chakraborty*, A. Dewanji, J.K. Ghosh & P.P. Majumder

*Indian Statistical Institute & *AIDS Control Programme, West Bengal, Calcutta*

Accepted March 31, 1998

Starting with the base year of 1991, the HIV infection projection for 1992-99 for the total, as well as various high-risk sub-populations of Calcutta, the first of its kind is provided. These projections are based on statistical methodology developed in this paper. Our methodology for spread of HIV infection takes into account various social interactions and practices and also uses available data. Rates of these interactions and practices and estimates of demographic parameters used in making projections were obtained primarily from surveys and census data. Since one of these estimated rates, that of HIV transmission rate through heterosexual encounters between an infected and an uninfected had a large range, we have provided two sets of projections based on the largest of these rates (worst-case scenario) and another that is consistent with the available data. The total projection of the number of HIV infected cases in Calcutta for 1999 is between 49,000 and 1,26,000. Separate projections are also provided for high-risk sub-groups. Among these, the sex workers expectedly will continue to manifest the highest numbers of newly infected cases. The temporal rate of increase in prevalence is projected to be alarmingly higher in the general population than even among sex workers, although the actual prevalence will continue to be the lowest in the general population compared to all other sub-groups of the population.

Key words AIDS - epidemiology - statistical prediction - high risk

HIV seropositivity and its ultimate progression to AIDS has been a major concern in the Western countries since the early 1980s. In India, categorized by World Health Organization (WHO) as a Pattern III country[1], the problem was of a lesser magnitude till the late 80s. However, WHO estimates also indicate that currently 80 per cent ($\approx$ 4000 cases) of the daily additional number of HIV infections occur in the developing countries and nowhere is the disease spreading at a faster rate than in Asia[2]. Many feel that an outbreak of a wide-scale HIV epidemic in India may be imminent.

The first case of seropositivity in Calcutta was found among the city's commercial sex-workers (CSWs) in 1986[3]. In 1988 the first AIDS case in a Calcutta hospital was reported[3]. Serosurveillance activities are being undertaken since 1987 by the Government of West Bengal at the State level to determine the seroprevalence rate in the population. However, a statistical analysis of data on HIV seroprevalence in Calcutta with an aim of future projection is lacking. In the current report we present the findings of such a statistical study which began with 1991 as the base year and ended with the annual projected HIV infected adult population of Calcutta in 1999. This has been done by identifying several high risk and low risk groups, modelling interactions among susceptibles and infected persons within and between groups and by using appropriate estimates of the necessary demographic, transmission and infection rates within each group. In general, Calcutta is looked upon as a metropolis with an (as yet) low

prevalence rate of seropositivity compared to the other Indian cities. A realistic projection, however will be of immense value to the planners and health workers.

The high risk groups pertinent to a study on HIV infection are (1) commercial sex-workers, (2) slum dwellers and (3) pavement dwellers. The rest of the population is termed as (4) general population. The high risk activities which puts the males of groups 2, 3 and 4 at risk are (*i*) promiscuity (having multiple sex partners, primarily through regular visits to CSWs); and (*ii*) sharing of needles for taking injectable-drugs. The females of group 1, and a proportion of females in groups 2 and 3 who are part time sex-workers, are both at risk of getting the virus and also transmitting it *via* unprotected sexual contact with clients. Thus unprotected sexual contact with infected spouses is a source of risk to individuals of both sexes in groups 2, 3 and 4. Another potential source of infection is blood transfusion, but experts are of the opinion that the careful screening presently done by the blood banks of the city leaves little room for the propagation of the virus *via* this route. Homosexuality, although prevalent in a low degree, may not be an important source of HIV transmission in India.

A fifth group is the mobile truckers population among whom a very high rate of promiscuity exists. Through visits to the city's red-light district they contribute directly towards infecting the city's sex-workers. However, not being residents, they are not enumerated among the resident population of the city.

Material & Methods

Data for this study comprised primarily of demographic information about the city of Calcutta, sources of HIV infection and transmission rates. Most data were taken from census records, several unpublished surveys and a published report of the Government of West Bengal[3]. However, because of lack of availability of certain types of data in the public domain we had to sometimes rely on experts' opinion.

The 1991 census count of the total population in Calcutta was 10.68m (million). Slum dwelling population of Calcutta in 1991 is estimated to be 2.5m of whom 1.6m (0.8m male and 0.8m female) are in the potentially sexually active age group. Among the slum dwelling males (SDM) 78.5 per cent are married and most live with their families[3]. We assume that 75 per cent of the married SDMs live with their wives in the city. Therefore, of 0.8m slum dwelling females (SDFs) about 0.47m are married and live with their husbands and the rest are either unmarried, widowed or separated from their husbands.

A survey[3] showed that 46.2 per cent of the SDMs regularly visited CSWs. It also indicated that the number of visits varied; as an approximate central figure, we assume that a promiscuous person visits sex workers once in 5 days. We assume the same rate of visit for the promiscuous males in other groups. The same survey[3] estimated that 5.6 per cent of the SDMs shared needles for injectable drug use. It is estimated that a small portion (about 2%) of the married SDFs (about 9,420) act as part time sex workers. Sharing of needles for drug injections is not known to be present among any sub-group of the female population of Calcutta.

For the pavement dwelling population, the underage and old populations are relatively small. We estimate the pavement dwelling population in 1991 to be 1.1m of which 0.7m are males (PDMs) and 0.3m are females (PDFs) in the sexually active category. It is assumed that all PDFs are married. Among PDMs 30.4 per cent are regular visitors to commercial sex-workers[3]. Approximately 12 per cent of the PDFs *i.e.*, about 36,000 are part-time sex workers[3]. The practice of needle-sharing is prevalent among 4.6 per cent of PDMs[3].

It is difficult to pinpoint the total number of CSWs (all females) in Calcutta, and estimates put it anywhere between 30,000 - 1,00,000 including both full-time and part-time workers[3]. The full-time workers stay in this profession for 1 to 5 yr and then leave[3]. Assuming a period of 4 yr in the profession on the average, approximately 25 per cent of the full-time workers leave each year. They are replaced by fresh recruits. On a working day, a full-time worker serves an estimated 3 to 4 clients[3]. Adjusting for non work days, average number of clients per day for a full-time sex-worker is taken to be 2.5. We assume that the part time sex workers (2% of the married SDFs and 12% of PDFs) serve one client a day on the average. By matching the number of clients [coming from the SDMs, PDMs general population males

(GPMs) and truck drivers] with the number of servers, we estimate the number of full time sex-workers in Calcutta in 1991 to be about 45,000. Some of the above numbers are reproduced from Facts on AIDS[3] and the remaining are based on the experiences of surveyors and health workers.

In the general population the number of males (GPMs) and females (GPFs) are assumed to be 2m in each group after adjusting for the young and old age groups. There are no data on the per cent of GPMs who are promiscuous. However, it seems reasonable to assume that among the GPMs the proportion of promiscuity will be smaller than that among the SDMs and PDMs. For our purpose we have assumed that 10 per cent of GPMs are promiscuous. We also assume that there are no part time sex-workers among GPFs.

An unpublished study on the sexual behaviour pattern of truck drivers and helpers (Rao *et al*, 1993) estimated that between 3000 and 4500 truckers come into the city of Calcutta everyday and a survey of 50 truckers showed that 38 per cent visited sex-workers on the day they came to the city. Assuming 3750 to be the average daily number of truckers coming to the city, about 1425 truckers visit the sex-workers of Calcutta every day.

Next we consider HIV seropositivity rates. The slum dwelling and the pavement dwelling males (SDMs and PDMs) are assumed to have a HIV infection rate of 1 per 1000 in 1991, and the general population males have an infection rate of 1 per 10,000. A weighted average of these gives a combined infection rate close to 5 per 10000 in these groups, which is the rate of HIV infection estimated in West Bengal in 1991 (unpublished results of a survey by the School of Tropical Medicine, Calcutta - the STM survey) on the basis of 21727 observations. Much of the survey was done in Calcutta, and it is expected that the results of this survey represent crude approximations of the actual Calcutta figures.

For the STM survey, the corresponding sentinel surveillance programme over the period July 1987 to July 1996 resulted in zero positive cases among 12400 individuals screened. A simple Bayesian statistical technique of the estimation of a population proportion based on the above data (using an uniform prior distribution) leads to an estimated proportion of about 1 in 12401, providing empirical justification for the infection rate used for the GPMs.

The STM survey revealed an infection rate of approximately 5 per 1000 among CSWs in 1991. Approximately 10 per cent of the sex-workers use condoms regularly[3]. Infection among the GPFs and the non sex-working SDFs and PDFs is assumed to have been absent at the beginning of 1991. Surveys conducted by Bhoruka Public Welfare Trust (HIV/AIDS intervention project for truckers) based on 619 and 2406 truckers in 1993 and 1994 respectively showed 39 and 124 among them to be infected. The weighted average is slightly less than 6 per cent. There is no evidence of the rate having gone up subsequently. We have used a constant infection rate of 6 per cent among the truckers throughout our study.

To determine the rate of progression of HIV in a population, the probabilities of several biological events need to be known. We have assumed that the main avenues of HIV infection in the population of Calcutta are heterosexual contacts and sharing of needles. Based primarily on experts' opinion in this field, we have made two suppositions : (a) an uninfected female is four times more likely to get the virus through a single sexual contact with an infected male than the other way around; (b) the rate of contracting HIV for an uninfected person through sharing of needles with an infected partner is half the rate at which an uninfected female gets the infection through sexual contact with an infected male. Supposition (a) stems from rough estimates of total viral load in body fluids exchanged between partners in one sexual encounter (P. Gupta, personal communication). Prevailing wisdom that the chance of infection from needle sharing is substantially lower than that from infected male to uninfected female *via* sexual encounters forms the basis of our supposition (b). In any case, since drug usage through needle-sharing does not appear to be widely practised in Calcutta, supposition (b) does not have a critical effect on our projections. WHO estimates suggest that an uninfected person will contract HIV in 100-1000 sexual encounters with an infected person[1]. Since the given range is extremely wide, and one would like to know the sensitivity of the projected

numbers to changes in these rates, we have considered two alternative scenarios. In the first scenario it is assumed that it takes 100 unprotected sexual contacts with infected male(s) on the average for an uninfected female to get HIV infection. (This is at one extreme of the rates suggested by WHO and corresponds to the worst case scenario). The corresponding probability of infection from female to male, and through sharing of needles, are then determined according to the suppositions (a) and (b) above. In the second scenario it is assumed that 150 sexual contacts are necessary on the average for the same event to happen. We emphasize that these figures (100 or 150 sexual encounters) are unrelated to the sexual activity of a person *per se*, but relate only to sexual contacts between an uninfected person and an infected person. The second scenario has empirical validity because in this case the predicted rates of HIV infection among the CSWs of Calcutta closely match the observed figures of the STM survey. The estimated rates of seroprevalence among sex workers in this case are approximately 9.5 per 1000 in 1992, 14.5 per 1000 in 1993 and 20.6 per 1000 in 1994. The corresponding observed rates in the STM survey are 7.23 per 1000, 14.07 per 1000 and 19.98 per 1000 in 1992, '93 and '94 respectively.

The demographic composition and the infection rates described earlier are for the year 1991. Our procedure annually updates the composition of the population taking into account the estimated demographic growth rates and estimates the additional number of infected persons every year. Our methodology considers the interactions between the various relevant subsections of the population and using the rates given earlier, calculates the probability of an uninfected individual belonging to a subsection becoming infected during the year. This probability multiplied by the total number of uninfected persons in the subsection yield the expected number of new cases in the subsection. The total of these new cases added to the number of infected persons at the beginning of the year provides the total number of HIV infected persons at the beginning of the next year.

Results

We first provide results for certain sub-groups of the total population along with relevant details on how the demographic and other rates were used in making these projections. Subsequently, we combine results of sub-groups to arrive at projections for Calcutta as a whole.

Projections for sub-groups : In the following, we provide one year increments only for scenario 1. Simply changing the appropriate probabilities will yield the corresponding figures for scenario 2. The overall number of infected individuals, as well as the number of infected individuals in each high risk group over the years for the period of our study are given in Tables I and II for the two scenarios and the corresponding increments in 1991 for scenario 1 are given in Table III.

Slum dwellers, male (SDM) and female (SDF) : A given male may be at risk from several possible sources (promiscuity, needle sharing, sexual contact with infected wife); the proportion of males at risk from the other two sources are considerably smaller compared to promiscuity. In our treatment, therefore, we have not considered the males (of any subpopulation) to be at risk from two sources simultaneously unless one of them is promiscuity. The hierarchy in which we consider the separate groups are given in the classification scheme below. First we look at the needle-sharing group, and within them consider the promiscuous and non-promiscuous groups. Next we look at the sub-groups of the non needle-sharing population (himself promiscuous and wife sex worker, himself promiscuous and wife not sex-worker, himself non-promiscuous and wife sex worker, himself non-promiscuous and wife not sex-worker). The males are not considered to be at risk from all three sources simultaneously.

For the SDMs the classification scheme is as follows:

1. Injectable drug users (SDM.IDU)
    1(A). promiscuous (SDM.IDU.P)
    1(B). non-promiscuous (SDM.IDU.NP)
2. non injectable drug users (SDM.NIDU)
    2(A). promiscuous (SDM.NIDU.P)
        2(A)(I). wife sex worker (SDM.NIDU.P.WSW)
        2(A)(II). wife not sex worker (SDM.NIDU.P.NWSW)
    2(B). non-promiscuous (SDM.NIDU.NP)
        2(B)(I). wife sex worker (SDM.NIDU.NP.WSW)
        2(B)(II). wife not sex worker (SDM.NIDU.NP.NWSW)

Table I. Total number of seropositive persons and rate of seropositivity in each category for the years 1991-1999 : Scenario 1

Year	Infection	SDM	PDM	GPM	SW	SDF (NSW)	PDF (NSW)	GPF	Total
1991	Total no.	800	700	200	452	0	0	0	2152
	Rate	0.0010	0.0010	0.0001	0.005	0.0	0.0	0.0	0.0002
1992	Total no.	1149	945	366	1878	300	171	65	4874
	Rate	0.0014	0.0013	0.0002	0.0117	0.0004	0.0006	0.0000	0.0004
1993	Total no.	1928	1470	759	3219	535	291	142	8344
	Rate	0.0023	0.0020	0.0004	0.0198	0.0006	0.0011	0.0001	0.0007
1994	Total no.	3250	2342	1430	5126	909	462	296	13815
	Rate	0.0039	0.0032	0.0007	0.0310	0.0011	0.0017	0.0001	0.0012
1995	Total no.	5343	3708	2495	7960	1535	735	568	22344
	Rate	0.0063	0.0050	0.0012	0.0475	0.0018	0.0026	0.0003	0.0020
1996	Total no.	8574	5801	4138	12206	2539	1165	1009	35432
	Rate	0.0099	0.0077	0.0019	0.0717	0.0030	0.0041	0.0005	0.0031
1997	Total no.	13481	8971	6638	18485	4100	1828	1696	55199
	Rate	0.0154	0.0117	0.0030	0.1070	0.0047	0.0063	0.0008	0.0047
1998	Total no.	20805	13698	10371	27566	6485	2836	2748	84509
	Rate	0.0234	0.0176	0.0047	0.1572	0.0074	0.0097	0.0012	0.0071
1999	Total no.	31486	20592	15821	40232	10063	4347	4329	126870
	Rate	0.0349	0.0261	0.0070	0.2261	0.0113	0.0146	0.0019	0.0105

SDM, slum dwelling male; PDM, pavement dwelling male; GPM, general population male; SW, sex-worker; SDF (NSW), slum dwelling female (non sex-worker); PDF (NSW), pavement dwelling female (non sex-worker); GPF, general population female

Table II. Total number of seropositive persons and rate of seropositivity in each category for the years 1991-1999 : Scenario 2

Year	Infection	SDM	PDM	GPM	SW	SDF (NSW)	PDF (NSW)	GPF	Total
1991	Total no.	800	700	200	452	0	0	0	2152
	Rate	0.0010	0.0010	0.0001	0.005	0.0	0.0	0.0	0.0002
1992	Total no.	1037	862	310	1516	232	132	50	4139
	Rate	0.0013	0.0012	0.0001	0.0095	0.0003	0.0005	0.0000	0.0004
1993	Total no.	1467	1150	522	2358	415	229	103	6244
	Rate	0.0018	0.0016	0.0002	0.0145	0.0005	0.0008	0.0000	0.0006
1994	Total no.	2124	1585	850	3408	631	332	182	9112
	Rate	0.0025	0.0022	0.0004	0.0206	0.0008	0.0012	0.0001	0.0008
1995	Total no.	3063	2204	1325	4773	929	465	304	13063
	Rate	0.0036	0.0030	0.0006	0.0285	0.0011	0.0016	0.0001	0.0011
1996	Total no.	4371	3055	1987	6584	1349	648	483	18477
	Rate	0.0051	0.0040	0.0009	0.0387	0.0016	0.0023	0.0002	0.0016
1997	Total no.	6166	4225	2899	8996	1937	901	739	25863
	Rate	0.0070	0.0055	0.0013	0.0520	0.0022	0.0031	0.0003	0.0022
1998	Total no.	8609	5814	4138	12210	2750	1248	1095	35864
	Rate	0.0097	0.0075	0.0019	0.0697	0.0031	0.0042	0.0005	0.0030
1999	Total no.	11902	7948	5811	16466	3861	1722	1583	49293
	Rate	0.0132	0.0101	0.0026	0.0926	0.0043	0.0058	0.0007	0.0041

SDM, slum dwelling male; PDM, pavement dwelling male; GPM, general population male; SW, sex-worker; SDF (NSW), slum dwelling female (non sex-worker); PDF (NSW), pavement dwelling female (non sex-worker); GPF, general population female

Table III. The estimated total population, the total number of infected individuals, and the total number of new cases during the year 1991 among the different population sub-groups

Category	Total	Infected	New cases	
			Scenario 1	Scenario 2
Table III(a). Slum dwellers (excluding the part time sex-workers)				
SDM	800000	800	349	237
SDM.IDU	44802	45	52	39
SDM.IDU.P	20700	21	32	25
SDM.IDU.NP	24102	24	20	14
SDM.NIDU	755198	755	297	198
SDM.NIDU.P	348902	349	291	194
SDM.NIDU.P.WSW	4108	4	8	5
SDM.NIDU.P.NWSW	344794	345	283	189
SDM.NIDU.NP	406296	406	6	4
SDM.NIDU.NP.WSW	4784	5	6	4
SDM.NIDU.NP.NWSW	401512	401	0	0
SDF.NSW	791580	0	300	232
Table III(b). Pavement dwellers (excluding part time sex workers)				
PDM	700000	700	245	162
PDM.IDU	32199	32	38	24
PDM.IDU.P	9789	10	18	11
PDM.IDU.NP	22410	22	20	13
PDM.NIDU	667801	668	207	138
PDM.NIDU.P	203011	203	179	119
PDM.NIDU.P.WSW	4108	10	21	14
PDM.NIDU.P.NWSW	192570	193	158	105
PDM.NIDU.NP	467461	465	28	19
PDM.NIDU.NP.WSW	24041	24	28	19
PDM.NIDU.NP.NWSW	443420	441	0	0
PDF.NSW	264000	0	171	132
Table III(c). General population				
GPM	2000000	200	166	110
GPM.IDU	20000	2	3	2
GPM.IDU.P	2001	0	2	1
GPM.IDU.NP	18000	2	1	1
GPM.NIDU	1980000	198	163	108
GPM.NIDU.P	198000	20	163	108
GPM.NIDU.NP	1782000	178	0	0
GPF	2000000	0	65	50

(Table III. Contd.)

Table III(d). Commercial sex workers (including part time)

Category	Total	Infected	New cases	
			Scenario 1	Scenario 2
SW	90420	452	551	368
SW.FT	45000	225	358	239
SW.PT	45420	227	193	129
Total	10680000	2152	2722	1987

SDM, slum dwelling male; IDU, injectable drug users; P, promiscuous; NP, non-promiscuous, NIDU, non injectable drug users, WSW, wife sex worker; NWSW, wife not sex worker; SDF, slum dwelling female; NSW, non sex worker; PDM, pavement dwelling male; PDF, pavement dwelling female; GPM, general population male; GPF, general population female; SW, sex worker; FT, full time; PT, part time

The injectable drug users are believed to share needles in intimate and closed groups of small sizes. We have, therefore, assumed that the groups are of size 3, they do not change partners, and engage in needle-sharing activities twice a week.

First we look at the non-promiscuous sub-group of the injectable drug users. In 1991 the existing groups had 0, 1, 2 or 3 already infected members with probabilities $(1-p)^3$, $3p(1-p)^2$, $3p^2(1-p)$ and p^3, respectively, where p is the prevalent HIV infection rate among SDMs for that year. Hence the number of groups with 0, 1, 2 or 3 infected members is n times the appropriate probability where n is the total number of groups. Let these numbers be n_0, n_1, n_2 and n_3. For 1991 these numbers are 8010, 24, 0 and 0 respectively, n is 8034 and p is 0.001.

During the year the probability that one uninfected person will be infected only through sharing needle with one infected partner is $p_1 = 1 - (1-1/200)^{52\times2}$. This may be understood as follows : (1-1/200) is the probability of not getting infected in one needle sharing session with an infected partner; $(1-1/200)^{52\times2}$ is then the probability that the individual does not contract the virus over the year, throughout which he engages in two weekly needle-sharing sessions with an infected partner; p_1 which is the probability that the individual gets the disease during the year, is then $1-(1-1/200)^{52\times2}$. Individuals of the SDM.IDU.NP group are subject to the needle-sharing risk factor only. Groups which started with 0 or 3 infected members continue to have the same number of infected members. But after one year some groups that started with 1 infected member will have 2 or 3 infected members and some groups having 2 infected members earlier will have all 3 members infected. This can be calculated noting that an individual sharing needles with one infected partner gets the infection with probability p_1, and an individual sharing needles with two infected partners gets the infection with probability $p_2 = p_1 + p_1 - p_1 \times p_1$ (this is the combination of two events, namely that the individual gets the infection either from partner 1 or partner 2, both of whom are infected, assuming that the impacts of the events are independent). Thus if n_1^*, n_2^*, n_3^* and n_4^* represent the new numbers for the groups with 0, 1, 2 and 3 infected, then

$$n_0^* = n_0$$

$$n_1^* = (1-p_1)^2 n_1$$

$$n_2^* = 2p_1 (1-p_1) n_1 + (1-p_2) n_2$$

$$n_3^* = p_1^2 n_1 + p_2 n_2 + n_3$$

For 1991 these numbers were 8010, 8, 12 and 4. Thus while the total of previously infected individuals in this group was $n_1 + 2n_2 + 3n_3$ (24 in 1991), the total of new infected persons is $n_1^* + 2n_2^* + 3n_3^*$ (44 in 1991), and the number of new cases during the year is the difference of the two (20 in 1991). The probabilities p_1 and p_2 (equal to 0.4062 and 0.6475 respectively) do not change in future years.

Now we consider the needle sharing promiscuous group. As the infection rate among sex workers in 1991 is 5 per 1000, and only 10 per cent of them are assumed to use condoms regularly, the probability that an uninfected promiscuous male will get the

infection during the year 1991 on account of his promiscuity is p_3 = I - (1-0.005 × 0.9/400)$^{365/5}$ = 0.00082 (here 0.005 × 0.9 represents the probability of getting an infected, non condom-using sex worker at any visit, 0.005 × 0.9/400 is the probability of getting the infection at any visit), and the individuals in the SDM.IDU.P group are at risk from promiscuity as well as needle sharing. The probability that an uninfected promiscuous male sharing needles with one infected partner will be infected at the end of the year is $p_4 = p_1 + p_3 - p_1 \times p_3$ (again a probability of the combination of two events, the individual getting the infection through either promiscuity or needle sharing). Similarly the corresponding probability for a promiscuous person sharing needles with two infected partners is $p_5 = p_3 + 2p_1 - 2p_3 \times p_1 - p_1^2 + p_3 \times p_1^2$ (a probability of the combination of three events). Using these probabilities the changed composition of the three-member group is calculated. Note that here all group are subject to risk of infection. In this case, if n_0, ..., n_3 and n_0^* ,..., n_3^* represent the old totals and the new totals for the number of groups, we get

$$n_0^* = (1-p_3)^2 n_0$$

$$n_1^* = 3p_3(1-p_3)^2 n_0 + (1-p_4)^2 n_1$$

$$n_2^* = 3p_3^2 (1-p_3) n_0 + 2(1-p_4) p_4 n_1 + (1-p_5) n_2$$

$$n_3^* = p_3^3 n_0 + p_4^2 n_1 + p_5 n_2 + n_3$$

For 1991 the old totals for the number of groups were 6879, 21, 0 and 0, and the new totals were 6862, 24, 10 and 3. Then the number of new cases and the infected total in the SDM.IDU.P group is similarly determined (which is 32 for 1991). Combining these, the total increase among the slum dwelling, injectable drug user males is calculated (and equals 52 in 1991). The numbers are given in Table III(a). The probabilities p_4 and p_5 were 0.4067 and 0.6477 respectively in 1991.

Next we consider the non needle-sharing group. It is assumed that spouses on the average have sexual contacts twice a week. For an SDM whose wife is HIV seropositive, the probability of infection at the end of one year is p_6 = 1 - (1-1/400)$^{2\times52}$ = 0.2292. A portion of SDM.NIDU.P.WSW are subjected to this risk in addition to the risk of acquiring HIV through promiscuity. For an uninfected promiscuous male whose wife is infected, the risk of infection at the end of one year is $p_7 = p_3 + p_6 - p_3 \times p_6$ (in 1991 this was 0.2298). Considering that among sex workers 5 per 1000 were infected at the beginning of 1991 and applying the same rate of infection among the sex worker wives of the SDMs, the expected number of new cases among SDM.NIDU.P.WSW and SDM.NIDU.NP.WSW are determined. For example, the number of new cases of SDM.NIDU.NP.WSW is 0.005 × (4784-5) × p_6 (≈ 6) in 1991. Among the 4098 susceptibles in the SDM.NIDU.P.WSW group, 5 per 1000 have infected wives. Thus the expected number of new cases in this group is 4098 × 0.005 × p_7 + 4098 × (1-0.005) × p_3 (≈ 8). The probability that SDM.NIDU.P.NWSW will get infected at the end of one year is p_3 and thus the number of new cases in this group can be calculated. It may be noted that SDM.NIDU.NP.NWSW are not subjected to any risk factor and there will not be any new cases in this group. By adding the new cases of HIV infection over all groups the total number of new cases among SDMs in that year is obtained. The figures are shown in Table III(a). Note that p_6 does not change over the years.

The classification scheme for SDFs is as follows

1. Part time sex workers (SDF.SW)
2. Not sex workers (SDF.NSW)
   - 2(A). Husbands infected (SDF.NSW.HINF)
   - 2(B). Husbands not infected (SDF.NSW.NHINF)

The spread of infection among all sex workers is considered later. In group 2(B), the females are not subjected to any risk factor and there will be no new cases among them. For a woman whose husband is HIV seropositive (group 2A) the risk of getting the disease at the end of one year is p_8 = 1-(1 - 1/100)$^{2\times52}$ = 0.6484. Using this the number of new cases among SDF.NSW is determined [Table III(a)].

Pavement dwellers and general population : Calculations for these sub-groups are similar to that of the slum dwellers; projections are presented in Tables III(b) and (c).

Sex worker population (SW) : The part time sex workers come from married SDF and PDF populations at an estimated rate of 2 and 12 per cent respectively. That fixes the number of part time

workers to be 45,420. It is also supposed that they get 1 client per night, on the average. Since full time workers serve 2.5 clients daily, by balancing the number of clients and servers the total number of full time SWs is estimated to be 45,000. The probability of a sex worker getting an infected client at any encounter is

$$\frac{1/5\,[8 \times 0.462 \times 0.001 + 7 \times 0.304 \times 0.001 + 20 \times 0.1 \times 0.0001] + [\,0.03750 \times 0.38 \times 0.06]}{1/5[8 \times 0.462 + 7 \times 0.304 + 20 \times 0.1] + [0.03750 \times 0.38]}$$

$$= 0.0013$$

This is a combination of the infection rates of the different populations (1 per 1000 among SDMs, PDMs, 1 per 10000 among GPMs, and 6 per 100 among the truckers) by the sizes of the promiscuous populations (0.8×0.462m among SDMs, 0.7×0.304m among PDMs, 2.0×0.1m among GPMs, and 3750×0.38 among truckers), taking into account the fact that the promiscuous males among the SDMs, PDMs and GPFs visit sex workers once in five days on the average, whereas 3750×0.38 truckers visit the sex workers everyday.

Then the probability that an uninfected and unprotected full time sex worker gets the infection during the year 1991 is

$$1 - (1 - 0.0013/100)^{365 \times 2.5} = 0.0118.$$

The corresponding probability for a part time sex worker is 0.0047. In 1991 rate of HIV infection among sex workers was 5 per 1000 and 10 per cent of commercial sex workers were regular condom users. Among 45,000 full time sex workers 225 were already infected and 4478 regular condom users were not at risk at this point. Hence 40297 full time sex workers are at risk and the estimated number of new cases among them is approximately 477. Assuming that 25 per cent of full time sex workers leave each year, 119 of those infected would leave. The net increase in the infected group of Calcutta's full time sex worker population is 358. Similarly the number of new cases among the part time sex workers is 193. Therefore, in the SW population in Calcutta the number of new cases in 1991 is estimated to be 551 [Table III (d)].

Combining all these the number of new HIV infections during the year 1991 is 2722 (Table III). It is to be noted that the total population in the last row of Table I includes the non susceptible groups also and is not just the sum of the totals in Tables III(a) - (d).

The general mechanism of the growth of the infected population in the following years is similar to the process described here for 1991. However, some further details of the procedure need to be explained. We assume a population growth rate of 1.5 per cent for all the population sub-groups (this is roughly equal to the national average growth rate). Thus the population in any particular group in any of the following years is 1.015 times the population in the previous year. We assume that this increment is entirely in the uninfected group. For example the number of groups with 0, 1, 2 and 3 infected in the SDM.IDU category are 6965, 24, 10 and 3 at the beginning of 1992, with 103 being added to the number of groups with zero infected.

We finally note that for each sub-group, the infected population during any of these years is the total of the infected population and the number of new cases during the year. This allows us to calculate any rate necessary for estimating the infected total in the following year. The infection rate among the SWs is calculated as a weighted average of the infection rates among the full-timers and the part-timers, the weights being the number of clients these groups receive per day (rather than a simple average). The remaining calculations then proceed as above.

Projections for the entire population : Starting with the base year 1991, the annual projections of HIV infection until 1999 are provided in Tables I and II for the two scenarios. The projection for 1999 in the total population of Calcutta is between $\approx$ 49,000 and $\approx$ 1,26,000 HIV infected cases. These are quite sensitive to changes in the assumed probabilities of HIV infection through heterosexual encounters between uninfected and infected, although due to the complex pattern of interactions among individuals it is difficult to determine a simple relation between the two. Sensitivity of projections on probabilities of transmission through needle sharing is nominal.

In Table III(a) - (d) the number of new cases of HIV infection in 1991 in various categories are provided. The first column gives the total population

in various sub-categories and the second column gives the number already infected at the beginning of 1991. The last two columns show the increase in these categories under the two scenarios considered. In Table III this exercise has been carried out for the first year only, *i.e.*, from 1991 to 1992. Similar calculations may be done for each year in the range 1992-1999, based on the rates of the previous year.

Fig. 1(a) - (c) graphically show the increase in the rate of HIV infection among the SDM, PDM and GPM under both scenarios. Fig. 2(a) - (f) compares the three chief risk factors for HIV infection among males, namely use of injectable drugs, promiscuity and sexual contacts with infected wives. Infected numbers in the three male populations under each scenario are plotted in the same graph for comparison. Graphs for scenarios 1 and 2 are drawn to the same scale so that the sensitivity of infection rates to the given rates of drug use and promiscuity become apparent. Rate of HIV infection among non-sex worker females are given in Fig. 3(a) and (b) separately under the two scenarios. Fig. 4(a) and (b) provide, respectively, the projected rates of HIV infection in the highest risk group (CSWs) and in the total population of Calcutta.

Discussion

Even the conservative projections of the World Health Organization pertaining to HIV infection are frightening. By the year 2000, it is estimated that 10 m people in Asia will be infected[2]. These estimates have been generated for broad geographical regions. However, for effective planning of AIDS control programmes, it is necessary to generate these

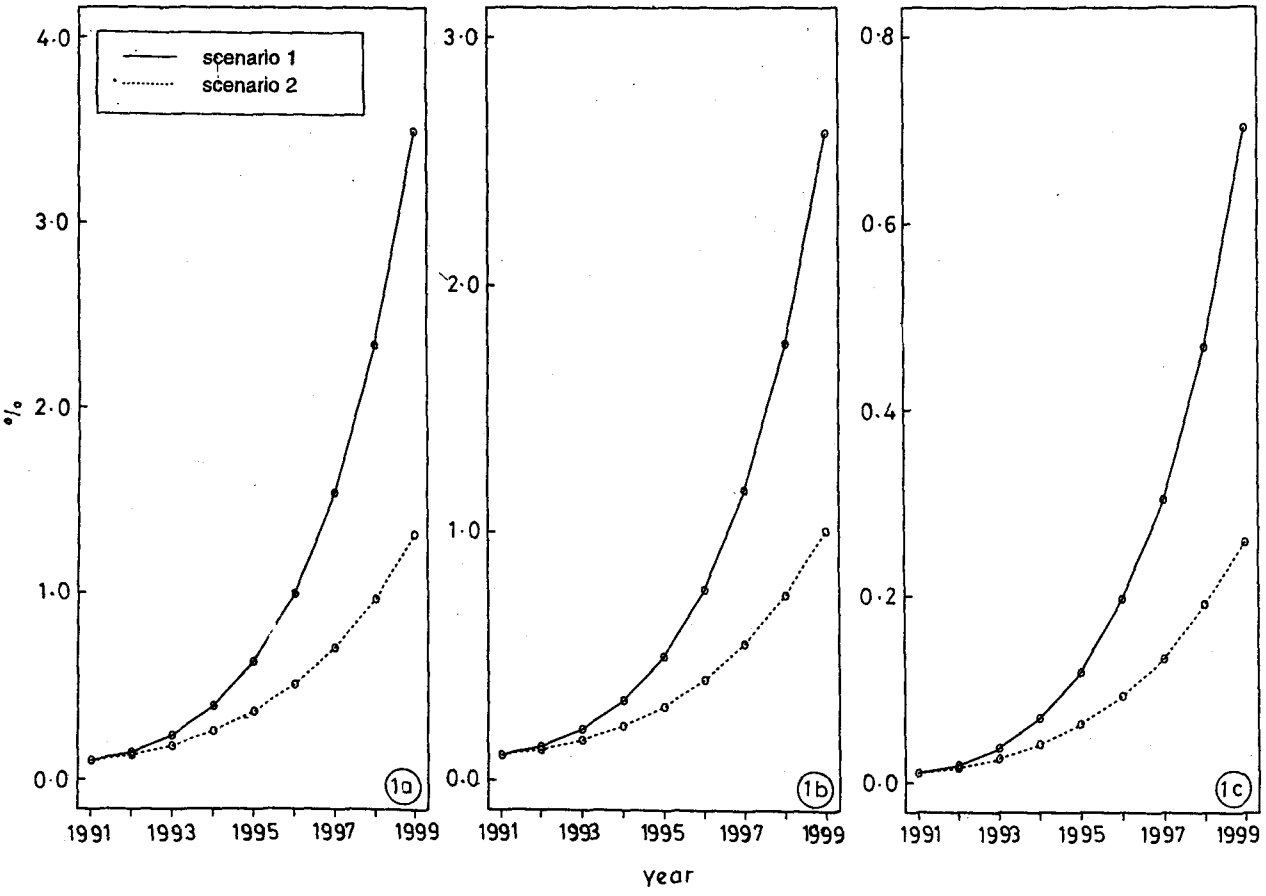

Fig. 1. Projected rates of increase of HIV infection during 1991-1999 in various male subpopulations of Calcutta. **(a)** Slum dwelling, **(b)** Pavement dwelling, **(c)** General population.

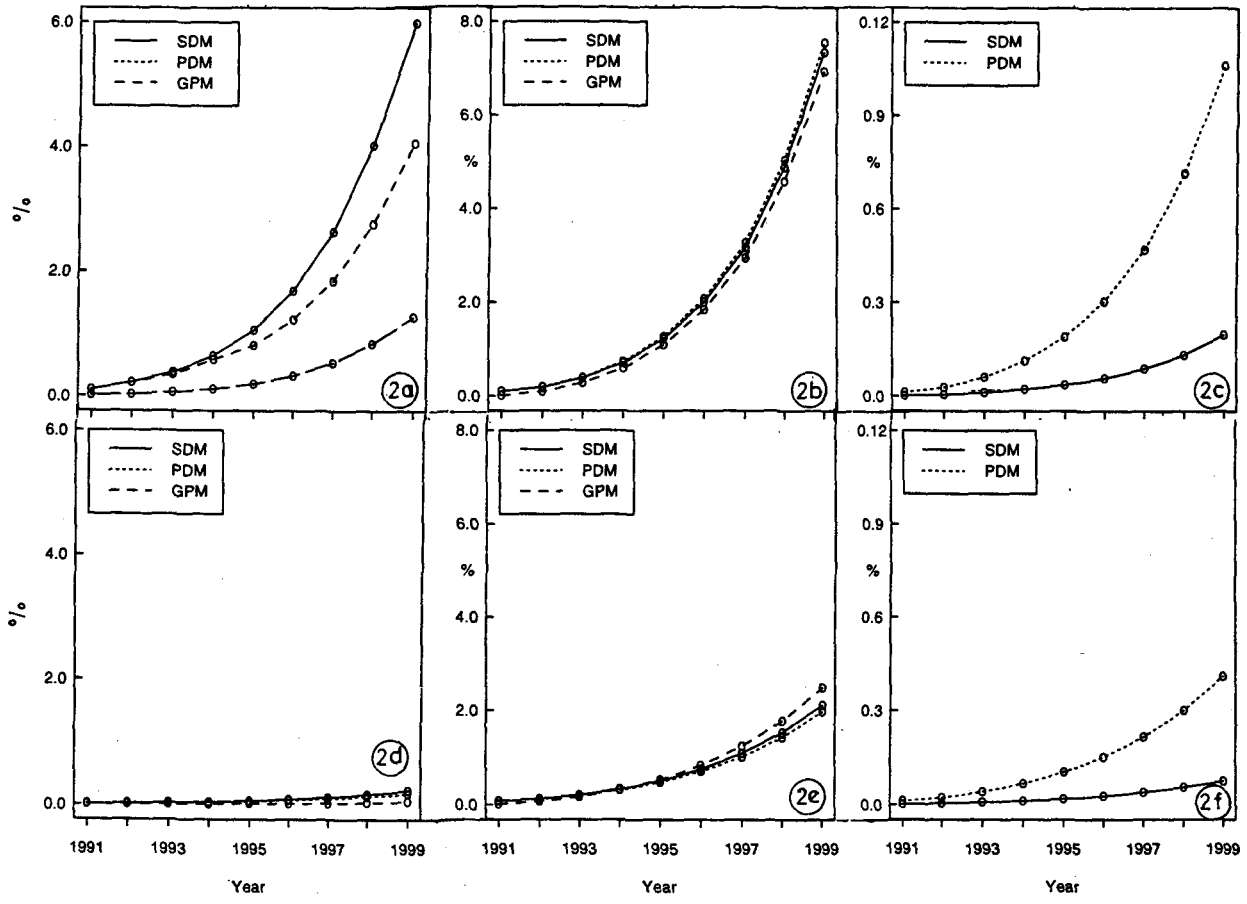

Fig. 2. Projected rates of increase of HIV infection during 1991-1999 due to the high risk behaviours leading to infection in various male sub-populations [Slum dwelling (SDM), Pavement dwelling (PDM), General population (GPM) of Calcutta]. **(a)** Injectable drug use, Scenario 1, **(b)** Promiscuity, Scenario 1, **(c)** Infection from wife due to commercial sex work, Scenario 1, **(d)** Injectable drug use, Scenario 2, **(e)** Promiscuity, Scenario 2, **(f)** Infection from wife due to commercial sex work, Scenario 2.

estimates for smaller regions because of the enormous variability in prevalence of HIV infection that surely exists within large geographical regions.

The quality of data required for this purpose is not uniformly satisfactory; the reasons being the nature of the sensitive information solicited, current low prevalence of HIV infection rate, absence of statistically designed pilot studies and lack of adequate infrastructure for serosurveillance. Therefore, sophisticated epidemiological modelling[4,5] or the use of other standard methods of projection (*e.g.*, back-calculation) of HIV incidence[6-10] with desirable accuracy is difficult. Even WHO is aware of the lack of data in India and their predictions are delphic, *i.e.*, based on iterated guesstimates and expert's opinion[1]. On the other hand, standard trend growth modelling approaches have been used[11] to estimate nation-wide rates of infection. However, such an attempt requires a justification of the particular choice of the model; besides the same approach may not necessarily be appropriate to model the infection rates across cities with variable population densities and social patterns.

We have proposed an objective statistical method for making projections of HIV infection which makes much less demand for data in comparison with some of the existing methods, such as the method of back-calculation[6,12,13]. The proposed methodology has been applied to make short-term projection (1991 to 1999) of HIV infection of Calcutta.

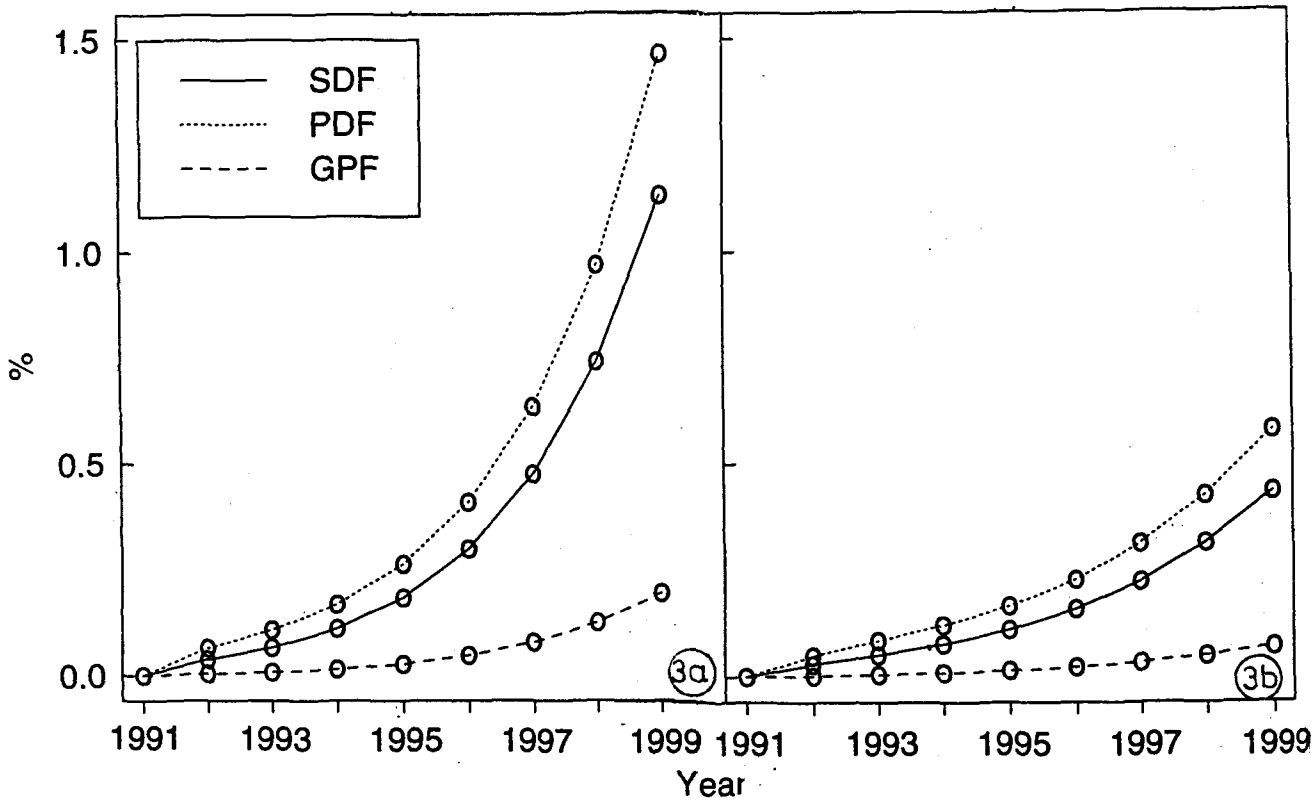

Fig. 3. Projected rates of increase of HIV infection during 1991-1999 in various non-sex working female sub-populations [Slum dwelling (SDF), Pavement dwelling (PDF), General population (GPF)] **(a)** Scenario 1**(b)** Scenario 2.

Our prediction is that by the turn of the century, there will be approximately between 49,000 and 1,26,000 HIV infected individuals in Calcutta. In the total population as well as in most high risk sub-groups, such as CSWs, slum and pavement dwellers *etc.*, the infection is spreading at an exponential rate, although the rates seem to have declined slightly towards the end of the century. In spite of the fact that the current rate of HIV infection in Calcutta is much lower than those in Mumbai and Chennai, Calcutta's burden of HIV infection will become very heavy by the turn of the century.

Why does our projection for 1999 indicate such a broad range (49,000-1,26,000) of cases? The reason is the approximate exponential growth of the number of infected persons. Even a small change in the parameter of this annual exponential growth can result, over a period of 8 yr, in a vastly different prediction of the number of infected persons. The change in this annual growth rate is caused because of our lack of precise knowledge of one of the key parameters to which our projections are very sensitive. How many sexual encounters with HIV infected men does it take, on the average, for an uninfected woman to become infected? Current biological wisdom[1] places this number anywhere between 100 and 1000 encounters. We have used two estimates : 100 - which corresponds to the worst-case scenario leading to a projection of a total of 1,26,000 cases in 1999, and 150 - which results in a projection of a total of 49,000 cases in 1999 and also provides a close match between the observed and expected rates of seropositivity among sex workers over 1992-1994. We believe that 150 may be a realistic estimate since it yields predictions consistent with observations. Thus 49,000 HIV infected cases in 1999 may be the most realistic projection for Calcutta. Our projections should be tested by collecting data through well-designed sample surveys. If our projections are found to be close to observed rates, then the methodology can be used to make projections into the twenty-first

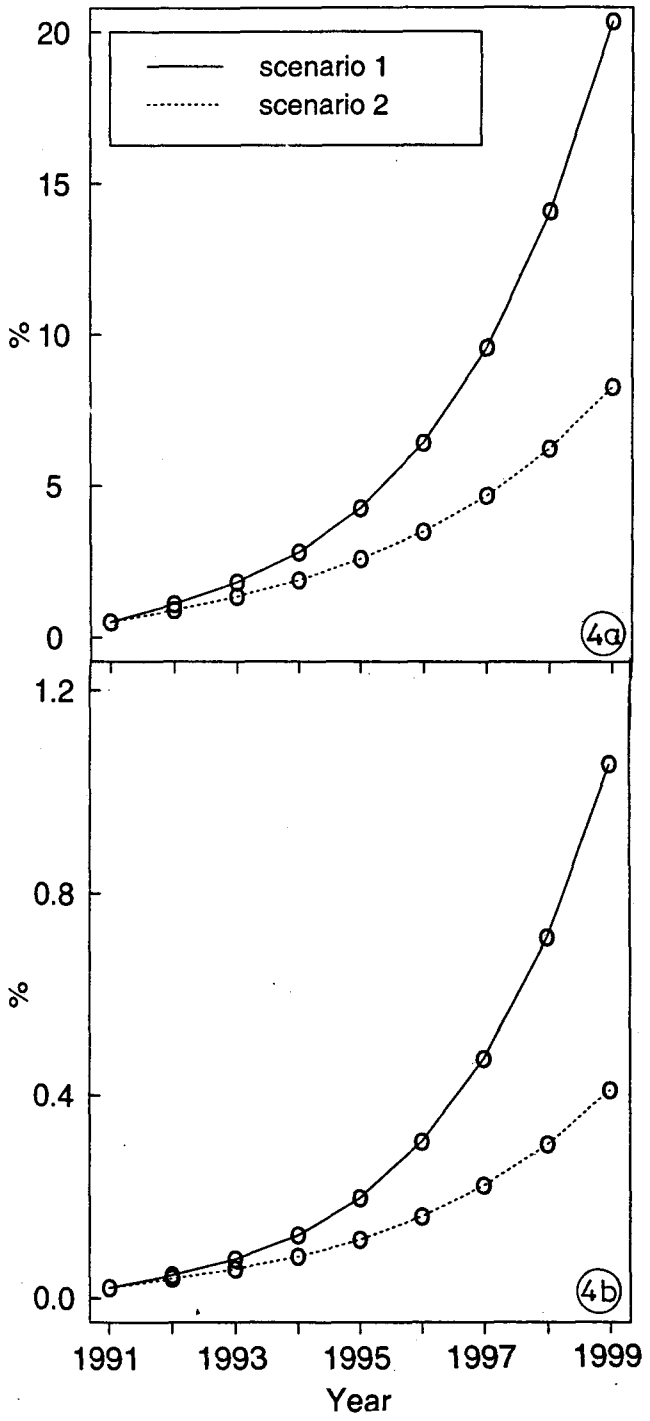

Fig. 4. Projected rates of increase of HIV infection during 1991-1999 in the **(a)** highest risk group-commercial sex workers, and **(b)** total population of Calcutta.

century. Further, our methodology can also be used for making projections for other Indian cities provided current estimates of appropriate parameters are available.

In our projections we have ignored the adjustments in the infected population due to death for the following reasons. Firstly, majority of the HIV infected population (sex-workers and promiscuous males) belong to the younger generation and the natural death rate among them is very small for a short-term period. Secondly, the passage to full blown AIDS from HIV contraction may take several years of incubation. In Tables I and II it may be observed that in any given year the majority of the HIV cases are infected during the three previous years. Thus even if the four year old cases die or become sexually inactive due to becoming full blown AIDS patients they do not cause a significant dent in the infected population. Besides in four to five years the full time sex-worker population gets almost completely replaced. In view of the above, projected incidences of HIV infection in any year is easily obtained by subtracting from the projected prevalence for the year the previous year's prevalence estimate.

We finally note that our finding of exponential spread of HIV infection lends some support to the earlier usage[11] of the exponential trend growth model. Further our findings also confirm the belief[14] that the alarming spread of HIV infection will not be restricted to the high risk groups, but will also spread rapidly to the general population through heterosexual transmission.

Acknowledgment

Authors are grateful to Sir David Cox, Nuffield College, Oxford, for drawing our attention to a set of papers compiled and edited by him with R.M. Anderson and H.C. Hillier. Authors are also grateful to Dr P. Gupta, University of Pittsburgh, for helpful discussions.

References

1. World Health Organization. *Current and future dimensions of the HIV/AIDS pandemic : a capsule summary*. Geneva : World Health Organization. September 1990.
2. Different approaches to AIDS research and intervention [editorial]. *Biotechnol Dev Monitor* 1997; *30* : 1-4.
3. National AIDS Control Programme, Department of Health & Family Welfare, Govt. of West Bengal. *Facts on AIDS, West Bengal facing the challenge*. March 1994.
4. Bailey NTJ. *The mathematical theory of infectious diseases and its applications*. London : Charles Griffin, 1975.

5. Breslow N, Day NE. *Statistical methods in cancer research.* Vols. 1 & 2. IARC Scientific Publications Nos 32 and 82. Lyon : International Agency for Research on Cancer 1980 and 1987.

6. Bacchetti P, Segal MR, Jewell NP. Back-calculation of HIV infection rates. *Statis Sci* 1993; *8* : 82-101.

7. Brookmeyer R. Reconstruction and future trends of the AIDS epidemic in the United States. *Science* 1991; *253* : 37-42.

8. Brookmeyer R, Gail MH. A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. *J Am Statist Assoc* 1988; *83* : 301-8.

9. Chin J. Current and future dimensions of the HIV/AIDS pandemic in women and children. *Lancet* 1990; *336* : 221-4.

10. Chin J. Estimation and projection of HIV infection and AIDS cases. *Curr Sci* 1995; *69* : 828-32.

11. Singh P. Projections on AIDS and HIV. *CARC Calling* 1993; *6* : 20-2.

12. Isham V. Estimation of the incidence of HIV infection. *Philos Trans R Soc Lond [Biol]* 1989; *325* : 113-21.

13. Day NE, Gore SM, McGee MA, South M. Predictions of the AIDS epidemic in the U.K. : The use of the back projection method. *Philos Trans R Soc Lond [Biol]* 1989; *325* : 123-34.

14. Lal S, Thakur BB. The problem of HIV and AIDS in India. *Curr Sci* 1995; *69* : 833-4.

Reprint requests : Dr A. Basu, Applied Statistics Unit, Indian Statistical Institute
203 B.T. Road, Calcutta 700035